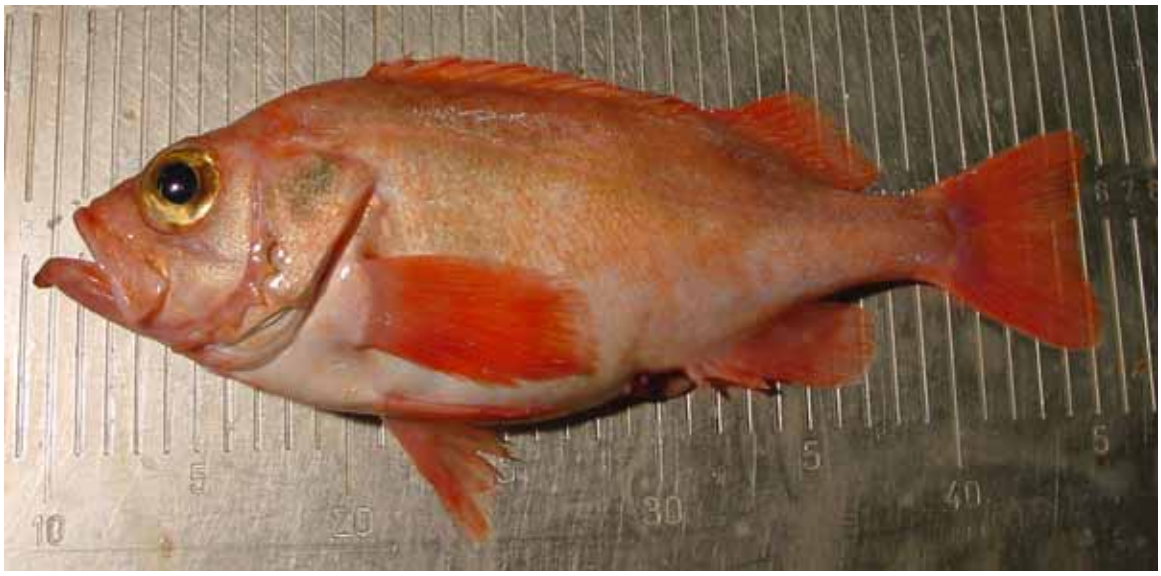
MOLECULAR GENETIC STUDIES ON SPECIES AND POPULATION STRUCTURE OF NORTH ATLANTIC REDFISH (GENUS *SEBASTES*; CUVIER 1829)



Dissertation

zur Erlangung des Doktorgrades des Fachbereiches Biologie der Universität Hamburg

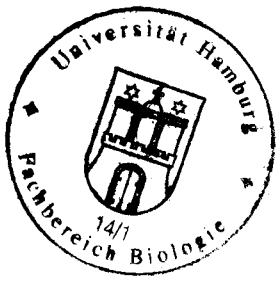
vorgelegt von

Cathrin Schmidt

Hamburg 2005

Genehmigt vom
Fachbereich Biologie der
Universität Hamburg
auf Antrag von Professor Dr. A. TEMMING
Weiterer Gutachter der Dissertation:
Professor Dr. R. HANEL
Tag der Disputation: 17. Juni 2005

Hamburg, den 02. Juni 2005



Professor Dr. Arno Frühwald
Dekan

University of Hamburg
Department of Biology

ZMK
ZENTRUM FÜR MEERES- UND KLIMAFORSCHUNG

March 9th, 2005

This is to certify that the quality of the English employed in the PhD thesis submitted by the candidate Cathrin Schmidt entitled "Molecular genetic studies on species and population structure of North Atlantic redfish (genus *Sebastes*; Cuvier 1829)" is of a standard sufficient to fulfill the language requirements of the University of Hamburg.

Prof. Dr. Michael St. John
Vice Director, Center for Marine and Climate Research
Institute for Hydrobiology and Fisheries Science
University of Hamburg
Olbersweg 24
22767 Hamburg
Germany
Tel +4940428386600
FAX +4940428386618
email michael.st.john @uni-hamburg.de

TABLE OF CONTENTS

SUMMARY

In this study, the genetic structure of species and populations of North Atlantic redfish (genus *Sebastes*; Cuvier 1829) was investigated, with special emphasis on the species complex occurring in the Irminger Sea and on the continental slopes of Greenland, Iceland and the Faroe Islands. For this purpose, DNA sequence analysis of the mitochondrial ND3 gene, microsatellite analysis and amplified fragment length polymorphism (AFLP) analysis were used to determine and assess inter- and intraspecific levels of genetic variation and genetic differentiation within and among samples of the four currently known North Atlantic *Sebastes* species *S. mentella* (Travin 1951), *S. marinus* (Linné 1758), *S. fasciatus* (Storer 1854) and *S. viviparus* (Krøyer 1845).

The results from these molecular methods provided evidence that the currently recognised species *S. mentella*, *S. fasciatus* and *S. viviparus* are genetically distinct and therefore represent valid species. Furthermore, the results of the genetic analyses revealed that samples of specimens pre-classified as *S. marinus* according to external morphological characters - collected on the continental slopes of Greenland, Iceland and the Faroe Islands - were genetically extremely heterogeneous. Findings from all methods indicated the existence of two genetically isolated groups of individuals within these samples. These two groups revealed genetic differences of the same order of magnitude as the genetic differences observed between all other North Atlantic *Sebastes* species. One of these genetically distinct groups was found in all areas studied, from Norway to the Flemish Cap, and most likely represents the species *S. marinus.* The occurrence of the second genetically distinct group of individuals suggested that another - apparently cryptic - species of *Sebastes* occurs on the continental slopes of Greenland, Iceland and the Faroe Islands. In particular, the results of the ND3 gene sequence analysis and the microsatellite analysis provided strong evidence of cryptic speciation due to (i) the prevalence of two distinct ND3 gene haplotype lineages with a relatively high level of intraspecific genetic divergence in the *S. marinus* samples from Iceland and Greenland and (ii) the occurrence of two genetically isolated groups of individuals within samples from Greenland, Iceland and the Faroe Islands, identified using microsatellite polymorphisms.

Even though the five groups - representing the four currently recognised species and the newly identified cryptic species - were genetically distinct, the results of the ND3 gene analysis revealed low levels of sequence divergence and a star-like phylogeny. In congruence with the prevalent perception about the phylogeny of this species complex, these results indicated that the North Atlantic *Sebastes* species are closely related and most likely derived from an explosive expansion after a population bottleneck or other demographic effects like founder events. Furthermore, the evolutionary scenario revealed by the phylogenetic analysis of the ND3 gene data provided support for the hypothesis that *S. mentella* is the representative

of the most basal lineage of North Atlantic *Sebastes*, which gave rise to the other North Atlantic *Sebastes* lineages - with *S. viviparus* representing the earliest split of the basal lineage.

Even though the results of the ND3 gene analysis indicated that the North Atlantic *Sebastes* species are closely related and evolved relatively recently, no evidence of recent broad-scale hybridisation between the species was found. The small number of admixed genotypes identified by microsatellite analysis indicated that there is only restricted, if any, hybridisation between *S. marinus*, *S. mentella* and the potential cryptic species in the areas off the Faroe Islands, Iceland and East and West Greenland, as well as between *S. fasciatus*, *S. marinus* and *S. mentella* on the Flemish Cap. Although no evidence of recent hybridisation between the North Atlantic *Sebastes* species was found, the observed incorporation of *S. fasciatus* specific ND3 gene haplotypes into individuals of *S. marinus* on the Flemish Cap - without any evidence of nuclear introgression - indicated that ancient introgressive hybridisation events occurred between *S. marinus* and *S. fasciatus*. Similarly, the observation of *S. mentella* specific ND3 gene haplotypes in several individuals of "giant" *S. marinus* without any evidence of nuclear introgression supports the hypothesis of ancient introgressive hybridisation events between *S. marinus* and *S. mentella*.

The results of the present study have also provided new insights into the intraspecific population structure of *S. mentella*, *S. marinus* and the newly identified potential cryptic species of North Atlantic *Sebastes*. The general patterns of population structure of these species - revealed by microsatellite polymorphism - were congruent with the weak genetic structuring usually reported for marine fish species.

Microsatellite analyses indicated population structure in *S. marinus* and *S. mentella*, but did not reveal such intraspecific structure within the newly identified cryptic species. The results indicated that there are at least three genetically distinct populations of *S. marinus* in the geographic area covered in this study, represented by the samples from (i) the Flemish Cap, (ii) Greenland and (iii) Norway, Iceland and the Faroe Islands. However, there appears to be some overlap in occurrence of the latter two populations in some areas, e.g. on the Dohrn Bank located between Iceland and Greenland and in some areas on the Icelandic shelf.

Similarly, microsatellite analysis indicated at least three separate populations of *S. mentella* in the area covered in the present study: One population occurring on the Flemish Cap, one in the central Irminger Sea and a third population distributed off East Greenland, West Greenland, the southern Irminger Sea and the NAFO areas 1F and 2J. Results from the other two analysis techniques particularly supported the genetic distinctness of *S. mentella* samples from the central Irminger Sea - consisting of relatively large and probably also old individuals - from samples from other regions throughout the North Atlantic. The observed genetic differences could either reflect differences between different cohorts (due to sweepstakes chance effects) or could be due to population separation caused by restricted adult dispersal and/or larval transport. Both hypotheses may be consistent with the findings of this study.

In this study, it was demonstrated that the combination of several molecular markers, in particular the combination of mtDNA sequences and microsatellite data, can provide new insights into the species and population structure of North Atlantic *Sebastes*. However, given the relatively small number of temporal samples combined with the longevity and mobility of all *Sebastes* species, more temporal replicates are needed before definite conclusions about the population structure of the *Sebastes* species complex in the central North Atlantic can be drawn.

GENERAL INTRODUCTION

Fisheries scientists and managers long have been concerned about population structure and the implications of subcomponents of fishery harvest, such as multiple species and different stocks (or both), for sustainable, long term management of the resource (Shaklee and Currens 2003).

A clear understanding of the species and population structure of exploited species and species complexes is essential for an adequate analysis of population dynamics and for the conservation of genetic resources – and therefore a prerequisite to a proper management of exploited species (Carvalho and Hauser 1995; Ruzzante et al. 1996). Failure to recognise the existence of species and population structure can lead to overexploitation of less reproductive species or populations and in the long term to erosion of genetic resources via depletion of some or all of the populations' spawning components (Allendorf et al. 1987; Carvalho and Hauser 1995; Ruzzante et al. 1996).

Determining the population structure of any fish species is a complex task: Numerous methods, such as data on fish abundance, geographic and temporal distributions, catch statistics, meristic and morphometric characteristics, immunological differences, population dynamics, variation in life history, parasite burdens and other biological characteristics, as well as artificial tags, have been used to gain insight into patterns of population structure and mixed-species and mixed-stock harvests (Ihssen et al. 1981; Shaklee and Currens 2003).

Molecular genetic methods have long been recognised for their usefulness in species identification and systematics, as well as in defining population structure of fish (see Ward 2000 and references therein). Therefore, during the last 10 to 15 years, fisheries biologists increasingly have used genetic methods to investigate the inter- and intraspecific structure of exploited fish species (Shaklee and Currens 2003).

Even though North Atlantic redfish (genus *Sebastes*; Cuvier 1829) represents an important component of the commercial fishery, our knowledge of the species and population structure of the genus is limited. During the last decades, several North Atlantic redfish stocks have severely declined due to overfishing and are considered to be outside safe biological limits (ICES 2002; 2004c).

The central topic of this thesis is the application of molecular markers for the investigation of the genetic structure of species and populations of North Atlantic redfish (genus *Sebastes*), with special emphasis on the species complex in the Irminger Sea and on the continental slopes of Greenland, Iceland and the Faroe Islands.

In the following, I will introduce the species and populations studied and will give an overview of the different kinds of molecular markers applied. The main aims and outlines of the thesis will be summarised.

An introduction to redfish

The genus Sebastes (Cuvier 1829)

Redfishes of the genus *Sebastes* (family Scorpaenidae, order Scorpaeniformes; Cuvier 1829) are the most species-rich group of scorpaenid fishes (Rocha-Olivares et al. 1999a; Love et al. 2002). Currently about 102 species of the genus *Sebastes* are recognised worldwide.
All species of the genus *Sebastes* are found in the marine environment and characterised by an antitropical distribution (Rocha-Olivares et al. 1999a).
While almost all of the *Sebastes* species are limited to the North Pacific and Gulf of California, only four occur in the North Atlantic and two in the Southern Hemisphere (Love et al. 2002).
These four North Atlantic species are *Sebastes mentella* (deep-water redfish, ocean perch; Travin 1951), *S. marinus* (golden redfish; Linné 1758), *S. fasciatus* (rosefish, Acadian redfish; Storer 1854) and *S. viviparus* (small redfish; Krøyer 1845).
Except *S. viviparus*, all North Atlantic redfish species are commercially important fishery resources.

Distribution and habitat of North Atlantic Sebastes

S. mentella and *S. marinus* are widely distributed throughout the North Atlantic, while *S. viviparus* and *S. fasciatus* are essentially restricted to the Northeast Atlantic and to the Northwest Atlantic, respectively (Barsukov et al. 1990). *S. mentella* is practically found in all regions of *Sebastes* occurrence throughout the North Atlantic, from the Barents Sea, along the Norwegian coast, around the Faroe Islands, Iceland, Greenland, the Irminger Sea and the Atlantic waters outside Canada and the USA except the North Sea and the Gulf of Maine (Barsukov et al. 1985). *S. fasciatus* is distributed from the Gulf of Maine and along North America to the Scotian Shelf and the Grand Banks (ICES 1998a; Roques et al. 2000). *S. marinus* is common in the southern part of the Barents Sea, off the coast of Norway, Iceland and Greenland. It is more rare than *S. fasciatus* and *S. mentella* off the coast of North America, but it has been reported to occur more often on the Northern and Great Newfoundland Banks, the Flemish Cap and the Gulf of St. Lawrence (Barsukov et al. 1985). *S. viviparus* is found in the Northeast Atlantic from the Barents Sea to the west of Iceland, but only rarely off the coast of Greenland. Little information exists about how far east the distribution of S. *fasciatus* extends and if the limits of dispersal of *S. viviparus* end around Iceland or South Greenland (Barsukov et al. 1985; Whitehead et al. 1986; ICES 1998a).
The North Atlantic *Sebastes* species are generally categorised as demersal, occurring along the slopes of fishing banks and deep water channels (Gascon 2003), but *S. mentella* is also found pelagically in the Irminger Sea (Magnússon and Magnússon 1995).

The four species have different depth preferences, but overlap in their depth distribution (Barsukov et al. 1985). *S. marinus* mainly inhabits depths of less than 300-370 m (sometimes it is found at depths of 700-750 m), while *S. mentella* has its main distribution between 200 m and 800 m, except in the Irminger Sea where it occurs already at depths of about 50 m. *S. fasciatus* occurs at depths between 150 m and 300 m (Roques et al. 2000). *S. viviparus* is found in shallower waters than the other three species, close to the shore, between 10 m and 120 m depth, although some individuals have been found in depths between 600-700 m (Barsukov et al. 1985; Whitehead et al. 1986; Magnússon and Magnússon 1995).

Life history and reproduction

The genus *Sebastes* shows several characteristic biological features that differ considerably from those of other commercially important fish species (Gascon 2003).
Redfishes are long-lived: Specimens of *S. mentella* have been aged to at least 75 years (Campana et al. 1990). They have slow growth rates and mature late. Maturity is attained at about 55% of the maximum length of the species (Wourms 1991). The maximum reported lengths of North Atlantic *Sebastes* range from 35 cm for the smallest species *S. viviparus* to 100 cm for the largest North Atlantic species *S. marinus* (Whitehead et al. 1986).
Age at maturity for North Atlantic *Sebastes* ranges from 5 years (*S. fasciatus*) to 15 years (*S. marinus*), depending on the species (ICES 1998a; Gascon 2003).
On average, commercial size is reached at about 25 cm (Gascon 2003).
Redfishes are viviparous with internal fertilisation. Generally, the viviparity is considered to be lecithotrophic (larvae feeding exclusively on energy stored in the yolk), but additional matrotrophy is assumed (Wourms 1991; Love et al. 2002). In the North Atlantic, mating takes place in autumn and the females carry the developing larvae until spring when they are released from April to July (Gascon 2003).
The females seek certain areas for the extrusion of the larvae and males and females are frequently separated during that period (Magnússon and Magnússon 1995). Redfish are highly fecund: For example, it is reported that females of *S. marinus* release between 50,000 and 350,000 larvae (Whitehead et al. 1986; Muus and Nielsen 1999).
All species of *Sebastes* have pelagic larvae that metamorphose into pelagic juveniles. After several months, the pelagic juveniles migrate vertically and move to the nursery areas (Kelly and Barker 1961; Wourms 1991).

Species structure and species identification of North Atlantic Sebastes

The four North Atlantic redfish species are morphologically remarkably similar and show in many regions widely overlapping meristic and morphological characters. This makes it

difficult to distinguish them consistently in the field (McGlade et al. 1983; Barsukov et al. 1985).

Over the last few decades, numerous characters have been studied to examine the differences among the four North Atlantic *Sebastes* species, to assess their taxonomic status, to study their relationship and to find characters useful for species discrimination. Differences among species were observed in meristic and morphometric characters such as the shape of the symphysial tubercle (e.g. Barsukov 1973; Barsukov et al. 1985; Power and Ni 1985), the number of soft rays in the anal fin (AFC = anal fin ray count: Ni 1981), the development and structure of the gas bladder muscles and their attachment to the axial skeleton (EGM = extrinsic gas bladder muscle: Ni 1981; Power and Ni 1982; 1985), otolith shape (Stransky and MacLellan 2005), larval characteristics (Templeman and Sandeman 1959; Templeman 1980), parasite infestation (Yanulov 1962a; 1962b), immunological (O'Rourke 1961) and biochemical characters such as the thermostability of isolated muscle tissue (Althukov et al. 1968) and the fatty acid profiles of heart and gill tissue and skull and otolith oil (Joensen and Grahl-Nielsen 2001).

Several studies have used molecular markers for species identification and to investigate the relationships among North Atlantic *Sebastes* species. These studies have been dominated by investigations on the electrophoretic mobility patterns of proteins, such as blood serum proteins (Altukhov and Nefyodov 1968), serum haptoglobin types (Nefyodov 1971), hemoglobins (Nævdal 1978; Nedreaas and Nævdal 1989; 1991a; 1991b; Nedreaas et al. 1994; Johansen et al. 1998), various enzymes (Payne and Ni 1982; McGlade et al. 1983; Trottier et al. 1989; Nedreaas and Nævdal 1989; 1991a; 1991b; Rubec et al. 1991; Nedreaas et al. 1994; Johansen et al. 1998; Johansen 2003) and sarcoplasmic proteins (Rehbein 1983).

DNA-based studies of North Atlantic *Sebastes* species structure included the analysis of mitochondrial DNA, such as tandem repeat polymorphisms in the control region, restriction fragment length polymorphism (RFLP) analysis of mtDNA (Bentzen et al. 1998) and sequence analysis of the mitochondrial gene coding for 16S rRNA (Sundt and Johansen 1998), as well as randomly amplified polymorphic DNA (RAPD) analysis (Johansen 2003), RFLP analysis of nuclear DNA (Desrosiers et al. 1999) and microsatellite analyses (Roques et al. 1999a; 1999b; 2001).

In general, low levels of inter- and intraspecific genetic variation have been observed with proteic markers: Differences between the taxa were mainly based on differences in allele frequencies and only few fixed differences were observed (Payne and Ni 1982; Nedreaas and Nævdal 1989; 1991a; Rubec et al. 1991; Johansen et al. 1998).

Moreover, most DNA-based studies did not show higher levels of genetic differentiation (Sundt and Johansen 1998; Desrosiers et al. 1999; Johansen 2003).

Exceptions were the microsatellite studies by Roques et al. (1999a; 1999b; 2001), which revealed higher levels of genetic differentiation. Roques et al. (1999a) demonstrated that microsatellite data from eight loci could be used to assign individual fish to the correct

species more precisely than other genetic techniques. As yet, the method has not been routinely applied in field surveys.

Therefore, a combination of several morphological and proteic characters, such as the extrinsic gas bladder muscle (EGM) rib passage patterns, the anal fin ray count (AFC), hemoglobin (Hb) and malate dehydrogenase (MDH) electrophoretic mobility patterns, has still to be used to confidently distinguish the taxa in adult life stages, but these methods are time-consuming and, in contrast to DNA-based methods, often not applicable to young life history stages (Nedreaas and Nævdal 1991a; Rubec et al. 1991; Gascon 2003).

Due to the serious difficulties in species identification, redfish in the North Atlantic is still exploited and managed as one species (Barsukov et al. 1985; Gascon 2003).

The low levels of genetic differentiation between the North Atlantic redfish species have not only hampered the unambiguous species identification, but also our understanding of the evolutionary relationship of the four species. Even though several authors proposed hypotheses on the evolutionary relationship of the four species (Barsukov and Zakharov 1973; Roques et al. 1999a; reviewed in Kendall 2000), the results of the different genetic methods are still ambiguous (e.g. Payne and Ni 1982; Nedreaas and Nævdal 1989; Nedreaas and Nævdal 1991a; Roques et al. 1999a; Johansen 2003).

Nevertheless, the resolution of the phylogenetic relationships among these closely related species is fundamental to our understanding of their evolution and radiation of biological diversity.

Intermediate individuals and hybridisation

Discrimination between redfish species is even more complicated when the possibility of hybridisation between species is considered, e.g. in areas of sympatric occurrence (Desrosiers et al. 1999).

In the western North Atlantic, an overlap in characters used to differentiate *S. mentella* and *S. fasciatus* has been observed in areas where the two species occur sympatrically, for example in the Gulf of St. Lawrence and adjacent waters (Desrosiers et al. 1999). The results of several genetic studies indicate that broad-scale introgressive hybridisation between *S. mentella* and *S. fasciatus* occurs in these areas (Rubec et al. 1991; Desrosiers et al. 1999; Sevigny et al. 2000; Roques et al. 2001). At present it is not known whether broad-scale hybridisation occurs in other areas than the Gulf of St. Lawrence and whether other species than *S. mentella* and *S. fasciatus* are involved.

However, several authors also mention fish that possess characters of both *S. mentella* and *S. marinus* in a varying manner (e.g. Kotthaus 1961; Nævdal 1978). Apart from the overlap in morphological characters, an overlap in electrophoretic mobility patterns of hemoglobins, blood serum proteins and enzymes has been observed in individuals of *S. mentella* and *S. marinus*, for example in Icelandic waters and off East and West Greenland (Altukhov and

Nefyodov 1968; Nedreaas and Nævdal 1991a; Nedreaas et al. 1994). Therefore, hybridisation between *S. marinus* and *S. mentella* in these areas cannot be ruled out (Altukhov and Nefyodov 1968; Nedreaas et al. 1994).

Population structure of commercially exploited redfish in the Irminger Sea and adjacent waters

The outlined difficulties in species discrimination have also hampered our understanding of the population structure of *Sebastes*. Consequently, little is known about the population structure of the commercially exploited species *S. marinus* and *S. mentella* in the central and eastern North Atlantic.

In the central and eastern North Atlantic *S. mentella* is currently separated into three stock components, *S. mentella* in the Northeast Arctic, "deep-sea" (or demersal) *S. mentella* on the continental slopes of Greenland, Iceland and the Faroe Islands and "oceanic" (or pelagic) *S. mentella* occurring pelagically in the Irminger Sea in depths between 50 m and 1000 m (ICES 1992; 1998a; Magnússon and Magnússon 1995).

The population structure of *S. mentella* in the Irminger Sea and on the continental slopes of Greenland, Iceland and the Faroe Islands has been discussed extensively elsewhere (e.g. ICES 1998a; Saborido-Rey et al. 2005), but is briefly summarised below.

Historically *S. mentella* was fished on the shelves and banks of the Faroe Islands, Iceland and East Greenland and was regarded as one stock, but with the start of a new pelagic fishery in the open Irminger Sea a new stock for *S. mentella* was defined (Saborido-Rey et al. 2005). In 1992, the ICES Study Group on Redfish Stocks decided to define the two types as "deep-sea" *S. mentella* and "oceanic" *S. mentella* (ICES 1992).

In the early 1990's, the pelagic fishery in the Irminger Sea started fishing in oceanic layers deeper than 500 m and some researchers considered that some of the redfish caught below 500 m were different from those living above 500 m but resembled more the demersal *S. mentella* living on the shelves (Magnússon and Magnússon 1995; ICES 1998a; reviewed in Saborido-Rey et al. 2005). This new type of *S. mentella* living below 500 m has been called "pelagic deep-sea" *S. mentella* (ICES 1998a).

The reported distinction of the *S. mentella* types in the Irminger Sea has been based on several criteria, such as differences in colour, length-weight relationships, length-at-maturity, infestation with the parasitic copepod *Sphyrion lumpi* and differences in the size of newly extruded larvae (Magnússon and Magnússon 1995).

There has been a large controversy about whether the different *S. mentella* types are more than one stock and different hypotheses have been put forward to describe the population structure of the species in these areas (ICES 1998a; Saborido-Rey et al. 2005):

The "single stock hypothesis" suggests that the mature individuals of a single stock - distributed from the Faroe Islands to Greenland - segregate according to age/size.

The "two stock hypothesis" proposes that the demersal *S. mentella* living on the shelves (so-called "deep-sea" *S. mentella*) and the *S. mentella* occurring in deeper pelagic waters in the Irminger Sea (so-called "pelagic deep-sea" *S. mentella*) constitute one stock unit, which is separated from the "oceanic" *S. mentella* occurring pelagic in the Irminger Sea.
According to the "three stock hypothesis" each of these three types constitutes a distinct stock. Several non-genetic and genetic studies - such as morphometric and meristic analyses (Reinert and Lastein 1992; Garabana 2005), otolith shape analyses (Stransky 2005), analyses of otolith microchemistry (Stransky et al. 2005), electrophoretic mobility patterns of hemoglobin and allozymes (Dushchenko 1987; Nedreaas and Nævdal 1989; Nedreaas and Nævdal 1991a; Nedreaas et al. 1994; Johansen et al. 2000b) and microsatellites (Roques et al. 2002) - have investigated the population structure of *S. mentella* in the North Atlantic. However, the interpretations of the population structure in the Irminger Sea and adjacent waters are still diverging. In addition, questions about the general ecology of the species, such as the location of spawning (= larval release) and nursery areas, are still discussed controversially (Magnússon and Magnússon 1995; Saborido-Rey et al. 2001; 2005). Therefore, the stock structure of *S. mentella* remains uncertain (ICES 2002).
The knowledge on the population structure of *S. marinus* is also limited and remains uncertain. *S. marinus* occurring in the central and eastern North Atlantic has been separated into two stock components, *S. marinus* in the Northeast Arctic (ICES subareas I and II) and *S. marinus* on the continental slopes of Greenland, Iceland and the Faroe Islands (ICES Divisions V, VI, XII, XIV and NAFO area 1; ICES 1998a). However, some studies based on meristic and morphometric measurements (Reinert and Lastein 1992), as well as on the electrophoretic mobility patterns of hemoglobin and on allozymes (Nedreaas and Nævdal 1991a; Nedreeas et al. 1994; Johansen et al. 2000a), indicate a more complex population structure within the latter stock component.
In addition, several authors have described a group of redfish with an average total length of above 60 cm, which are separated from the normal range of *S. marinus* and *S. mentella* length curves (Kotthaus 1961; Altukhov and Nefyodov 1968; Johansen et al. 2000a). These so-called "giant" redfish (or "giant" *S. marinus*) occurred on the shelves of Greenland and Iceland and along the Reykjanes Ridge in depths between 500 m and 800 m (ICES 1998a) and were morphologically identified as *S. marinus*, but showed some differences in the shape and number of gill rakers and in the internal structure of the otoliths when compared with *S. marinus* and *S. mentella* (Kotthaus 1961). The status of "giant" *S. marinus* is controversial: Barsukov (1973) proposed that the distinguishing features could be explained by processes of intraspecific variation with age, whereas genetic analysis could indicate that the "giants" are hybrids of *S. mentella* and *S. marinus* (Altukhov and Nefyodov 1968) or even a genetically isolated group of *S. marinus* (Johansen et al. 2000a). At present, the taxonomic status of "giant" redfish remains unclear (Johansen et al. 2000a).

The possible relationships between the different redfish stocks in the Irminger Sea and along the continental slopes of Greenland, Iceland and the Faroe Islands are illustrated in Figure 1. In summary, our knowledge about the species and population structure of *Sebastes* inhabiting the shelves and continental slopes off Greenland, Iceland, Faeroe Islands is still limited. However, the identification of species as well as populations is essential for a proper management of the genus (Gascon 2003).

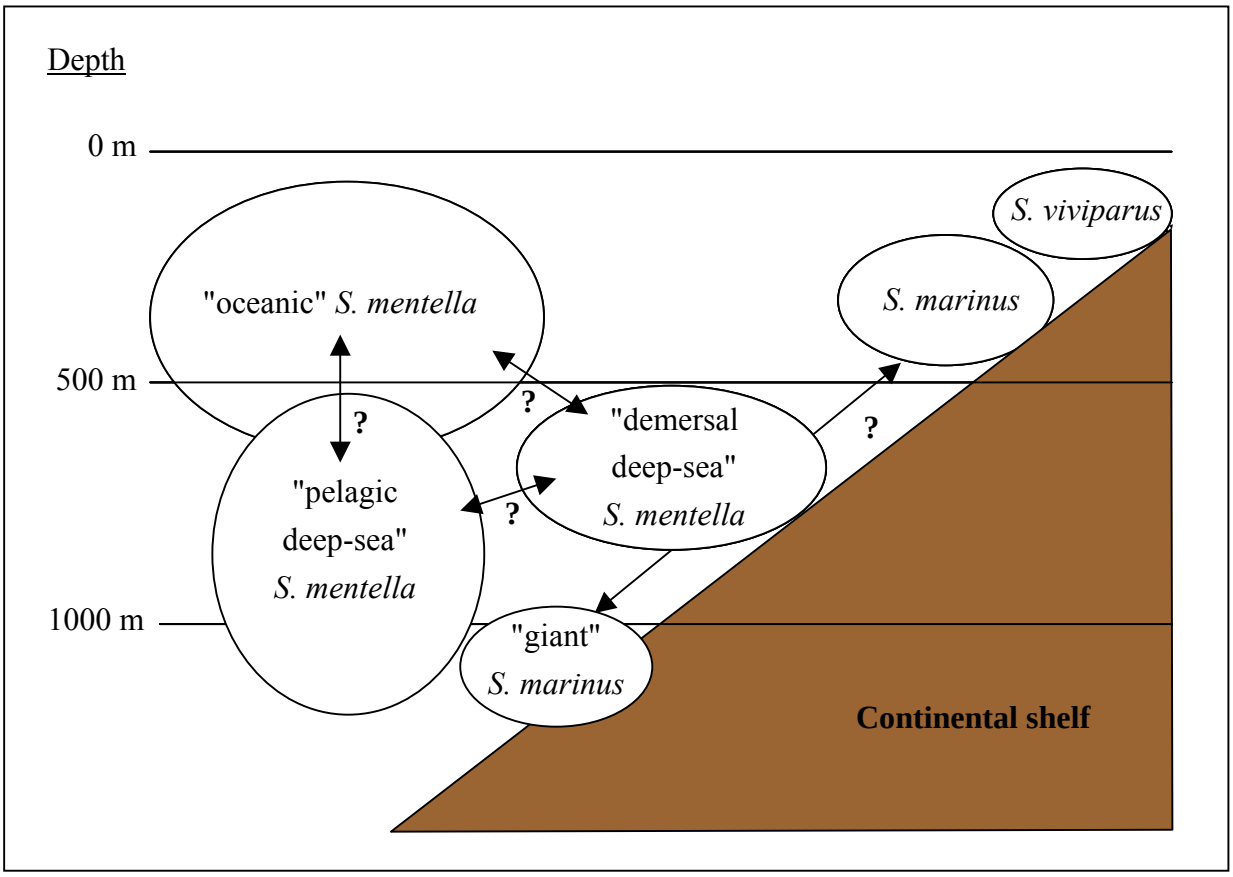


Figure 1. Possible relationships between the different redfish stocks in the Irminger Sea and along the continental slopes of Greenland, Iceland and the Faroe Islands (modified from: ICES 1998a).

Molecular markers

Molecular genetic methods have proved invaluable in distinguishing between morphologically very similar species, in revealing previously unrecognised cryptic sibling species, but also in defining population structure of fish (see Ward and Grewe 1995 and references therein).

The choice of a particular molecular method for detecting species and population structure must account for the geographical and temporal scales on which evolutionary processes operate within and between species (Grant et al. 1999). Therefore, some methods are more suited than others to detect structure on different scales (Grant et al. 1999).

In redfish, proteic markers, as well as some mitochondrial markers, generally revealed low levels of genetic variability, both on species and on population level (Payne and Ni 1982; McGlade et al. 1983; Rehbein 1983; Trottier et al. 1989; Nedreaas and Nævdal 1989; 1991a; 1991b; Rubec et al. 1991; Nedreaas et al. 1994; Bentzen et al. 1998; Johansen et al. 1998; Sundt and Johansen 1998; Johansen 2003).

Therefore, there was a need for genetic markers with a higher resolution and the ability to detect genetic differences at species and at population level, and at every developmental stage (as hemoglobin and several allozyme markers are often not applicable to young life history stages; Nedreaas and Nævdal 1991a).

During the last decade, the development of DNA-based molecular methods produced new marker systems with the potential to overcome the limitations of hemoglobin and allozyme analysis in North Atlantic redfish, as already shown by Roques et al. (1999a; 1999b; 2000; 2001; 2002), as well as in numerous studies for Pacific *Sebastes* species (e.g. Rocha-Olivares and Vetter 1999; Rocha-Olivares et al. 1999a; 1999b; 1999c; 1999d; 2003; Miller et al. 2000; Buonaccorsi et al. 2002; Kai et al. 2002a; 2002b; 2003; Cope 2004; Matala et al. 2004a and 2004b).

In this study, I investigated the species and population structure of North Atlantic redfish (genus *Sebastes*) using three different DNA-based methods:

One method is based on mitochondrial DNA sequence variations, two methods are based on nuclear markers, namely microsatellite polymorphism and amplified fragment length polymorphism (AFLP). All three methods are PCR-based (polymerase chain reaction; Mullis et al. 1986) techniques.

Polymerase chain reaction (PCR)

A major development that revolutionised the use of DNA markers in population and phylogenetic studies is the polymerase chain reaction (PCR; Mullis et al. 1986). The PCR is a technique used to amplify a specific segment of DNA that lies between two regions of known sequence (Mullis et al. 1986; Mullis and Falcoona 1987; Saiki et al. 1988). Two

oligonucleotides, usually about 20 nucleotides in length and complementary to the flanking sequences of the target DNA, are used as primers for a series of reactions that are catalysed by a heat stable DNA polymerase. Only a minute amount of tissue is needed as the target DNA is amplified many thousand-fold from the small amount initially available, permitting non-invasive and non-lethal sampling, or the analysis of museum specimens and scales (e.g. Nielsen et al. 1997).
The particular value of the PCR in relation to fisheries is that - once appropriate primers are available - large numbers of individuals can be assayed quickly, thereby facilitating the comprehensive screening of population variability (Kocher et al. 1989; Ward 2002).

Mitochondrial DNA (mtDNA)

Most of the DNA in eukaryotic cells is found within the nucleus, but a small proportion (< 1%) is non-nuclear and located in the organelles (mitochondria, and - in plants - chloroplasts). The mitochondrial (mt) genome is a single small, double stranded, circular DNA molecule contained in multiple copies in mitochondria. Typically, the size of animal mitochondrial genomes is about 16,000 to 18,000 base pairs (bp). It contains 13 genes coding for proteins, two genes coding for ribosomal RNAs, 22 genes coding for transfer RNAs and one major non-coding region that contains the initiation sites for mtDNA replication and RNA transcription (Meyer 1993; 1994). The two strands of the mitochondrial genome are designated light (L) and heavy (H). These names reflect differences in the G + T content. With few exceptions, all genes in the vertebrate mitochondrial genome are encoded by the H-strand (Meyer 1994).
Generally, fish mtDNA is assumed to be homoplasmic (all molecules in an organism are identical; Billington 2003), even though cases of heteroplasmy have been reported (e.g. Hoarau et al. 2002a).
As mitochondrial genomes are haploid and apparently non-recombining, the mtDNA molecule is effectively one single locus, with the composite genotypes equivalent to alleles, in contrast to microsatellite or allozyme markers (Stepien and Kocher 1997).
Owing in particular to two features, mitochondrial DNA has been the most widely studied region of eukaryotic genomes and has played a critical role in the development of population and evolutionary genetics (Avise et al. 1987; Avise 1998; Rand 2001a; Rand 1994):
Firstly, mtDNA is maternally inherited. Therefore, the effective population size of mtDNA is approximately only ¼ that of nuclear loci, when males and females are equally frequent, which increases the effect of genetic drift and results in stronger genetic differentiation as compared to nuclear DNA (Birky et al. 1983; Birky et al. 1989).
Secondly, mtDNA evolves more rapidly than nuclear DNA. The rates of substitutions in mtDNA are typically between five to ten times greater than in single copy nuclear genes,

averaging approximately 2% sequence divergence per million years (Meyer 1994; Mindell and Thacker 1996).

These properties of mtDNA make it a useful genetic marker for molecular systematics, for examination of phylogenetic relationships among closely related species (Thomas and Beckenbach 1989; Avise 1994; Stepien and Kocher 1997), as well as for population genetic studies (Avise 1998).

Several studies successfully used mitochondrial gene sequence data, mainly of the cytochrome b gene and the control region, to study the molecular evolution, systematics and the population structure of Pacific and South Atlantic *Sebastes* species and also applied the method to juvenile identification (Rocha-Olivares and Vetter 1999; Rocha-Olivares et al. 1999a; 1999b; 1999c; 1999d; Kai et al. 2002a; 2002b; Cope 2004).

As the four North Atlantic *Sebastes* species represent a group of very closely related species, a gene with a relatively high rate of evolution is needed to determine the species structure and the phylogenetic relationships among the taxa.

Therefore, the mitochondrial ND3 gene - coding for the NADH dehydrogenase subunit 3 (an enzyme subunit involved in the electron transport chain) - was chosen to investigate the species structure and the phylogenetic relationships of the North Atlantic *Sebastes* species, as it has a higher substitution rate than cytochrome b and most other nuclear and mitochondrial genes (Thomas and Beckenbach 1989; Domanico and Phillips 1995). The NADH dehydrogenase subunit 3 has a length of 351 base pairs and is located between the genes coding for the transfer RNAs for Arginin and Glycin ($tRNA^{Arg}$ and $tRNA^{Gly}$).

ND3 gene sequence data have been successfully applied to investigate the phylogenetic relationships among the Salmoninae (Phillips and Oakley 1997), Pacific salmon and trout species (genus *Oncorhynchus*; Thomas and Beckenbach 1989; Domanico and Phillips 1995; Domanico et al. 1997; McKay et al. 1996; Trautner 2000), as well as among species of lamprey (genus *Lampetra*; Docker et al. 1999).

Hence, it is probably one of the most appropriate regions of the mtDNA for species identification and for the examination of the phylogenetic relationships among North Atlantic *Sebastes* species.

Microsatellites

Microsatellites are co-dominant DNA markers, which are common and widespread in most eukaryotic nuclear genomes (Jarne and Lagoda 1996; Estoup and Angers 1998).

They consist of short segments of nuclear DNA in which a specific motif of 1-6 bases is tandemly repeated. Microsatellites are usually classified into three types, perfect microsatellites, which are composed of non-interrupted sequences of repeat units, imperfect microsatellite loci, which have a few interrupting bases in an otherwise perfect repeat array

and compound microsatellites, consisting of neighbouring sequences of different repeat types (Estoup and Angers 1998).

The total length of microsatellite arrays is generally quite short, usually less than 300 base pairs (Tautz 1989). The regions flanking the microsatellite are generally non-repeat sequences that are conserved among closely related species (Estoup and Angers 1998).

Microsatellites are non-coding and thus likely to be selectively neutral (Jarne and Lagoda 1996; Morris et al. 1996). As their mutation rate is very high, with values between 10^{-5} and 10^{-2} mutations per generation, they often show high levels of polymorphism, the variation arising from differences in the number of repetitive sequences (Jarne and Lagoda 1996; Estoup and Angers 1998). The microsatellite polymorphism is mainly caused by replication slippage, which occurs when the new strand mispairs with the template strand during the DNA replication (Schlötterer and Tautz 1992).

Traditionally, two mutation models have been considered for microsatellites, the infinite allele model (IAM: Kimura and Crow 1964) and the stepwise mutation model (SMM: Kimura and Ohta 1978). Under the IAM, a mutation involves any number of tandem repeats and always results in an allelic state not previously encountered in the population (Estoup and Angers 1998). The SMM predicts that mutation will result in an allele that is one repeat larger or smaller, with decreases and increases being equally likely. Therefore, alleles can mutate towards allele states already present in the population. More recently a model has been developed that more accurately explains the variation observed at dinucleotide repeat microsatellites (Di Rienzo et al. 1994). This model, the two phase model (TPM), incorporates the mutational process of the SMM, but also allows for mutation steps of several repeats. Most studies indicate that the TPM is the most realistic mutation model among the three models described above, and that microsatellites indeed mutate in a stepwise fashion at a relatively high rate (reviewed in Estoup and Angers 1998 and in Schlötterer 2000). Nevertheless, mutation rate and model can differ among the different types of microsatellites (see above), mutation rates can differ for different alleles at a locus, and microsatellites do not mutate into infinite lengths (Estoup and Angers 1998). Therefore, the mutational processes are still not fully understood (Jarne and Lagoda 1996; Goldstein and Pollock 1997). Deviations from the models have consequences for the selection of appropriate statistics that make specific assumptions about mutation models (reviewed by Goldstein and Pollock 1997).

Since the early to mid-1990s, microsatellites have gradually replaced allozymes and other molecular markers as the tool of choice for population genetic studies, due to several reasons (reviewed in Hansen 2003):

Firstly, microsatellite analysis is relatively simple. Due to their short length, microsatellite loci can easily be amplified using the polymerase chain reaction (PCR), if the sequences of the conserved flanking regions are known. This makes them useful for studies of DNA even from fossils, otoliths, old scales etc. (Nielsen et al. 1997; see also Luikart and England 1999).

The length of the amplified PCR products can subsequently be determined by analysis on a denaturing polyacrylamide gel in an automated sequencer.
Secondly, the codominant mode of inheritance in combination with the high levels of variability observed in microsatellites - compared for example to allozymes - results in new, powerful applications. Recently, a number of new statistical tools were developed, which allow to infer the number of populations in a given data set and the precise assignment of individuals to a set of baseline populations. Additionally, the genetic contribution of these baseline populations to the genotype of an individual, i.e. individual admixture proportions, can be estimated (e.g. Cornuet et al. 1999; Pritchard et al. 2000; Falush et al. 2003). These methods can be used to detect hybrid-zones (Nielsen et al. 2003; 2004) and allow for the assignment of individuals to species and stocks (Schmidt 2000; Nielsen et al. 2001).
In general, microsatellites are useful for detecting population structure in fish species with low levels of allozyme or mtDNA variation, e.g. in species where gene flow or recency of isolation has limited divergence (e.g. Bentzen et al. 1996; Ruzzante et al. 1999; 2000). Microsatellites are therefore potentially capable of detecting genetic structure on small spatial scales and over short periods of time (Grant et al. 1999).
Therefore, they are the molecular markers of choice for detecting the inter- and intraspecific genetic structure of North Atlantic *Sebastes*, according to previous studies on Northwest Atlantic redfish (Roques et al. 1999a; 1999b; 2000; 2001; 2002) and Pacific *Sebastes* populations (Miller et al. 2000; Buonaccorsi et al. 2002; Rocha-Olivares et al. 2003; Matala et al. 2004a and 2004b).

Amplified fragment length polymorphism (AFLP)

The AFLP method is a relatively new molecular tool combining the reliability and reproducibility of RFLP (restriction fragment length polymorphism) analyses with the advantage of RAPD (randomly amplified polymorphic DNA) analyses of screening a high number of loci in a single PCR reaction (Liu et al. 1998). The technique, first described by Vos et al. (1995), is a powerful random fingerprinting technique, which was originally developed to allow the construction of genetic maps for application in genome research and positional cloning of genes (Vos and Kuiper 1997).
The AFLP procedure is based on the detection of DNA restriction fragments by PCR amplification (Zabeau and Vos 1993; Vos et al. 1995). It consists of four steps - restriction, ligation of adapters, pre-amplification and selective amplification - outlined briefly below.
Firstly, DNA is cut by two different restriction enzymes (e.g. a six-base cutter and a four-base cutter). Then double-stranded adapter sequences are ligated to the end of the restriction sites, which can subsequently serve as universal binding sites for primer annealing in PCR. In this way, restriction fragments of a particular DNA can be amplified with "universal" AFLP primers corresponding to the restriction site and adapter sequences (Vos and Kuiper 1997).

In the two subsequent amplification steps (pre-amplification and selective amplification), subsets of restriction fragments are amplified using selective AFLP primers. Hence, only those fragments are amplified, in which the primer extensions match the nucleotides flanking the restriction site. This reduces the complexity of the restriction fragment mixture, as for most genomes the number of fragments that will be simultaneously detected in this way will be too high to be resolved in any fragment analysis system, such as polyacrylamide gels. Consequently, the number of fragments to be amplified can be "tuned" by the selection of the number of selective bases in the AFLP primers (Vos and Kuiper 1997). Typically, 50-200 bands are generated per individual after electrophoresis of the PCR amplified products on an analytical polyacrylamide gel. These restriction fragment patterns are called AFLP fingerprints, which are a rich source for restriction fragment length polymorphisms.

These polymorphisms are caused by different types of alterations in the DNA sequence: Mutations abolishing the restriction enzyme target site, mutations creating new target sites, insertions, deletions or inversions between two restriction sites (Vos and Kuiper 1997).

The AFLP technique assays the entire genome for polymorphic markers and requires relatively small amounts of genomic DNA. It provides 10-100 times more markers and is thus more sensitive than other fingerprinting techniques (e.g. allozymes and RFLP analysis: Lu et al. 1996; Sharma et al. 1996).

The large number of AFLP markers gives an estimate of variation across the entire genome, thus giving a good general picture of the level of genetic variation (Zabeau and Vos 1993; Vos et al. 1995). The AFLP data usually must be treated as dominant markers: At a particular locus, the plus-allele (the presence of a band) dominates over the null-allele (unamplifiable with PCR). Therefore, heterozygous genotypes with a given band cannot be directly distinguished from homozygotes with the band.

The AFLP technique differs significantly from other random fingerprinting techniques such as RAPD by its robustness and reproducibility (Zabeau and Vos 1993; Vos et al. 1995). Unlike in microsatellite analysis, no taxon-specific primer sets are required. Therefore, the AFLP technique can be applied to a wide variety of organisms with no prior sequence information. Thus AFLP analysis allows a relatively quick development of markers, which is often important in conservation and endangered species management (Gaudeul et al. 2000).

In recent years AFLP analysis is quickly becoming the tool of choice for many applications and organisms. Potential applications in fish (as well as in other species) include genetic mapping (Ransom and Zon 1999; Nakamura et al. 2001), parentage analysis (Liu et al. 1998), DNA sex identification (Griffith et al. 2000), estimation of genetic diversity in cultivated, natural and rare populations (e.g. Gaudeul et al. 2000; Trautner 2000; Micket et al. 2003), characterisation of hybridisation and introgression (Congiu et al. 2001; Young et al. 2001), phylogenetic studies of fish species in rapidly evolving systems (Albertson et al. 1999), species and population assignment (Campbell et al. 2003) and population genetics (Seki et al. 1999).

A recent study demonstrated that AFLP analysis might be an excellent alternative to microsatellite analysis in order to enhance resolution in studies of population assignment, especially when population differentiation is weak (Campbell et al. 2003).
Also in redfish, the AFLP technique was successfully applied to study genetic differentiation: Kai et al. (2002a) used the AFLP technique to identify diagnostic markers for three colour morphotypes of the black rockfish (*Sebastes inermis*).
Therefore, the AFLP technique could be a useful tool to detect inter- and intraspecific genetic structure in North Atlantic *Sebastes*.

Aims and outline

A detailed knowledge of the species structure and differentiation of the North Atlantic redfish species complex, the evolutionary relationships among the species and their intraspecific population structure is relevant and necessary for the protection and management of these commercially important species.
This thesis consists of three chapters, focussing on two main objectives:
The first objective was to gain insight into the genetic structure of the North Atlantic redfish species as well as into their evolutionary relationships and to determine the most reliable and applicable method for species identification, using the molecular markers described above.
The second objective was to investigate the genetic population structure of the commercially important species *S. marinus* and *S. mentella,* with special emphasis on the species complex in the Irminger Sea and on the continental slopes of Greenland, Iceland and the Faroe Islands, and to find a powerful method for detection of population structure and for population identification.

In the study presented in **Chapter 1**, the genetic variability in the mitochondrial ND3 gene was studied to investigate (i) the species structure of the North Atlantic *Sebastes* species and their morphological types and (ii) the phylogenetic relationships among the North Atlantic *Sebastes* species in comparison to a North Pacific and a South Atlantic *Sebastes* species as well as a species of the related genus *Helicolenus*.

In the study presented in **Chapter 2**, eight microsatellite loci were used to investigate the species structure of the four North Atlantic redfish species and the possibility of hybridisation between them. The population structure within *S. mentella* and *S. marinus* in the central North Atlantic (in the Irminger Sea and on the continental slopes of Greenland, Iceland and the Faroe Islands) was studied, with special emphasis on *S. marinus*. The use of microsatellites for the assignment of single individuals to species and populations was evaluated.

In the study presented in **Chapter 3**, the genetic variability at AFLP loci was studied for all four species of North Atlantic *Sebastes*. The AFLP technique was evaluated with respect to its usefulness in species identification and its ability to detect genetic population structure in *S. mentella* in the Irminger Sea, Iceland and Greenland area.

In the **Final Discussion** chapter, the results of these three studies were compared and the three methods were evaluated with respect to their use in redfish genetics. Furthermore, the findings were set in relation to the ecological and life history features of the species complex. Different hypotheses were put forward to explain the inter- and intraspecific structure of the North Atlantic redfish species complex.

CHAPTER 1

Species structure and phylogenetic relationships among North Atlantic redfish (genus *Sebastes*; Cuvier 1829) based on mitochondrial NADH dehydrogenase subunit 3 DNA sequence data

1. Abstract

The species structure and the phylogenetic relationships among the four North Atlantic redfish species, *Sebastes mentella*, *S. marinus*, *S. fasciatus* and *S. viviparus*, and their morphological types, as well as their phylogenetic relationships with the North Pacific *Sebastes* species *S. alutus*, the South Atlantic species *S. capensis* and a North Atlantic species of the related genus *Helicolenus* (*H. dactylopterus*) were investigated by examining mitochondrial DNA sequence variation. For this purpose, the complete mitochondrial NADH dehydrogenase subunit 3 (ND3) gene was sequenced for 337 individuals of these species from different localities. The sequence data were used to determine the levels of inter- and intraspecific genetic differentiation and for phylogenetic reconstruction.

In general, the level of genetic differentiation within the genus *Sebastes* was low, with interspecific sequence divergences between 0.3% and 7.4%. The low level of genetic differentiation of 0.3% to 2.3% sequence divergence between all pairs of North Atlantic *Sebastes* haplotypes, the low resolution in the phylogenetic trees and the star-like topology of the statistical parsimony network indicate a population bottleneck and/or founder event with a subsequent population expansion and a rapid speciation in a relatively short time. Nevertheless, the degree of genetic variability was large enough to recognise diagnostic mutations for several of the species analysed. The presence of diagnostic mutations can be interpreted as evidence of recent reproductive isolation between the lineages, even though there is a strong indication of ancient introgressive hybridisation events, in particular between *S. marinus* and *S. fasciatus* on the Flemish Cap and, in the case of the so-called "giant" *S. marinus*, between *S. marinus* and *S. mentella*. The phylogenetic results provide evidence that the North Pacific species *S. alutus* is more closely related to the North Atlantic species than *S. capensis*, which occurs in the South Atlantic. These findings support the general hypothesis that the North Atlantic species have arisen from a single ancestral species originating from the North Pacific, which then radiated to form the four North Atlantic species. The evolutionary scenario revealed by the phylogenetic analysis of the ND3 sequence data leads to the hypothesis that *S. mentella* is the representative of the most basal lineage of North Atlantic *Sebastes*. The results further indicate a direct descendance of the *S. fasciatus* and *S. marinus* lineages from the basal *S. mentella* lineage, whereas the *S. viviparus* lineage probably originates from an earlier split from this basal lineage. Furthermore, the observation of two haplotype lineages with a relatively high level of intraspecific genetic divergence, abundant in the *S. marinus* samples from Iceland and Greenland, gives evidence of two genetically isolated groups within those samples and suggests cryptic speciation within the North Atlantic *Sebastes* species complex.

2. Introduction

The viviparous genus *Sebastes* (Cuvier 1829) is the most species-rich assemblage of scorpaenid fishes, with about 102 species worldwide (Love et al. 2002; Rocha-Olivares et al. 1999a). While most of the *Sebastes* species (about 96) are found in the North Pacific, only few species are known from the South Pacific, South and North Atlantic (Rocha-Olivares et al. 1999c; Love et al. 2002).

In the North Atlantic four closely related *Sebastes* species are recognised: *Sebastes mentella* (Travin 1951), *S. marinus* (Linné 1758), *S. fasciatus* (Storer 1854) and *S. viviparus* (Krøyer 1845). *S. marinus* and particularly *S. mentella* are widely distributed in the North Atlantic, while *S. viviparus* and *S. fasciatus* are essentially restricted to the Northeast Atlantic and to the Northwest Atlantic, respectively (Barsukov 1973; Barsukov et al. 1990). Except *S. viviparus*, all North Atlantic *Sebastes* species are commercially important fishery resources.

The North Atlantic *Sebastes*, or redfish, species appear morphologically remarkably similar, with widely overlapping meristic and morphometric characters (Barsukov et al. 1985; Rubec et al. 1991). Previous genetic studies based on proteins and mitochondrial markers have generally revealed low levels of intra- and interspecific variation (Payne and Ni 1982; McGlade et al. 1983; Rehbein 1983; Trottier et al. 1989; Nedreaas and Nævdal 1989; 1991a; 1991b; Rubec et al. 1991; Nedreaas et al. 1994; Bentzen et al. 1998; Sundt and Johansen 1998). At present, the most precise genetic technique for the identification of North Atlantic *Sebastes* species is microsatellite analysis, but the technique is not routinely applied in field surveys (Roques et al. 1999a; reviewed in Kendall 2000).

Hence, species identification is difficult and time-consuming compared to other commercially exploited fish species, because several morphological and protein characters have to be combined to confidently distinguish the taxa at adult stage (Nedreaas and Nævdal 1989; Rubec et al. 1991; Gascon 2003). Consequently, redfish in the North Atlantic is exploited and managed as one species (Rubec et al. 1991; Gascon 2003).

The low levels of morphological and genetic differentiation between the North Atlantic redfish species have not only hampered the unambiguous species identification, but also our understanding of the evolutionary relationship of the four species.

Even though it is generally believed that the North Atlantic *Sebastes* species are monophyletic and derived from a single North Pacific ancestor, most likely Pacific ocean perch (*Sebastes alutus*) or its ancestor (Rocha-Olivares 1999a; Love et al. 2002), the interpretations of the relationships among the four species are still diverging, depending on the characters studied by the different authors (e.g. Barsukov and Zakharov 1973; Payne and Ni 1982; McGlade et al. 1983; Roques et al. 1999a). Nevertheless, the resolution of the phylogenetic relationships among

these closely related species is fundamental to our understanding of their evolution and radiation of biological diversity.

In addition to the difficulties in identifying species and inferring phylogenetic relationships, the genetic status of certain morphological phenotypes has been controversially discussed. Some authors separate *S. mentella* in the Irminger Sea into a so-called "oceanic" and a "deep-sea" phenotype with differences in certain morphological (Magnússon and Magnússon 1995) and molecular (Johansen et al. 2000b) characters. There has been a large controversy about whether the different *S. mentella* phenotypes represent different stocks (ICES 1998a; Saborido-Rey et al. 2005). Also the taxonomic status of the so-called "giant" redfish (or "giant" *S. marinus*) - individuals with an average total length above 60 cm that are morphological identified as *S. marinus*, but are separated from *S. marinus* due to some morphological and hemoglobin characteristics - remains unclear (Kotthaus 1961; Altukhov and Nefyodov 1968; Johansen et al. 2000a).

Mitochondrial DNA is a useful genetic marker for molecular systematics, for the examination of phylogenetic relationships among closely related species (Thomas and Beckenbach 1989; Avise 1994; Stepien and Kocher 1997), but also for population genetic studies (Avise 1998), due to its predominantly maternal mode of inheritance, clonal nature, lack of recombination and rapid rate of evolution. Several studies successfully used mitochondrial gene sequence data to gain insight into the molecular evolution, systematics and the population structure of Pacific and South Atlantic *Sebastes* species (Rocha-Olivares and Vetter 1999; Rocha-Olivares et al. 1999a; 1999b; 1999c; 1999d; Kai et al. 2002a and 2002b; Cope 2004).

In this study, the mitochondrial ND3 gene (coding for the NADH dehydrogenase subunit 3) was chosen to investigate the species structure and the phylogenetic relationship of the North Atlantic *Sebastes* species, as it has a higher substitution rate than cytochrome b and most other nuclear and mitochondrial genes (Thomas and Beckenbach 1989; Domanico and Phillips 1995; Domanico et al. 1997). Therefore, the ND3 gene may be one of the most appropriate regions of the mtDNA for species identification and for the examination of the phylogenetic relationships among the North Atlantic *Sebastes* species.

The aim of this study was to estimate the inter- and intraspecific variability of the complete ND3 gene for all four North Atlantic *Sebastes* species, the related Pacific species *S. alutus* (Gilbert 1890), the South Atlantic species *S. capensis* (Gmelin 1789) and a North Atlantic species of the closely related scorpaenid genus *Helicolenus*, *H. dactylopterus* (Delaroche 1809), in order to gain insight into the species structure and the phylogenetic relationships of the four North Atlantic *Sebastes* species as well as into the genetic relationship between *S. alutus*, *S. capensis* and *Helicolenus dactylopterus*. In addition, the genetic status of the different morphological types of North Atlantic *Sebastes* was investigated.

3. Material and methods

3.1 Sample collection

North Atlantic redfish species and their morphological types were sampled by different scientists from institutes from Norway, Iceland, Spain and Germany (see Appendix I) during research surveys across the North Atlantic from 1996 to 2001.

Individual redfish were pre-classified into species on board using morphometrical (e.g. eye diameter, length and inclination of the preopercular spines) and morphological characters (e.g. spines, body shape, symphysial tubercle, colour) usually used for redfish (genus *Sebastes*) identification (Barsukov 1973; Barsukov et al. 1985). Samples from the shelf of Greenland were pre-classified into species using hemoglobin electrophoresis (Nedreaas and Nævdal 1989; 1991a), because the species identification - particularly the differentiation of *S. marinus* and *S. mentella* specimens - using standard characters was ambiguous. Hemoglobin analyses were performed by a Norwegian scientist[1] directly on board. These individuals were then frozen whole and were later analysed in the laboratory by a specialist in redfish morphometrics[2].

Specimens of *S. mentella* collected in the Irminger Sea were typed as "oceanic" and "deep-sea" type by Icelandic scientists[3], according to characters described by Magnússon and Magnússon (1995), e.g. colour and parasite infestation. In addition, DNA samples of 37 extremely large *S. marinus* (average length 78 cm) collected in the Irminger Sea in 1996, defined as "giant" *S. marinus* by Icelandic scientists[3], were provided by the cooperating institute from Iceland (see Appendix I). Gill samples for genetic analyses were taken either directly from the fish on board the research vessels or afterwards in the laboratory at the different institutes (see Appendix I) from fish that had been frozen whole for subsequent morphometrical analyses. The gill samples were preserved in 100% ethanol.

In total 313 adult individuals of the genus *Sebastes* (130 *S. mentella,* 107 *S. marinus,* 34 *S. fasciatus* and 42 *S. viviparus*) and of four specimens of *Helicolenus dactylopterus* were collected from different locations across the North Atlantic. In addition, gill samples of ten individuals of *Sebastes alutus* (Cape St. James, Northeast Pacific) and *Sebastes capensis* (Tristan da Cunha, South Atlantic) were provided by cooperating institutes from Canada and Tristan da Cunha (see Appendix I). Total lengths were recorded for all North Atlantic *Sebastes* samples. Sex was also recorded for all North Atlantic *Sebastes* samples except the *S. viviparus* sample from Norway (sample VINO01), the *S. mentella* sample from Norway (sample MENO01) and two *S. mentella*

[1] Hemoglobin analyses were performed by Torild Johansen, University of Bergen (UIB), Norway.

[2] Morphometrical analyses were performed by Dolores Garabana, Institute of Marine Research, Vigo, Spain.

[3] Scientists from the Marine Research Institute (MRI) in Reykjavík, Iceland

samples from the Irminger Sea collected in the year 1997 (14 individuals of each morphological type "deep-sea" and "oceanic"; samples MEds97 and MEoc97).

Table 1.1 gives an overview of the samples collected. In Appendix I the exact sampling information (e.g. position, month, depth) is listed. The North Atlantic sampling locations of the *Sebastes* samples are pictured in detail in Figure 1.1 (S. *marinus*, *S. viviparus*, *S. fasciatus* and *Helicolenus dactylopterus*) and 1.2 (*S. mentella*). Size distributions of sampled *S. marinus* and *S. mentella* are given in Appendix II.

Table 1.1. Summary statistics for redfish samples collected showing species names, morphological types, geographical origins, sampling years, sample codes, sample sizes (*n*) and sex (number of males and females). The species names are based on the pre-classification on board using standard characters (Barsukov 1973; Barsukov et al. 1985) and, in case of the samples of *S. marinus* and *S. mentella* from Greenland, on hemoglobin electrophoretic patterns (Hb; Nedreaas and Nævdal 1989; 1991a). Dashes indicate missing information.

Species	Type	Location	Year	Code	*n*	Males	Females
Sebastes mentella	deep-sea	Irminger Sea	1997	MEds97	12	7	5
Sebastes mentella	deep-sea	Irminger Sea	1997	MEds97	14	-	-
Sebastes mentella	deep-sea	Irminger Sea	2001	MEds01	15	8	7
Sebastes mentella	oceanic	Irminger Sea	1996	MEoc96	12	10	2
Sebastes mentella	oceanic	Irminger Sea	1997	MEoc97	14	-	-
Sebastes mentella	oceanic	Irminger Sea	2001	MEoc01	15	11	4
Sebastes mentella		Norway	2001	MENO01	14	-	-
Sebastes mentella		SW-Iceland	2001	MESWIc01	10	7	3
Sebastes mentella [Hb]		Greenland East	2001	MEEGr01	10	4	6
Sebastes mentella		Flemish Cap	2001	MEFC01	14	9	5
Sebastes mentella					**130**	**56**	**32**
Sebastes marinus		Norway	2001	MANO01	10	10	0
Sebastes marinus		SW-Iceland	1997	MASWIc97	11	2	9
Sebastes marinus		SE-Iceland	2001	MASEIc01	10	9	1
Sebastes marinus [Hb]		Greenland East/Dohrn Bank	2001	MADB01	9	4	5
Sebastes marinus [Hb]		Greenland East	2001	MAEGr01	10	7	3
Sebastes marinus		Flemish Cap	2001	MAFC01	20	15	5
Sebastes marinus	giant	Irminger Sea	1996	MAG96	37	13	24
Sebastes marinus					**107**	**60**	**47**
Sebastes fasciatus		Flemish Cap	1997	FAFC97	13	10	3
Sebastes fasciatus		Flemish Cap	2001	FAFC01	21	6	15
Sebastes fasciatus					**34**	**16**	**18**
Sebastes viviparus		Norway	2001	VINO01	14	-	-
Sebastes viviparus		SW-Iceland	1996	VISWIc96	12	7	5
Sebastes viviparus		SW-Iceland	2001	VISWIc01	16	7	9
Sebastes viviparus					**42**	**14**	**14**
Sebastes alutus		Cape St. James/Pacific	2001	ALUT	10	-	-
Sebastes capensis		Tristan da Cunha/S-Atlantic	2002	CAP	10	-	-
Helicolenus dactylopterus		Shetland Islands (West)	2001	HEL	4	-	-
Total					**337**		

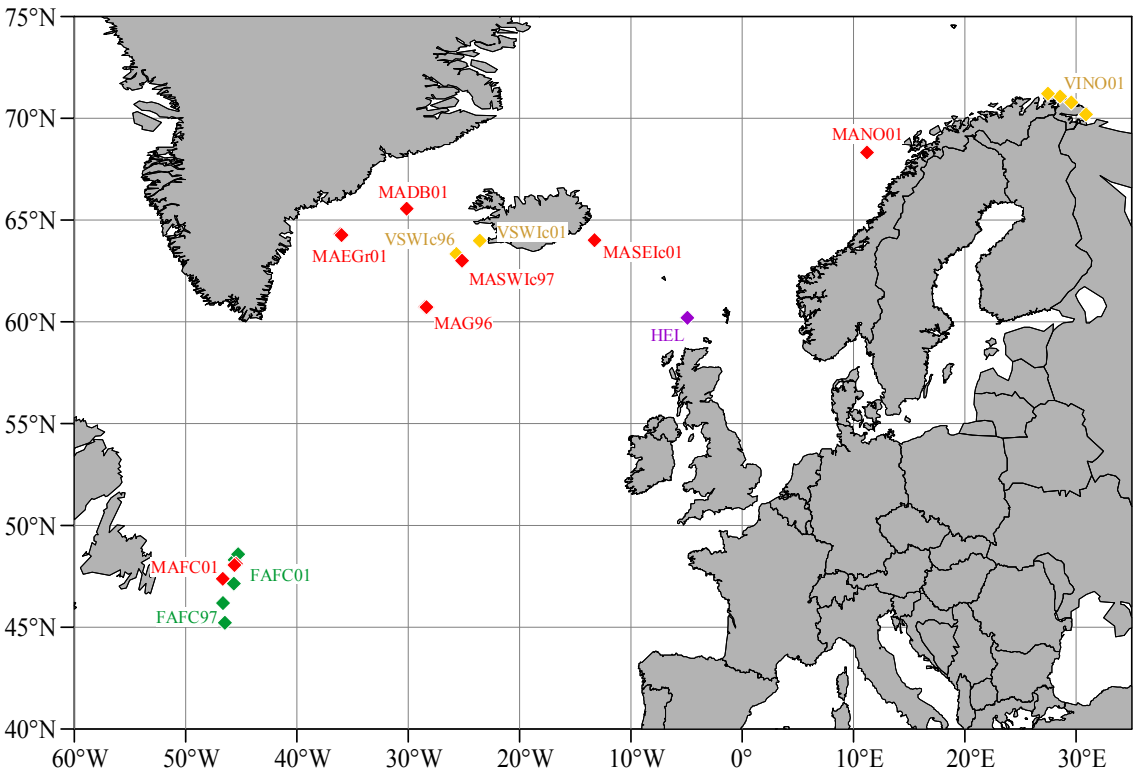


Figure 1.1. North Atlantic sampling locations for the ND3 gene sequence analysis. Red: S. *marinus*; yellow: *S. viviparus*; green: *S. fasciatus*; purple: *Helicolenus dactylopterus*. Sample codes are explained in Table 1.1.

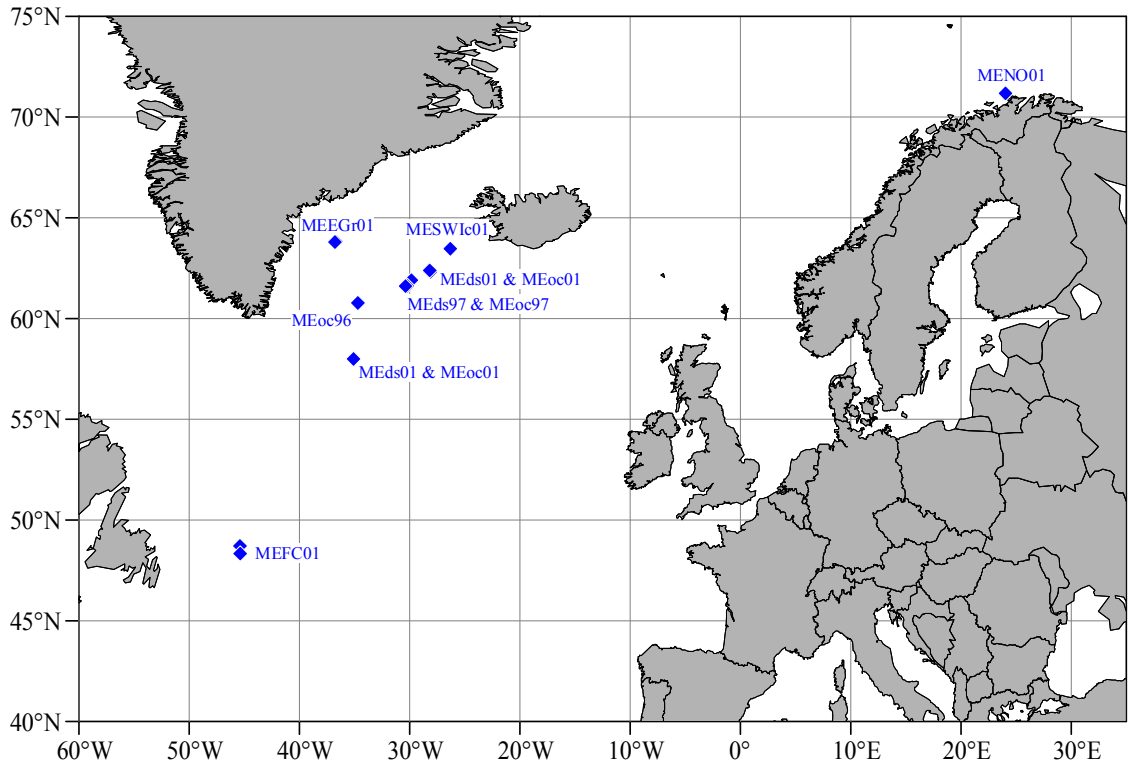


Figure 1.2. North Atlantic sampling locations for the ND3 gene sequence analysis. Blue: *S. mentella*. Sample codes are explained in Table 1.1.

3.2 Sample analysis

3.2.1 DNA extraction

Total genomic DNA was extracted with the DNeasy™ Tissue Kit (Qiagen) following manufacturers instructions. During this procedure, RNA was digested with RNase purchased from Qiagen following manufacturers instructions.

3.2.2 DNA quantitation

A sensitive method for reliable quantitation of double stranded DNA is the fluorometric DNA assay employing Hoechst 33258 Dye. This dye is a bisbenzimide DNA intercalator that binds to the AT rich regions of double stranded DNA and excites in the near UV (350 nm) and emits in the blue region (455 nm) (Labarca and Paigen 1980; Downs and Wilfinger 1983). The fluorescence intensity at 455 nm corresponds to the quantity of DNA present in the sample. The DNA concentration (µg/ml) in each sample was measured with a fluorescence/luminescence spectrophotometer (Perkin Elmer).
Samples, standards and blanks were prepared according to Downs and Wilfinger (1983) with the following modifications:
5 µl of DNA sample, 95 µl of assay buffer (100 mM NaCl, 10 mM EDTA, 10 mM Tris-HCl, pH 7.0) and 2000 µl Hoechst 33258 Dye working solution (100 ng/ml) were mixed in a quartz cuvette (10 mm pathlength). For calibration, two additional quartz cuvettes were prepared using calve thymus DNA (50 ng/µl) as standard and deionised water as blank instead of the sample DNA. The measurements were carried out following manufacturers instruction for the fluorescence spectrophotometer.

3.2.3 In vitro amplification of the mitochondrial ND3 gene by polymerase chain reaction (PCR) using universal primers

A mitochondrial DNA fragment containing the ND3 gene was amplified by polymerase chain reaction (PCR: Mullis et al. 1986; Mullis and Falcoona 1987; Saiki et al. 1988), using the universal primer pair ARG and GLY, based on conserved regions of the genes coding for the transfer RNAs for arginine and glycine, $tRNA^{Arg}$ and $tRNA^{Gly}$, which flank ND3 in vertebrate mitochondrial genomes (Meyer 1993; 1994; see Figure 1.3). The primer sequences were taken from McKay et al. (1996). A list of primer sequences is given in Table 1.2.

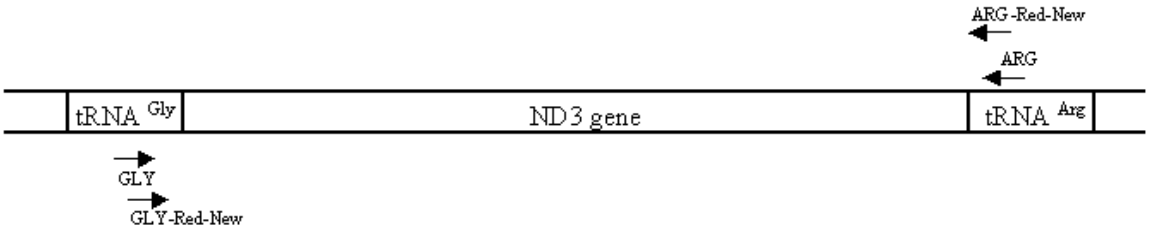


Figure 1.3. Schematic diagram of the location of mitochondrial ND3 gene amplification and sequencing primers. The positions of the primers are represented by horizontal arrows.

PCR amplifications of DNA samples of two specimens of each of the North Atlantic redfish species (*S. mentella, S. marinus, S. fasciatus* and *S. viviparus*) were performed following a protocol by Trautner (2000).

Reactions were carried out in 50 µl reaction volumes containing 5 µl DNA template (approximately 300 ng), 0.7 mM of each amplification primer (ARG and GLY; MWG Biotech), 200 µM of each deoxynucleoside triphosphate (dNTP; Gibco/Life Technologies), 1.8 mM $MgCl_2$ (MWG Biotech), 1x *Taq* polymerase specific PCR buffer (MWG Biotech) and 2 units *Taq* polymerase (MWG Biotech).

A negative (sterile PCR grade dH_2O) and a positive (rainbow trout DNA, as described in Trautner 2000) control DNA template were run for each PCR amplification performed.

The PCRs were performed with a T-Gradient thermocycler (Biometra), starting with an initial denaturing step of 3 minutes at 95°C, followed by 30 cycles of 30 seconds denaturation at 95°C, one minute primer annealing at 45°C and one minute and 30 seconds primer extension at 72°C and a final primer extension step for 3 minutes at 72°C.

Table 1.2. Oligonucleotide primer sequences. All primers were purchased from MWG Biotech.

Primer	Sequence (5' to 3')	Source
ARG	ATGCGGATCCT(C/T)TTGAGCCGAAATCA	McKay et al. 1996
GLY	ACGTGAATTCGTA(G/T)(A/G)(A/C)GTG(A/G)CTTCCAA	McKay et al. 1996
M13 universal sequence primer	GTAAAACGACGGCCAGT	Universal
M13 reverse sequence primer	CAGGAAACAGCTATGAC	Universal
ARG-Red-New	CCTTTTGAGCCGAAATCAAAT	This study
GLY-Red-New	GCTTCCAATCACCCGGTCTT	This study

3.2.4 Gel electrophoresis

The PCR products were visualised by gel electrophoresis in a 1% agarose gel stained with ethidium bromide. 1 µl PCR product, 9 µl dH_2O and 1 µl loading buffer (0.1% SDS, 50% saccharose and 0.15% bromophenol blue) were mixed and loaded into the sample wells of the gel. 0.1 µg of a molecular weight ladder (1-kb-Plus-Ladder; Invitrogen) was run on every gel as a reference to determine the amplification product size.
To consider the amplification reaction as positive, the PCR product must be approximately 450 base pairs long, as this is the length of the mitochondrial target fragment. If the length of the amplified fragment differs from 450 bp or if more than one band is observed in the agarose gel, the amplified product represents either a different DNA fragment or, in the latter case, the PCR products are heterogeneous and represent therefore different DNA fragments. Such PCR artefacts can be caused by primer mismatch, when the PCR primer is too unspecific or the annealing temperature is too low.
As electrophoresis buffer TBE buffer (1x) was used. The electrophoresis was run at 70-100 V/20-80 mA for about an hour. The DNA was visualised on a UV transilluminator and documented by camera.

3.2.5 Cycle sequencing of PCR products

As the PCR products showed the appropriate length of approximately 450 bp and the reactions gave a sufficient high yield of PCR product, it was tested if direct sequencing of the amplified fragments is possible.
The cycle sequencing reaction is based on the standard dideoxynucleotide terminated chain elongation method (Sanger et al. 1977), but utilises a linear polymerase reaction to amplify DNA that is complementary to the target DNA (Murray 1989).
The sequencing test was performed using the "SequiTherm EXCEL™ II DNA sequencing Kit" (Epicentre Technologies) according to the manufacturers protocol with some modifications.
The reactions were performed bi-directionally with forward (GLY) and reverse primer (ARG). Each of the two fluorescence (IRD800)-labelled sequencing primers ARG and GLY was used in a separate reaction. 7.2 µl PCR product, 7.2 µl SequiTherm EXCEL™ II DNA sequencing buffer, 1 µl SequiTherm EXCEL™ II DNA Polymerase (5 u), 4 pmol IRD800-labelled ARG or GLY and 0.6 µl dimethyl sulfoxide (DMSO) were mixed in a 0.5 µl microtube. 4 µl of this so-called "Premix" was mixed with 2 µl of each termination mix (SequiTherm EXCEL™ II-LC Termination Mix A, C, G, T) in a microtiter plate. A drop of mineral oil was used to cover the sequencing reactions to avoid evaporation.

The sequencing reactions were initially denatured for 5 minutes at 95°C, followed by 30 cycles of 30 seconds denaturation at 95°C, 15 seconds primer annealing at 50°C and one minute primer extension at 70°C. All reactions were performed in a T-Gradient thermocycler (Biometra).

3.2.6 Analysis of the sequencing reactions on an automated sequencer

The reaction products were mixed with 3 µl stop solution (included in the SequiTherm EXCEL™ II DNA sequencing Kit), denatured for 5 minutes at 95°C, placed on ice and loaded on a 41 cm polyacrylamide gel (SequaGel™ XR, Biozym; prepared according to manufacturers recommendations).
DNA fragments were visualised and analysed in an automated sequencer (LI-COR ™ 4200 GeneReadIR DNA system) according to manufacturers recommendations.
The data were collected automatically and analysed on computer using the Windows®-based DNA sequencing software e-Seq version 2.0 (LI-COR™).

3.2.7 Cloning of PCR products

The sequence reaction worked poorly when the PCR product was used directly. There are several factors that can yield such poor-quality sequencing results from PCR templates. Some of the most common factors are contaminated sequencing template DNA or mismatching primers. In both cases, heterogeneous PCR products are formed that are of the same or nearly the same size (and will therefore migrate to the same position on an agarose gel), but differ in sequence. Attempts to directly sequence such templates can result in double ladders on the polyacryamide gel, or, like observed in this study, PCR amplifications can result in too few products of the desired size to allow direct sequencing.
A common approach to overcome these problems is to clone those PCR products.
Therefore, PCR products were inserted into plasmid vectors (Ligation). These plasmids were subsequently transformed into competent cells of *Escherichia coli* (Transformation), which were then multiplied in cell culture. The plasmids were extracted and then sequenced using plasmid specific primers. This procedure often yields higher-quality sequences than direct sequencing of PCR products. The most convenient and direct method is TA cloning as described in Marchuk et al. (1991): Many thermostable DNA polymerases, e.g. the *Taq* DNA polymerase (MWG Biotech) used in this study, have a terminal transferase activity that preferentially adds adenine to the 3' ends of PCR products. PCR products with such single 3' adenylate extensions can be directly cloned into a vector containing complementary 3' thymidine overhangs.

The pGEM-T Easy Vector System II (Promega) was used to clone the PCR products. The pGEM-T Easy Vector contains an ampicillin-resistance gene for ampicillin selection and a multiple cloning site (MCS) integrated in a *lac*Z gene for blue/white colour screening.
On a medium containing the antibiotic ampicillin, only successfully transformed bacteria carrying the vector with the ampicillin-resistance gene will grow. If no colonies are obtained, the transformation has failed.
The *lac*Z gene produces beta-galactosidase. When transformed cells are grown on agar plates containing X-Gal (5-bromo-4-chloro-3-indolyl-beta-D-galactoside) induced with IPTG (isopropylthio-beta-D-galactoside), beta-galactosidase hydrolyses X-Gal forming an indigo-type blue-green precipitate. Successfully transformed bacteria therefore produce blue colonies. IPTG induces the beta-galactosidase production by binding and inactivating the *E. coli lac* repressor and activates the *lac* operon.
If the DNA fragment ligates successfully into the multiple cloning site, the *lac*Z gene is disrupted, and the host bacteria appear as white colonies (= positive colonies). If only blue colonies are obtained, the ligation reaction has failed.
PCR products of two specimens of each redfish species (*S. mentella, S. marinus, S. fasciatus* and *S. viviparus*) were cloned according to the suppliers protocol[1] (Promega) with slight modifications. Buffers, solutions and media were prepared following the suppliers protocol[1].

a) Ligation of the PCR product into vectors
Vector DNA, ligase buffer and T4 DNA ligase were brought down from -70°C storage and placed on ice. 5 µl 2x Rapid Ligation Buffer (thoroughly shaken to resuspend ATP granules), 1 µl of pGEM-T Easy Vector (50 ng/µl), 1 µl of PCR product, 1 µl T4 DNA ligase and 2 µl dH_2O (sterile PCR grade) were added together. The reaction was incubated for one hour at room temperature.

b) Transformation of the vector into competent cells of *E. coli*
The vectors were inserted into competent cells (JM109 High Efficiency Competent Cells, Promega) as follows:
2 µl of the ligation preparation was placed in a 1.5 ml Eppendorf tube on ice. The competent cells were removed from -70°C storage and thawed on ice for 5 minutes. 25 µl of the competent cell solution was used per 2 µl ligation preparation. The Eppendorf tubes were gently pipette-mixed and left on ice for 20 minutes.
The Eppendorf tubes were then placed in a water bath at 42°C for 43 seconds and placed immediately back on ice for 2 minutes.

[1] pGEM®-T and pGEM®-T Easy Vector Systems (Promega); Technical Manual No. 042

The samples were taken off the ice and 475 µl of SOC medium was added to each tube. The tubes were gently inverted and placed on a shaker with 150 rpm at 37°C for 90 minutes.
After incubation, the transformed competent cell cultures were gently inverted and 50 µl were aliquoted to LB agar plates (1.5%) containing 0.5 mM IPTG, 80 µg/ml X-Gal and 100 µg/ml ampicillin and spread across the plates thoroughly. The petri dishes were then inverted and placed at 37°C over night (12-16 h). After the incubation the samples were removed from the incubator and screened for positive cells.

c) Selection of positive colonies
Two to four clones of each PCR product transformation were picked and transferred into sterile culture tubes with 5 ml aliquots of LB/ampicillin medium. The tubes were placed on a shaker with 150 rpm at 37°C overnight.
After incubation, 1.5 ml of each culture was used for the plasmid isolation. For storage, 0.85 ml of each culture was mixed with 0.15 ml of glycerol, vortexed and then stored at -80°C.

d) Plasmid isolation by alkaline lysis
Plasmid DNA was isolated using alkaline lysis as described in Maniatis et al. (1982). Alkaline lysis involves the breaking down of bacterial cells to release the plasmid DNA and further purifies the plasmid DNA for the sequencing reaction.

e) Test for insertion
It was tested whether the plasmid vectors had the required DNA insertion (containing the ND3 gene) present by determining the size of the insert.
To obtain the size of the insert, a digestion reaction was carried out using the restriction enzyme *Eco*RI to cut out the insert from the plasmid vector (the multiple cloning site of the pGEM-T Easy Vector contains two *Eco*RI target sites). The reactions were carried out in 10 µl reaction volumes containing 7.2 µl dH_2O (sterile PCR grade), 1 µl plasmid DNA, 1x *Eco*RI specific reaction buffer (Gibco/Life Technologies) and 0.8 µl *Eco*RI (0.8 units; Gibco/Life Technologies). The reactions were incubated for 2 h at 37°C and subsequently run on a 1% agarose gel stained with ethidium bromide, using a molecular weight ladder (1-kb-Plus-Ladder, Invitrogen) as sizing standard (as described in 3.2.4) and 1 µl untreated pGEM-T Easy Vector as control. The DNA bands were visualised on a UV transilluminator and documented by camera.
Only when the restricted inserts revealed a length of approximately 450 bp, the plasmid DNA of the corresponding clone was chosen for direct sequencing, as this was the expected length of the mitochondrial target fragment.

3.2.8 Sequencing of plasmid vectors

Two clones of each of the eight individual redfish (two individuals of each of the four species) were chosen for sequencing to obtain the sequence of the ND3 gene and the flanking regions for each individual. The vector plasmid contains a binding site for the universal M13 sequence primer and the M13 reverse sequence primer (primer sequences are given in Table 1.2). Therefore, it is possible to directly sequence the whole plasmid containing the inserts from both directions.

The sequencing reactions were performed as described in 3.2.5 with the following modifications:

The premix contained 3.2 µl H_20 (sterile PCR grade), 4 µl plasmid DNA, 7.2 µl SequiTherm EXCEL™ II DNA sequencing buffer, 1 µl SequiTherm EXCEL™ II DNA Polymerase (5 u), 4 pmol IRD800-labelled M13 universal sequence primer or M13 reverse sequence primer (MWG Biotech), respectively, and 0.6 µl DMSO.

Each of the two fluorescence (IRD800)-labelled sequencing primers M13 universal sequence primer and M13 reverse sequence primer was used in a separate reaction.

The sequencing reactions were analysed on an automated sequencer as described in paragraph 3.2.6.

The sequences of the plasmid clones were aligned using the software BioEdit version 5.0.1 (Hall 1999). The sequences of the H (heavy) strands (= coding strand) were proofread using the sequences of the complementary L (light) strands.

The inserted DNA fragment containing the ND3 gene and a part of the flanking $tRNA^{Arg}$ and $tRNA^{Gly}$ genes was identified by aligning the sequences with a salmon (*Salmo salar*) ND3 gene sequence taken from the database GenBank (accession number U28366; reference: McKay et al. 1996).

3.2.9 Construction of redfish specific ND3 primers

Based on the alignment of all eight individual sequences, redfish specific ND3 primers (ARG-Red-New and GLY-Red-New) in the consensus sequence of the flanking conserved $tRNA^{Arg}$ and $tRNA^{Gly}$ genes were designed.

The designed primers GLY-Red-New and ARG-Red-New were similar in length (20 and 21 bp, respectively) with a GC content of 38% (ARG-Red-New) and 55% (GLY-Red-New). The primer sequences are listed in Table 1.2.

Unlabelled primers for PCR amplification as well as IRD800-labelled sequencing primers were purchased from MWG Biotech.

3.2.10 Optimising PCR conditions for the newly designed primers GLY-Red-New and ARG-Red-New

The previously applied PCR protocol by Trautner (2000; see paragraph 3.2.3) was optimised and adjusted to the new PCR primers.
To determine the optimal annealing temperature for the new primers, a gradient PCR cycler (T-Gradient thermocycler; Biometra) was used. In this way, a variety of annealing temperatures between 50°C and 66°C was tested.
The optimal primer concentrations had to be determined empirically, by testing concentrations in the range of 0.3 mM and 2 mM.
The optimal $MgCl_2$ concentrations were determined by testing concentrations between 1 mM and 3 mM. The magnesium ion concentration affects *Taq* polymerase performance and primer annealing, as free Mg^{2+} ions are needed for enzyme activity, but too many Mg^{2+} ions lead to reduced fidelity and increased non-specific amplification. Mg^{2+} also binds free dNTPs, primers and template, reducing the free quantities available.
The PCR products were visualised by gel electrophoresis in a 1% agarose gel stained with ethidium bromide as described above.
An annealing temperature of 62°C and primer concentrations of 1 mM of each amplification primer were chosen as these conditions gave an abundant yield without additional non-specific products. The $MgCl_2$ concentration of 1.8 mM was maintained, as it gave the best yield.
The final PCR conditions were set as follows:
50 µl reaction volumes contained 2 µl DNA template (approximately 150 ng), 1 mM of each amplification primer (ARG-Red-New and GLY-Red-New), 200 µM of each deoxynucleoside triphosphate (dNTP; Gibco/Life Technologies), 1.8 mM $MgCl_2$ (MWG Biotech), 1x *Taq* polymerase specific PCR buffer (MWG Biotech) and 2 units *Taq* polymerase (MWG Biotech).
A negative (sterile PCR grade dH_2O) and a positive control DNA template (rainbow trout DNA, as described in Trautner 2000) were run for each PCR amplification performed.
The PCRs were performed as described previously (paragraph 3.2.3), but with a higher annealing temperature of 62°C.

3.2.11 Sequencing of PCR products using the newly developed redfish specific primers GLY-Red-New and ARG-Red-New

PCR reactions were directly sequenced, using the newly designed IRD800-labelled sequencing primers GLY-Red-New and ARG-Red-New and the protocol described in paragraph 3.2.5 but with modified reverse primer concentrations.

In case of the reverse primer ARG-Red-New, the fluorescence intensity of the fragments in the sequencing gel was not sufficient and therefore made the sequence analysis difficult. To optimise the sequencing reaction, the amount of IRD800-labelled ARG-Red-New primer was doubled (8 pmol instead of 4 pmol in the premix). The template volume was reduced to 5.2 µl in the premix. For the primer GLY-Red-New, the concentration of the sequencing components were maintained as described above. The sequencing cycles were maintained for both forward and reverse sequencing reactions. With the adjusted PCR and sequencing conditions, the ND3 gene of all individuals listed in Table 1.1 was sequenced. DNA fragments were visualised and analysed in an automated sequencer as described above.

3.3 Data analysis

3.3.1 Sequence alignment

Sequences were aligned using the software BioEdit version 5.0.1 (Hall 1999). The sequences of the H (heavy) strands (= coding strands) were proofread using the sequences of the complementary L (light) strands.

3.3.2 Sequence composition and variation

The 337 ND3 gene sequences were compared and identical sequences were pooled and defined as haplotypes (= mitochondrial genotype) for the subsequent analyses. Haplotype frequencies were calculated by direct counting of haplotypes.
Mitochondrial ND3 gene diversity within samples was estimated as the number of haplotypes, haplotype diversity ($\hat{h}$) and nucleotide diversity ($\hat{\pi}$) indices.
The haplotype diversity or gene diversity ($\hat{h}$; Nei and Tajima 1981) is a function of the numbers and frequencies of the different haplotypes in each sample, regardless of their sequence relationship. It is defined as the probability that two randomly chosen haplotypes are different in the sample.
Haplotype diversity is estimated as

$$\hat{h} = \frac{n}{n-1}(1 - \sum_{i=1}^{k} p_i^2)$$

where n is the number of gene copies (= number of sequenced individuals) in the sample, k is the number of haplotypes and p_i is the sample frequency of the i-th haplotype.

The nucleotide diversity ($\hat{\pi}$; Nei 1987) is a weighted sequence divergence measure and describes the average number of nucleotide differences between all pairs of haplotypes found in a sample (Nei and Kumar 2000).
It is given by

$$\hat{\pi} = \sum_{i=1}^{k} \sum_{j<i} p_i p_j \hat{d}_{ij}$$

where $\hat{d}_{ij}$ is an estimate of the number of mutations having occurred since the divergence of haplotypes *i* and *j*, *k* is the number of haplotypes, p_j is the frequency of haplotype *j* and p_i is the frequency of haplotype *i*.

Haplotype diversity ($\hat{h}$) and nucleotide diversity ($\hat{\pi}$) were estimated within each species sample using the software Arlequin version 2.000 (Schneider et al. 2000).

3.3.3 Site saturation

The sequences were examined for possible site saturation. Saturation is the equilibrium value of sequence divergence that is reached when multiple substitutions erase the record of previous substitutions at a site. The level of saturation in the sequences was assessed visually by plotting the number of pairwise transitional nucleotide substitutions (purine versus purine and pyrimidine versus pyrimidine) and transversional substitutions (purine versus pyrimidine) against uncorrected p-distance (= proportion of nucleotide sites at which the two sequences compared are different; see paragraph 3.3.4) for all pairwise comparisons of haplotypes. As sequences approach saturation, such a plot will show a decrease in transitions and transversion with increasing distance. Once saturation is reached, the number of substitutions should remain constant with increasing distance, indicating that little, if any, phylogenetic signal remains in the data and that the data are uninformative.

3.3.4 Phylogenetic tree reconstruction

There are two approaches to phylogenetic reconstruction. One approach uses genetic distances (distance method), the other approach uses discrete characters (character state method). In distance methods, pairwise distances are computed for all pairs of haplotypes and a phylogenetic tree is constructed by certain principles and algorithms (e.g. Kumar et al. 1993). In character state methods, data with discrete character states - such as nucleotide states in DNA sequences - are

used, and a tree is constructed by considering the evolutionary relationships of DNA sequences at each nucleotide position (Kumar et al. 1993).

To infer a phylogeny of the 351 bp ND3 gene fragment, phylogenetic trees were constructed using one distance based method and two character state based methods: (1) the neighbour-joining method (a distance based method), (2) maximum parsimony (MP) and (3) maximum likelihood (ML) procedures (the latter two are character state based methods).

The methods are reviewed in Swofford et al. (1996), therefore only the basic principles will be described here.

Genetic distances and neighbour-joining (NJ)

The neighbour-joining (NJ) method (Saitou and Nei 1987) is a widely used distance clustering algorithm that allows for unequal rates of divergences among lineages. The main advantage of neighbour-joining is its efficiency. It can be used on very large data sets for which other methods of phylogenetic inference (e.g. maximum parsimony, maximum likelihood) are computationally too intensive.

The procedure identifies the closest pairs of taxonomic units (in this case haplotypes) by the distances between them.

The following distance measure indices and substitution models were used for estimating pairwise sequence divergence between the different haplotypes and the construction of neighbour-joining trees:

(1) The number of nucleotide differences (= number of sites) at which the two sequences compared are different.

(2) p-distance (e.g. Kumar et al. 1993)

This distance is the proportion (p) of nucleotide sites at which the two sequences compared are different. It is obtained by dividing the number of nucleotide differences by the total number of nucleotides compared. It does not make any correction for multiple substitutions at the same site, substitutional rate biases (for example, differences in the transitional and transversional rates), or differences in evolutionary rates among sites.

(3) Kimura 2-parameter model (K2P; Kimura 1980)

In actual sequence data the rate of transitional nucleotide substitutions is often higher than that of transversional substitutions. This is particularly so for animal mitochondrial DNA (Brown et al. 1982; Avise 1994). The Kimura 2-parameter distance measure (Kimura 1980) is a corrected percentage of nucleotides for which two haplotypes are different. This distance measure corrects for multiple substitutions (= multiple hits) per site and takes into account different substitution rates between transitions and transversions, while assuming that the four nucleotide frequencies are the same and that rates of substitution

do not vary among sites (Kumar et al. 1993). The transition-transversion ratio is estimated from the data.

(4) Hasegawa-Kishino-Yano model (HKY85; Hasegawa et al. 1985)
This model is a generalisation of the Kimura 2-parameter model. HKY85 takes into account different substitution rates between transitions and transversions, but does not assume equal base frequencies.

Based on the distance matrices of pairwise genetic distances, neighbour-joining trees were constructed. The algorithm of the NJ method is explained in detail in Nei and Kumar (2000), but Swofford et al. (1996) give an overview of the principles of the NJ analysis:
Starting with the distance matrix based on e.g. Kimura 2-parameter's genetic distance, at each successive step of the analysis a modified distance matrix is constructed in which the separation between each pair of nodes is adjusted on the basis of their average divergence from all other nodes. The tree is then constructed by linking the least distant pair of nodes as defined by the modified matrix. When two nodes are linked, their common ancestor node is added to the tree and the terminal nodes are removed from the tree. This process converts the newly added common ancestor into a terminal node on a tree of reduced size. At each stage in the process two terminal nodes are replaced by one new node. The process is complete when two nodes remain, separated by a single branch (Swofford et al. 1996).
By minimising the sum of branch lengths at each step of sequence clustering (minimum-evolution criterion), the algorithm searches for a topology showing the smallest value of the sum of all branches - as an estimate of the correct tree.
The NJ method produces an unrooted tree, because it does not assume a constant substitution rate i.e., an evolutionary clock (Saitou and Nei 1987).
The neighbour-joining trees were constructed using the software package PAUP* (version 4.0b10; Swofford 2002) using simple pairwise distances (p-distance), the pairwise Kimura 2-parameter's genetic distance model (Kimura 1980) and the Hasegawa-Kishino-Yano model (HKY85; Hasegawa et al. 1985).
The robustness of the phylogenies (or internodes in the NJ tree) was tested using Felsenstein's (1985) bootstrap test.
The bootstrap test is a way to estimate the quality of a phylogenetic tree: Let there be *m* sequences, each with *n* nucleotides. A phylogenetic tree is reconstructed from these sequences using some tree building method (in this case neighbour-joining). From each sequence, *n* nucleotides are then randomly chosen with replacement, giving rise to *m* rows of *n* columns each. These now constitute a new set of sequences. A new tree is then reconstructed with these sequences. The topology of this tree is compared to that of the original tree. Each interior branch of the original tree that is different from the bootstrap tree in terms of the sequences it partitions

is given a score of 0; all other interior branches are given the value 1. This procedure of resampling the sites and subsequent tree reconstruction is repeated several hundred times (Nei and Kumar 2000). If a group appears in x percent of the bootstrap trees, the confidence level associated with that group is taken as x percent (Swofford 1993). This is the so-called bootstrap value (or the bootstrap support).
As a general rule, if the bootstrap value for a given interior branch is 95 or higher, then the topology at that branch is considered robust (Nei and Kumar 2000).
The reliability of each inferred neighbour-joining tree was tested by analysing 1000 bootstrap replicates, using the bootstrap option in PAUP* (version 4.0b10; Swofford 2002).

To estimate pairwise sequence divergence between the different haplotypes, the number of differences and p-distances between all pairs of haplotypes were calculated in the software package MEGA version 2.1 (Kumar et al. 2001).

Maximum parsimony (MP)

Maximum parsimony (MP) methods are widely used in phylogenetics (Swofford et al. 1996). MP attempts to select trees that minimise the number of evolutionary steps required to explain the observed data (= the total tree length). Parsimony methods have the advantage that they do not require explicit models of evolutionary change (Avise 1994; Swofford et al. 1996) and that data are not reduced into distances.
For constructing a maximum parsimony tree, only nucleotide sites at which there are at least two different kinds of nucleotides, each represented at least twice, are used (synapomorphies; Kumar et al. 2001). These sites are called parsimony-informative sites. Other variable sites are not used for constructing a MP tree, although they are informative for distance and maximum likelihood methods (Kumar et al. 2001).
For each possible tree topology, the sequences at each node are inferred to be those that require the least number of changes to give each of the two immediately descendant sequences (Weir 1996). The total number of changes required to traverse the whole tree is determined and the tree with the minimum number is the optimal (= most parsimonious).
As the examination of all possible tree topologies (exhaustive search) took excessively long calculation times caused by the large data matrix (351 characters and 31 haplotypes), a heuristic method was used to search for the most parsimonious tree. Heuristic tree searches generally operate by hill climbing methods, the peak representing the most parsimonious tree. An initial tree is used to start the process, and then the tree is improved by rearranging it in a way that minimises its length (Swofford et al. 1996). The initial tree is obtained by stepwise addition. Starting with three haplotypes the other haplotypes are connected - one at a time - to a developing tree until all taxa have been placed. Since heuristic searches are susceptible to become "stuck" in

a local optimum instead of finding the global optimum, initial trees were obtained by stepwise addition of sequences in a random input order, which results in different starting points (= initial trees) between search replicates.

This set of trees is then subjected to rearrangements that attempt to find shorter trees. This process is called branch-swapping.

In this study, the tree bisection and reconnection (TBR) branch-swapping algorithm as implemented in PAUP* was used. The tree is divided into two parts, and these are reconnected through every possible pair of branches in order to find a shorter tree. This is done after each taxon is added, and for all possible divisions of the tree.

Maximum parsimony analyses of relationships among haplotypes were conducted in PAUP* (version 4.0b10; Swofford 2002) using the heuristic search option, with the following parameters: Ignore uninformative characters, use all characters unordered and unweighted, accelerated transformation, starting trees obtained via stepwise addition with random haplotype additions and 100 random replicates (the search of the tree space will be started from 100 different "vantage points"), only one tree held at each step during stepwise addition (to reduce computation time), branch swapping by tree bisection and reconnection (TBR), steepest descent not enforced[1], retain minimal trees, collapse zero-length branches and save all minimal trees.

The most parsimonious trees were collapsed to obtain a strict consensus tree (including only those monophyletic branches occurring in all the most parsimonious trees).

The following indices were calculated in PAUP* to measure the fit of characters to the particular trees (Swofford 1993): The consistency index (CI; Kluge and Farris 1969), the retention index (RI; Farris 1989a; 1989 b) and the rescaled consistency index (RC; Farris 1989a; 1989b).

The consistency index (CI) is a measure of the average fit of the characters to a tree. It is calculated as the minimum possible number of character changes on any tree, divided by the actual number of steps on the current tree. The index ranges from one (no homoplasy, perfect fit) to a value asymptotically approaching zero (a lot of homoplasy, poorest fit).

The retention index (RI) is similar to the consistency index but removes bias due to autapomorphies. It varies from 0 to 1, with higher RI values indicating that characters in the data set are more congruent with each other and the tree. The rescaled consistency index (RC) is the product of CI and RI. Also the homoplasy index (HI = 1 - CI) was calculated in PAUP*.

The bootstrap option of PAUP* was used to examine the robustness of internal nodes. Clade support was assessed with heuristic searches of 1000 nonparametric bootstrap replications, using the same parameters as described above, and a 50% majority rule consensus tree was generated.

[1] If PAUP* finds a shorter tree during the branch-swapping process, it will discard any trees that have not been swapped yet, and use the new best tree as the starting tree for the next round of branch swapping. Swofford (2002) recommends not to apply this option when using random sequence addition.

Quartet puzzling maximum likelihood (ML) methods

Maximum likelihood methods (Felsenstein 1981) also do not reduce the data into distances, which is similar to MP methods. In contrast to MP methods, however, ML methods use all the data instead of just the parsimonious informative sites.

The maximum likelihood approach applies (like distance based methods) a defined model of sequence substitution to a given data set. The likelihood of observing a given set of sequence data for this specific substitution model is maximised for each tree topology, and the topology that gives the highest maximum likelihood is chosen as the final tree (Swofford et al. 1996). As the large data matrix caused excessively long computation times, the heuristic quartet puzzling maximum likelihood method (Strimmer and von Haesseler 1996) was used to construct the most-likely tree from the data set.

Quartet puzzling is a fast tree search algorithm based on the analysis of quartets of sequences. The method consists of three major steps. First, in the "maximum-likelihood step" the best four-haplotype tree is constructed from every possible quartet combination in a data set. Then a so-called "puzzling" step is used, in which these quartet trees are repeatedly combined to an overall tree (Strimmer and von Haesseler, 1996). During this step sequences are added sequentially in random order to a maximum–likelihood tree of a randomly chosen sequence quartet.

A new sequence is added to this subtree determined by a voting procedure, according to the quartets (for details see Strimmer and von Haesseler 1996). To examine the landscape of all possible optimal trees, the puzzling step is repeated several times.

Finally, in the "consensus step", the majority rule consensus of all intermediate trees is calculated, resulting in the quartet puzzling tree that shows groupings that are well supported.

The quartet puzzling tree displays reliability values from the puzzling step on each internal node, which indicate the number of times the node appeared during the puzzling process. If groups are found only occasionally in different runs, the resulting reliability values are low. The reliability values have been shown to be correlated with bootstrap values, at least for a few data sets (Strimmer and von Haesseler 1996).

ML analysis was performed using the quartet puzzling tree search procedure of Strimmer and von Haesseler (1996) with 1000 puzzling steps as implemented in PAUP* (version 4.0b10; Swofford 2002). Two different substitution models with the corresponding model parameters and the option "use empirical base frequencies" were used:

(1) The model of Felsenstein (F81; Felsenstein 1981), which assumes equal rates of substitution between all pairs of bases and allows for unequal base frequencies.

(2) The Hasegawa-Kishino-Yano model (HKY85; Hasegawa et al. 1985; see paragraph "Genetic distances and neighbour-joining") with a transition/transversion ratio of 3:1 (calculated from the data).

Helicolenus dactylopterus was used as outgroup taxon in all phylogenetic analysis. The presence of a more distantly related outgroup in a phylogenetic reconstruction has the effect of polarising the tree, making it possible to infer which taxa represent ancient lines and which have evolved more recently (McKay et al. 1996). *Helicolenus dactylopterus* was chosen as outgroup taxon, as a recent study has shown that the species is very closely related to the genus *Sebastes* (Smith and Wheeler 2004).

Congruence among methods (NJ, MP and ML) was assessed qualitatively by comparing similarity of tree topologies and corresponding nodal support.

Statistical parsimony network

Phylogenetic relationships among haplotypes were also inferred by constructing a statistical parsimony network (SPN; Templeton et al. 1992) using the program TCS version 1.13 (Clement et al. 2000).

TCS uses parsimony (as defined in Templeton et al. 1992) to construct pairwise distances (in this case number of mutational steps) between all haplotypes until the probability of parsimony (= probability that two haplotypes have a parsimonious relationship, i.e., the probability of having non unobserved mutations; Templeton et al. 1992) exceeds 95%. The number of mutational differences associated with the probability just above this 95% cutoff point represents the maximum number of mutational steps justified by the 95% parsimony criterion. This limit means that single nucleotide substitutions that interconnect haplotypes that differ only at a single nucleotide are regarded as certain (Templeton et al. 1992; Clement et al. 2000).

TCS then connects the haplotypes into a network, starting with the smallest distance (= one nucleotide difference), until all the sequences are included in a single network or the distance corresponding to the parsimony connection limit has been reached (Templeton et al. 1992). The resulting graph represents the reconstructed gene genealogy of the haplotypes (Clement et al. 2000).

The statistical parsimony was specifically designed for estimating haplotype trees when the level of sequence divergence is low, e.g. on intraspecific level. The method has been shown to have greater statistical power as well as accuracy and provides higher resolution for inferring phylogenetic relationships than the phylogenetic methods described above when there are few variable sites (Templeton et al. 1992).

3.3.5 AMOVA (analysis of molecular variance) and Φ_{ST} statistics

The partition of genetic diversity within and among the North Atlantic *Sebastes* samples was analysed by a hierarchical analysis of molecular variance (AMOVA; Weir and Cockerham 1984; Excoffier et al. 1992) as implemented in Arlequin version 2.000 (Schneider et al. 2000). The AMOVA approach used in Arlequin is based on analyses of variance of haplotype frequencies, but also takes into account the number of mutations between molecular haplotypes (Excoffier et al. 1992).

Samples of the same species that are to believed to be more related to each other than to samples of different species were grouped: All North Atlantic *Sebastes* samples were divided by species into four groups (one *S. viviparus* group, one *S. fasciatus* group, one *S. mentella* group and one *S. marinus* group). Then the proportion of the total genetic variation that can be found among species, among samples within species, and within samples, was calculated, using both the proportion of pairwise nucleotide differences (p-distances) and differences in haplotype frequencies estimated from the molecular data. The precise formulae for the estimation of the components of variance can be found in Excoffier et al. (1992).

As the sample size for "oceanic" *S. mentella* and "deep-sea" *S. mentella* was relatively high (41 individuals of each type), the pairwise Φ_{ST} between the two samples was calculated in Arlequin to investigate the degree of genetic differentiation between the two morphological types.

The fixation index Φ_{ST} (Excoffier et al. 1992) is a measure of population differentiation that summarises the degree of differentiation between population divisions, corresponding to Wright's (1951) F-statistics. Φ_{ST} is defined as the correlation of random pairs of haplotypes drawn from within populations relative to the correlation of pairs of random haplotypes drawn from the whole population (Excoffier et al. 1992). Φ_{ST} measures the amount of genetic variation in the total sample ("oceanic" and "deep-sea" *S. mentella* pooled) that is due to differences among populations (in this case subsamples) comprising that sample. This proportion can range from zero (no differentiation) to one (maximum of genetic differentiation).

The AMOVA and the pairwise Φ_{ST} value between the "oceanic" and the "deep-sea" *S. mentella* samples, based on information about haplotype frequencies and p-distances among pairs of haplotypes, were computed in Arlequin 2.000 (Schneider et al. 2000).

Significance and resulting P-values of the AMOVA variance components and of the pairwise Φ_{ST} value were tested in Arlequin version 2.000 (Schneider et al. 2000) using non-parametric permutation procedures as described in Excoffier et al. (1992) with 10,000 permutations.

4. Results

4.1 Sequence composition and diversity

The variation in the complete mitochondrial ND3 gene with a length of 351 base pairs was studied in 313 North Atlantic specimens of the genus *Sebastes* (*S. mentella, S. marinus*, *S. fasciatus* and *S. viviparus*), 10 individuals of *Sebastes alutus* (Northeast Pacific), 10 individuals of *Sebastes capensis* (Tristan da Cunha, South Atlantic) and four specimens of North Atlantic *Helicolenus dactylopterus,* which were used as outgroup.

In total 31 different ND3 gene haplotypes were detected among the 337 specimens sequenced in this analysis. 28 of the 31 haplotypes were found exclusively in the *Sebastes* samples (excluding the outgroup *Helicolenus dactylopterus*), with 23 haplotypes occurring only in North Atlantic *Sebastes* specimens.

The haplotypes were named according to the species in which they occurred either exclusively or (if the haplotype occurred also in other samples) with the highest frequency (see paragraph 4.2). In this way, haplotypes were labelled and numbered ME1 through ME8 (for *S. mentella*), MA1 through MA3 (for *S. marinus*), G1 through G4 (for "giant" *S. marinus*), FA1 through FA4 (for *S. fasciatus*), VI1 through VI4 (for *S. viviparus*), ALUT1 and ALUT2 (for *S. alutus*), CAP1 through CAP3 (for *S. capensis*) and HEL1 through HEL3 (for *Helicolenus dactylopterus*).

The complete haplotype alignment can be found in Appendix III.

Table 1.3 shows the number of polymorphic sites (parsimony informative and singleton) and substitutions (transitions and transversions) observed in the samples.

Table 1.3. Number of haplotypes, polymorphic sites (parsimony informative and singleton) and substitutions (transitions and transversions) observed in the samples.

	Total*	***Sebastes* total**	***Sebastes* North Atlantic**
Haplotypes	31	28	23
Polymorphic sites	59	42	23
1st position	9	7	3
2nd position	3	3	1
3rd position	47	32	19
Parsimony informative sites	49	32	10
Singleton sites	10	10	13
Substitutions	61	43	23
Transitions	47	38	21
Transversions	14	5	2

***** including the outgroup *Helicolenus dactylopterus*

The 31 different haplotypes showed 59 polymorphic sites (49 of these sites were parsimony informative) including 61 nucleotide substitutions. Within the genus *Sebastes,* haplotypes were defined by 42 polymorphic sites (32 parsimony informative) including 43 nucleotide substitutions. The North Atlantic *Sebastes* haplotypes were characterised by 23 polymorphic sites (10 parsimony informative) with 23 nucleotide substitutions.

The majority of the substitutions were transitions and most of the substitutions were observed at 3rd codon positions. The transition/transversion ratio was much higher for the North Atlantic *Sebastes* data set (approximately 10:1) and considerably lower when all haplotypes were included (approximately 3:1).

Haplotype diversity ($\hat{h}$), nucleotide diversity ($\hat{\pi}$) and total number of haplotypes in the different samples are presented in Table 1.4.

Table 1.4. Number of individuals analysed per species (n), number of haplotypes (n_h), haplotype diversity ($\hat{h}$) and nucleotide diversity ($\hat{\pi}$). Standard errors (S.E.) are given in parentheses.

Species		***n***	n_h	***h***	π
S. mentella	**total**	**130**	**9**	**0.553 (0.042)**	**0.0019 (0.0016)**
S. marinus	**total (without "giants")**	**70**	**6**	**0.745 (0.024)**	**0.0068 (0.0041)**
S. marinus	without the Flemish Cap sample	50	4	0.600 (0.037)	0.0046 (0.0031)
S. marinus	"giant"	37	7	0.757 (0.048)	0.0034 (0.0025)
S. fasciatus		**34**	**4**	**0.223 (0.093)**	**0.0007 (0.0009)**
S. viviparus		**42**	**4**	**0.582 (0.050)**	**0.0019 (0.0017)**
North Atlantic *Sebastes*	**total**	**313**	**23**	**0.845 (0.013)**	**0.0066 (0.0040)**
S. alutus		10	2	0.200 (0.154)	0.0006 (0.0009)
S. capensis		10	3	0.511 (0.164)	0.0019 (0.0018)
Helicolenus dactylopterus		4	3	0.833 (0.222)	0.0029 (0.0028)

The number of haplotypes within the different North Atlantic *Sebastes* samples varied between four (in the samples of *S. fasciatus*, *S. viviparus* and *S. marinus* - without the *S. marinus* sample from the Flemish Cap) and nine (in the sample of *S. mentella*) haplotypes. The sample with the

highest number of individuals, the *S. mentella* sample, revealed also the highest number of haplotypes. In the *S. alutus* sample two and in the *S. capensis* sample as well as in the *Helicolenus dactylopterus* sample three different haplotypes were identified.

A relatively high haplotype diversity of 0.845 (± 0.013 standard error) was observed in the total North Atlantic *Sebastes* data set. The *S. fasciatus* sample revealed the lowest haplotype diversity (0.223), whereas the highest haplotype diversities were observed in the "giant" *S. marinus* sample ($\hat{h}$ = 0.757) as well as in the total *S. marinus* sample ($\hat{h}$ = 0.745; $\hat{h}$ = 0.600 without the Flemish Cap sample). The haplotype diversities observed in the samples of *S. mentella* and *S. viviparus* were intermediate (0.553 and 0.582, respectively). In contrast, the haplotype diversities of the *S. mentella*, *S. viviparus* and *S. marinus* samples were quite similar (between 0.553 and 0.600), when the Flemish Cap sample of *S. marinus* was excluded. In the non-North-Atlantic *Sebastes* samples haplotype diversity varied between 0.200 in *S. alutus* and 0.511 in *S. capensis*, with high standard errors of ± 0.154 and ± 0.164, respectively. In the *Helicolenus dactylopterus* sample a haplotype diversity of 0.833 was observed, also with a high standard error of ± 0.222.

Nucleotide diversity ($\hat{\pi}$) in general was very low, with a high standard error (in all cases, the standard error was of the same order of magnitude as $\hat{\pi}$ itself). Nucleotide diversity observed in the total North Atlantic *Sebastes* data set was 0.0066 (SE ± 0.0040). *S. fasciatus* revealed the lowest nucleotide diversity (0.0007), *S. marinus* the highest (0.0068 including the Flemish Cap sample and 0.0046 excluding the Flemish Cap sample). *S. mentella* and *S. viviparus* showed a similar nucleotide diversity of 0.0019. The "giant" *S. marinus* sample revealed a nucleotide diversity of 0.0034. The nucleotide diversity found in the non-North-Atlantic *Sebastes* samples and the *Helicolenus dactylopterus* sample varied between 0.0006 in *S. alutus* and 0.0029 in *Helicolenus dactylopterus*.

In general, the haplotype diversities and nucleotide diversities within the North Atlantic *Sebastes* samples were lowest in the *S. fasciatus* sample, higher in the *S. viviparus* and *S. mentella* samples and highest in the total *S. marinus* sample. Within the North Atlantic *Sebastes* data set the *S. marinus* and the "giant" *S. marinus* samples were the most genetically diverse samples.

Table 1.5 shows the pairwise p-distances and number of nucleotide differences between the different haplotypes.

The average p-distance between all pairs of the 28 *Sebastes* haplotypes was 2.3% (eight nucleotide substitutions) and ranged from 0.3% (one substitution) to 7.4% (26 substitutions). The highest differences between the *Sebastes* haplotypes were 26 substitutions (between FA2 and CAP3, between FA3 and both CAP2 and CAP3 and between VI4 and both CAP2 and CAP3). The average p-distance of 1.0% between all pairs of North Atlantic *Sebastes* haplotypes was smaller than the overall average and ranged from 0.3% to 2.3% (one to eight nucleotide substitutions). The highest difference (eight nucleotide substitutions) was observed between haplotype VI4 and the haplotypes FA2 and FA3. Between *S. mentella* and *S. marinus* haplotypes

(MA and ME) the differences were small and ranged from one (p = 0.3%) to four (p = 1.1%) nucleotide substitutions.

Differences between haplotypes within one species were at maximum four nucleotide substitutions (p = 1.1%), observed between the *S. marinus* haplotypes MA1 and MA3.

The differences between the *Sebastes* haplotypes and *Helicolenus dactylopterus* (HEL) haplotypes were much higher than the differences observed within the genus *Sebastes* and varied between 29 (p = 8.3%; observed between CAP1 and HEL1) and 40 (p = 11.4%; observed between HEL3 and FA2 and between HEL3 and FA3) substitutions.

The variation in North Atlantic *Sebastes* haplotypes almost exclusively consisted of synonymous substitutions, only at three out of 23 polymorphic sites an amino acid change in the transcribed molecule was observed. One amino acid change (Methionine substituted Valine) occurred in all four *S. viviparus* (VI) haplotypes, one amino acid change in the haplotype MA3 (Isoleucine substituted Valine) and two at the same position in G3 (Alanine substituted Threonine), and G4 (Isoleucine substituted Threonine). Therefore, no more than two amino acid changes were observed between the North Atlantic *Sebastes* haplotypes (between the four *S. viviparus* haplotypes and MA3, G3 and G4, as well as between MA3 and G3 and G4, respectively; see Table 1.6). In case of *S. mentella*, *S. viviparus* and *S. fasciatus* haplotypes, no amino acid changes were observed between haplotypes within one species. In contrast, the *S. marinus* haplotype MA3 differed in one amino acid change from MA1 and MA2. Also G3 and G4 differed in one amino acid change from each other and from the other two haplotypes found in "giant" *S. marinus*, G1 and G2 (see Table 1.6). *S. alutus* haplotypes differed at maximum in two amino acid changes, *S. capensis* and *Helicolenus dactylopterus* haplotypes differed in four to six amino acid changes from the North Atlantic *Sebastes* haplotypes.

The curves of transitions and transversions against p-distance (saturation plots) are presented in Figure 1.4 (only *Sebastes* haplotypes) and 1.5 (all haplotypes).

The saturation plots show that saturation level has not been reached yet in the *Sebastes* haplotypes, since the curves of transitions and transversions were almost linear at sequence divergences smaller than 0.08 (the pairwise p-distances between *Sebastes* haplotypes were at maximum 0.074; see Table 1.5 and Figure 1.4). A slight trend towards asymptotic saturation in transversions and also in transitions was observed among the most divergent haplotypes (with divergences above 0.08 and more), among the haplotypes of the outgroup taxon *Helicolenus dactylopterus* and the *Sebastes* haplotypes (Figure 1.5). As no saturation was observed within the *Sebastes* sequence data set and a slight saturation when the outgroup taxon *Helicolenus dactylopterus* was included, all sequence data were included in the phylogenetic analysis.

Table 1.5. Pairwise p-distances (below the diagonal) and number of nucleotide differences (above the diagonal) between ND3 gene haplotypes. ME = *S. mentella*, MA = *S. marinus*, G = "giant" *S. marinus*, FA = *S. fasciatus*, VI = *S.viviparus*, ALUT = *S. alutus*, CAP = *S. capensis,* HEL = *Helicolenus dactylopterus*.

	ME1	ME2	ME3	ME4	ME5	ME6	ME7	ME8	MA1	MA2	MA3	G1	G2	G3	G4
ME1		1	2	2	3	2	3	2	3	2	3	2	2	3	2
ME2	0.003		1	1	2	1	2	1	2	1	2	1	1	2	1
ME3	0.006	0.003		2	1	2	1	2	1	2	3	2	2	3	2
ME4	0.006	0.003	0.006		3	2	3	2	3	2	3	2	2	3	2
ME5	0.009	0.006	0.003	0.009		3	2	3	2	3	4	3	3	4	3
ME6	0.006	0.003	0.006	0.006	0.009		3	2	3	2	3	2	2	3	2
ME7	0.009	0.006	0.003	0.009	0.006	0.009		3	2	3	4	3	3	4	3
ME8	0.006	0.003	0.006	0.006	0.009	0.006	0.009		3	2	3	2	2	3	2
MA1	0.009	0.006	0.003	0.009	0.006	0.009	0.006	0.009		3	4	3	3	4	3
MA2	0.006	0.003	0.006	0.006	0.009	0.006	0.009	0.006	0.009		1	2	2	3	2
MA3	0.009	0.006	0.009	0.009	0.011	0.009	0.011	0.009	0.011	0.003		3	3	4	3
G1	0.006	0.003	0.006	0.006	0.009	0.006	0.009	0.006	0.009	0.006	0.009		2	1	2
G2	0.006	0.003	0.006	0.006	0.009	0.006	0.009	0.006	0.009	0.006	0.009	0.006		3	2
G3	0.009	0.006	0.009	0.009	0.011	0.009	0.011	0.009	0.011	0.009	0.011	0.003	0.009		3
G4	0.006	0.003	0.006	0.006	0.009	0.006	0.009	0.006	0.009	0.006	0.009	0.006	0.006	0.009	
FA1	0.009	0.006	0.009	0.009	0.011	0.009	0.011	0.009	0.011	0.009	0.011	0.009	0.009	0.011	0.009
FA2	0.011	0.009	0.011	0.011	0.014	0.011	0.014	0.011	0.014	0.011	0.014	0.011	0.011	0.014	0.011
FA3	0.011	0.009	0.011	0.011	0.014	0.011	0.014	0.011	0.014	0.011	0.014	0.011	0.011	0.014	0.011
FA4	0.006	0.003	0.006	0.006	0.009	0.006	0.009	0.006	0.009	0.006	0.009	0.006	0.006	0.009	0.006
VI1	0.014	0.011	0.014	0.014	0.017	0.014	0.017	0.014	0.017	0.014	0.017	0.014	0.014	0.017	0.014
VI2	0.011	0.009	0.011	0.011	0.014	0.011	0.014	0.011	0.014	0.011	0.014	0.011	0.011	0.014	0.011
VI3	0.014	0.011	0.014	0.014	0.017	0.014	0.017	0.009	0.017	0.014	0.017	0.014	0.014	0.017	0.014
VI4	0.017	0.014	0.017	0.017	0.020	0.017	0.020	0.017	0.020	0.017	0.020	0.017	0.017	0.020	0.017
ALUT1	0.026	0.023	0.026	0.026	0.028	0.026	0.023	0.026	0.028	0.026	0.023	0.026	0.026	0.028	0.026
ALUT2	0.028	0.026	0.028	0.028	0.031	0.028	0.026	0.028	0.031	0.028	0.026	0.028	0.028	0.031	0.028
CAP1	0.066	0.063	0.066	0.066	0.068	0.066	0.063	0.066	0.063	0.060	0.057	0.066	0.066	0.068	0.066
CAP2	0.068	0.066	0.068	0.068	0.071	0.068	0.066	0.068	0.066	0.063	0.060	0.068	0.068	0.071	0.068
CAP3	0.068	0.066	0.068	0.068	0.071	0.068	0.066	0.068	0.066	0.063	0.060	0.068	0.068	0.071	0.068
HEL1	0.105	0.103	0.100	0.105	0.103	0.105	0.097	0.105	0.097	0.100	0.097	0.105	0.100	0.108	0.105
HEL2	0.103	0.100	0.097	0.103	0.100	0.103	0.094	0.103	0.094	0.097	0.100	0.103	0.097	0.105	0.103
HEL3	0.108	0.105	0.103	0.108	0.105	0.108	0.100	0.108	0.100	0.103	0.100	0.108	0.103	0.111	0.108

Table 1.5 (continued). Pairwise p-distances (below the diagonal) and number of nucleotide differences (above the diagonal) between ND3 gene haplotypes. ME = *S. mentella*, MA = *S. marinus*, G = "giant" *S. marinus*, FA = *S. fasciatus*, VI = *S.viviparus*, ALUT = *S. alutus*, CAP = *S. capensis,* HEL = *Helicolenus dactylopterus*.

	FA1	FA2	FA3	FA4	VI1	VI2	VI3	VI4	ALUT1	ALUT2	CAP1	CAP2	CAP3	HEL1	HEL2	HEL3
ME1	3	4	4	2	5	4	5	6	9	10	23	24	24	37	36	38
ME2	2	3	3	1	4	3	4	5	8	9	22	23	23	36	35	37
ME3	3	4	4	2	5	4	5	6	9	10	23	24	24	35	34	36
ME4	3	4	4	2	5	4	5	6	9	10	23	24	24	37	36	38
ME5	4	5	5	3	6	5	6	7	10	11	24	25	25	36	35	37
ME6	3	4	4	2	5	4	5	6	9	10	23	24	24	37	36	38
ME7	4	5	5	3	6	5	6	7	8	9	22	23	23	34	33	35
ME8	3	4	4	2	5	4	3	6	9	10	23	24	24	37	36	38
MA1	4	5	5	3	6	5	6	7	10	11	22	23	23	34	33	35
MA2	3	4	4	2	5	4	5	6	9	10	21	22	22	35	34	36
MA3	4	5	5	3	6	5	6	7	8	9	20	21	21	34	35	35
G1	3	4	4	2	5	4	5	6	9	10	23	24	24	37	36	38
G2	3	4	4	2	5	4	5	6	9	10	23	24	24	35	34	36
G3	4	5	5	3	6	5	6	7	10	11	24	25	25	38	37	39
G4	3	4	4	2	5	4	5	6	9	10	23	24	24	37	36	38
FA1		1	1	1	6	5	6	7	10	11	24	25	25	38	37	39
FA2	0.003		2	2	7	6	7	8	11	12	25	24	26	39	38	40
FA3	0.003	0.006		2	7	6	7	8	11	12	25	26	26	39	38	40
FA4	0.003	0.006	0.006		5	4	5	6	9	10	23	24	24	37	36	38
VI1	0.017	0.020	0.020	0.014		1	2	1	10	11	24	25	25	33	32	34
VI2	0.014	0.017	0.017	0.011	0.003		1	2	9	10	23	24	24	34	33	35
VI3	0.017	0.020	0.020	0.014	0.006	0.003		3	10	11	24	25	25	35	34	36
VI4	0.020	0.023	0.023	0.017	0.003	0.006	0.009		11	12	25	26	26	34	33	35
ALUT1	0.028	0.031	0.031	0.026	0.028	0.026	0.028	0.031		1	19	20	20	35	36	36
ALUT2	0.031	0.034	0.034	0.028	0.031	0.028	0.031	0.034	0.003		18	19	19	34	35	35
CAP1	0.068	0.071	0.071	0.066	0.068	0.066	0.068	0.071	0.054	0.051		1	1	29	30	30
CAP2	0.071	0.068	0.074	0.068	0.071	0.068	0.071	0.074	0.057	0.054	0.003		2	30	31	31
CAP3	0.071	0.074	0.074	0.068	0.071	0.068	0.071	0.074	0.057	0.054	0.003	0.006		30	31	31
HEL1	0.108	0.111	0.111	0.105	0.094	0.097	0.100	0.097	0.100	0.097	0.083	0.085	0.085		1	1
HEL2	0.105	0.108	0.108	0.103	0.091	0.094	0.097	0.094	0.103	0.100	0.085	0.088	0.088	0.003		2
HEL3	0.111	0.114	0.114	0.108	0.097	0.100	0.103	0.100	0.103	0.100	0.085	0.088	0.088	0.003	0.006	

Table 1.6. Number of amino acid changes between ND3 gene haplotypes. ME = *S. mentella*, MA = *S. marinus*, G = "giant" *S. marinus*, FA = *S. fasciatus*, VI = *S. viviparus*, ALUT = *S. alutus*, CAP = *S. capensis,* HEL = *Helicolenus dactylopterus*.

	ME1	ME2	ME3	ME4	ME5	ME6	ME7	ME8	MA1	MA2	MA3	G1	G2	G3	G4
ME1															
ME2	0														
ME3	0	0													
ME4	0	0	0												
ME5	0	0	0	0											
ME6	0	0	0	0	0										
ME7	0	0	0	0	0	0									
ME8	0	0	0	0	0	0	0								
MA1	0	0	0	0	0	0	0	0							
MA2	0	0	0	0	0	0	0	0	0						
MA3	1	1	1	1	1	1	1	1	1	1					
G1	0	0	0	0	0	0	0	0	0	0	1				
G2	0	0	0	0	0	0	0	0	0	0	1	0			
G3	1	1	1	1	1	1	1	1	1	1	2	1	1		
G4	1	1	1	1	1	1	1	1	1	1	2	1	1	1	
FA1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1
FA2	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1
FA3	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1
FA4	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1
VI1	1	1	1	1	1	1	1	1	1	1	2	1	1	2	2
VI2	1	1	1	1	1	1	1	1	1	1	2	1	1	2	2
VI3	1	1	1	1	1	1	1	1	1	1	2	1	1	2	2
VI4	1	1	1	1	1	1	1	1	1	1	2	1	1	2	2
ALUT1	1	1	1	1	1	1	1	1	1	1	0	1	1	2	2
ALUT2	1	1	1	1	1	1	1	1	1	1	0	1	1	2	2
CAP1	4	4	4	4	4	4	4	4	4	4	3	4	4	5	5
CAP2	4	4	4	4	4	4	4	4	4	4	3	4	4	5	5
CAP3	5	5	5	5	5	5	5	5	5	5	4	5	5	6	6
HEL1	5	5	5	5	5	5	5	5	5	5	4	5	5	6	6
HEL2	4	4	4	4	4	4	4	4	4	4	5	4	4	5	5
HEL3	5	5	5	5	5	5	5	5	5	5	4	5	5	6	6

Table 1.6 (continued). Number of amino acid changes between ND3 gene haplotypes. ME = *S. mentella*, MA = *S. marinus*, G = "giant" *S. marinus*, FA = *S. fasciatus*, VI = *S. viviparus*, ALUT = *S. alutus*, CAP = *S. capensis,* HEL = *Helicolenus dactylopterus*.

	FA1	FA2	FA3	FA4	VI1	VI2	VI3	VI4	ALUT1	ALUT2	CAP1	CAP2	CAP3	HEL1	HEL2
ME1															
ME2															
ME3															
ME4															
ME5															
ME6															
ME7															
ME8															
MA1															
MA2															
MA3															
G1															
G2															
G3															
G4															
FA1															
FA2	0														
FA3	0	0													
FA4	0	0	0												
VI1	1	1	1	1											
VI2	1	1	1	1	0										
VI3	1	1	1	1	0	0									
VI4	1	1	1	1	0	0	0								
ALUT1	1	1	1	1	2	2	2	2							
ALUT2	1	1	1	1	2	2	2	2	0						
CAP1	4	4	4	4	3	3	3	3	3	3					
CAP2	4	4	4	4	3	3	3	3	3	3	0				
CAP3	5	5	5	5	4	4	4	4	4	4	1	1			
HEL1	5	5	5	5	5	5	5	5	4	4	4	4	5		
HEL2	4	4	4	4	4	4	4	4	5	5	5	5	6	1	
HEL3	5	5	5	5	5	5	5	5	4	4	4	4	5	0	1

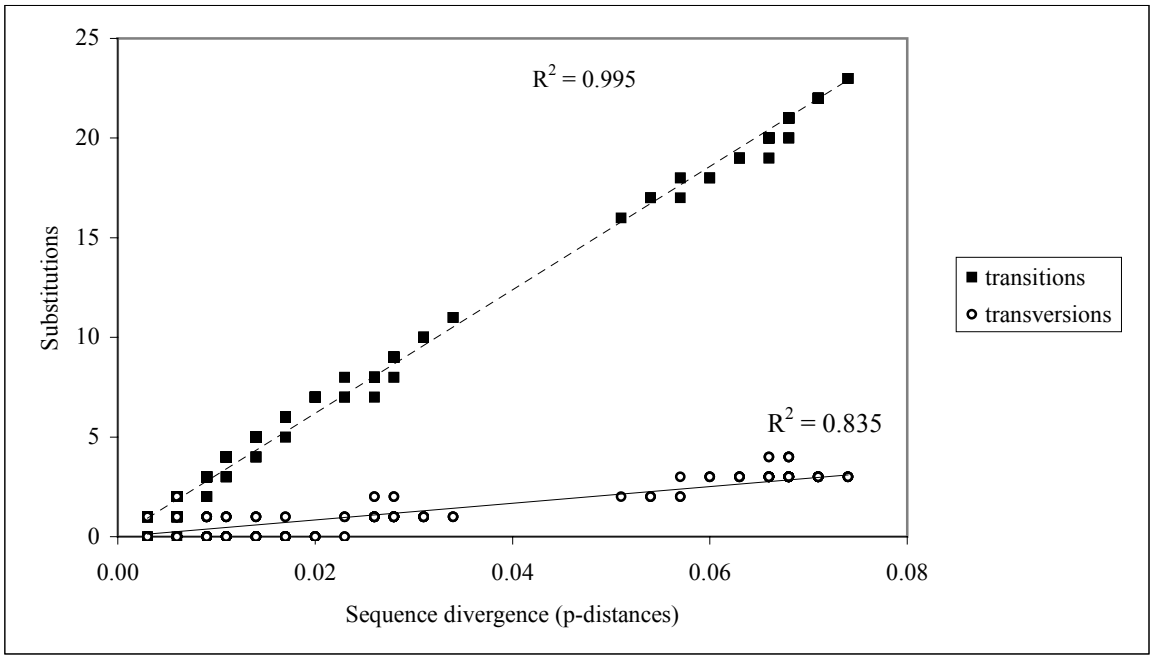


Figure 1.4. Plot of the number of transitions (black squares) and transversions (white circles) against uncorrected sequence divergence (p-distances) of pairwise comparisons among all 28 *Sebastes* haplotypes. The slope and the coefficient of determination (R^2) are depicted for transitions and transversions.

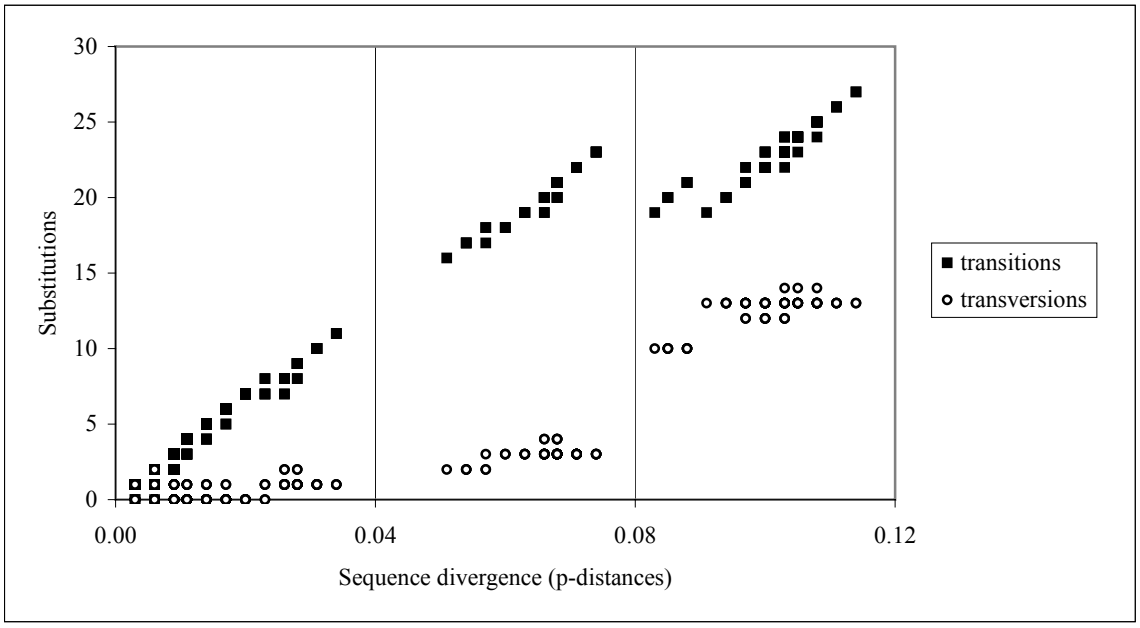


Figure 1.5. Plot of the number of transitions (black squares) and transversions (white circles) against uncorrected sequence divergence (p-distances) of pairwise comparisons among all 31 haplotypes. Vertical gridlines indicate the sequence divergence among groups of species. p-distances between 0.00 and 0.04: Comparisons among North Atlantic *Sebastes* and *S. alutus* haplotypes; p-distances between 0.04 and 0.08: *S. capensis* haplotype comparisons; p-distances above 0.08: *Helicolenus dactylopterus* (outgroup) haplotype comparisons.

4.2 Haplotype frequencies and distribution

Haplotype frequencies and haplotype distribution in the samples are shown in Table 1.7. In case of MEoc, MEds, FAFC and VISWIc, samples from different years were pooled (for sample codes see Table 1.1). In Table 1.8 the haplotype distribution in the samples of "deep-sea" and "oceanic" *S. mentella* collected in different years is given separately for each sample. On average, one to two haplotypes per species sample occurred in high frequencies. The remaining haplotypes were rare and occurred in low frequencies.

Table 1.7 shows that haplotypes observed in the samples of *S. viviparus*, *S. alutus*, *S. capensis* and *Helicolenus dactylopterus* did not occur in samples of the other species. For *S. viviparus* four specific haplotypes (of which VI1 and VI2 were the most common), for *S. alutus* two specific haplotypes, and for *S. capensis* as well as for *Helicolenus dactylopterus* three specific haplotypes were observed (see also Table 1.4).

Also, most of the haplotypes found in the samples of *S. mentella, S. marinus* and *S. fasciatus* were restricted to these species. Only three haplotypes (ME2, FA1 and FA3) occurred in more than one species.

S. mentella samples revealed three common haplotypes, ME1, ME2 and ME3, with absolute frequencies of 15, 82 and 26, respectively, accompanied by five rare variants (ME4 through ME8) which appeared in one or two individuals only.

In almost all *S. mentella* samples (Norway, Irminger Sea "oceanic" type, Iceland, Greenland, Flemish Cap) ME2 was the most abundant haplotype, except for the "deep-sea" *S. mentella* sample from the Irminger Sea. In this sample the haplotype ME3, which was recorded in only two other individuals (in the Irminger Sea "oceanic" type sample and the Flemish Cap sample), appeared in 24 of 41 individuals analysed. Table 1.8 shows that ME3 appeared in the "deep-sea" *S. mentella* samples from 1997 as well as in the sample from 2001.

In "oceanic" *S. mentella*, ME2 was the most abundant haplotype, but also ME1 was found in 12 of 41 individuals, while this haplotype appeared relatively rarely in any of the other samples of *S. mentella* (only three individuals of the samples MENO01 and MEEGr01 carried this haplotype). Table 1.8 reveals that ME1 was found in two of three samples of "oceanic" *S. mentella* collected in different years. The individuals collected in 1997 exclusively carried ME2.

ME2 was not only the most abundant haplotype in *S. mentella* but also in the entire data set: 82 out of 130 individuals of *S. mentella* revealed ME2, in total 99 individuals carried this haplotype. In two of the 70 *S. marinus* specimens ME2 was observed: In one individuals in the *S. marinus* sample from the Dohrn Bank/Greenland (MADB01) and in one individual in the *S. marinus* sample from Southeast Iceland collected in 2001 (MASEIc01).

ME2 was also found in 15 individuals of the "giant" *S. marinus* sample (see below).

Except for ME2, all other seven haplotypes (ME1 and ME3 through ME8) exclusively appeared in *S. mentella*.

Three haplotypes, MA1, MA2 and MA3, were exclusively found in *S. marinus*, the latter two also in "giant" *S. marinus*. Overall, MA2 was the most common *S. marinus* haplotype, it was found in 34 individuals of the 70 *S. marinus* specimens analysed. Also MA1 appeared in high frequencies (in 21 individuals), but was found only in the samples from Iceland and Greenland (MASWIc97, MASEIc01, MADB01 and MEEGr01). In the sample from East Greenland, all ten individuals carried MA1. In contrast, MA3 was relatively rare and appeared only in the "giant" *S. marinus* sample and in the *S. marinus* sample from the Dohrn Bank, MADB01.

In the *S. fasciatus* samples, four haplotypes were found: The most abundant haplotype FA1 and three rather rare haplotypes (FA2, FA3 and FA4), which occurred with low frequencies. FA1 and FA3 were also found in samples of *S. mentella* and *S. marinus* from the Flemish Cap. Whereas only one individual carrying FA1 occurred in the *S. mentella* sample from the Flemish Cap, all twenty individuals of the *S. marinus* sample from the Flemish Cap revealed FA1 or FA3, haplotypes that were observed almost exclusively in *S. fasciatus*.

Also in the sample of "giant" *S. marinus,* haplotypes from different species appeared in higher frequencies: Out of 37 individuals, morphologically identified as "giant" *S. marinus*, fifteen individuals carried haplotype ME2, the most abundant haplotype in the *S. mentella* samples. Fourteen individuals showed the haplotypes MA2 and MA3, exclusively found in *S. marinus* samples. Only eight individuals showed unique haplotypes, G1, G2, G3 and G4, which appeared exclusively in the "giant" *S. marinus* sample.

Table 1.7. Distribution of ND3 gene haplotypes of *Sebastes* and *Helicolenus* for species and population samples. Sample codes are explained in Table 1.1. In case of MEoc, MEds, FAFC and VISWIc samples from different years were pooled.

Sample	Haplotype ME1	ME2	ME3	ME4	ME5	ME6	ME7	ME8	MA1	MA2	MA3	G1	G2	G3	G4	FA1	FA2	FA3	FA4
MEoc	12	27	1	1															
MEds		16	24		1														
MENO01	2	10				2													
MESWIc01		9					1												
MEEGr01	1	9																	
MEFC01		11	1					1								1			
MANO01										10									
MASWIc97									6	5									
MASEIc01		1							4	5									
MADB01		1							1	4	3								
MAEGr01									10										
MAFC01																16		4	
MAG96		15								10	4	3	2	2	1				
FAFC																30	2	1	1
VINO01																			
VISWIc																			
ALUT																			
CAP																			
HEL																			
Total	15	99	26	1	1	2	1	1	21	34	7	3	2	2	1	47	2	5	1

Table 1.7. (continued)

Sample	Haplotype VI1	VI2	VI3	VI4	ALUT1	ALUT2	CAP1	CAP2	CAP3	HEL1	HEL2	HEL3	*n*
MEoc													41
MEds													41
MENO01													14
MESWIc01													10
MEEGr01													10
MEFC01													14
MANO01													10
MASWIc97													11
MASEIc01													10
MADB01													9
MAEGr01													10
MAFC01													20
MAG96													37
FAFC													34
VINO01	3	11											14
VISWIc	12	12	2	2									28
ALUT					9	1							10
CAP							2	7	1				10
HEL										2	1	1	4
Total	15	23	2	2	9	1	2	7	1	2	1	1	337

Table 1.8. Distribution of ND3 gene haplotypes for the samples of "oceanic" (MEoc) and "deep-sea" (MEds) *S. mentella* by sampling year. Sample codes are explained in Table 1.1.

	Haplotype								
Sample	ME1	ME2	ME3	ME4	ME5	ME6	ME7	ME8	*n*
MEoc96	6	6							12
MEoc97		14							14
MEoc01	6	7	1	1					15
MEds97		10	16						26
MEds01		6	8		1				15

4.3 Phylogenetic analysis

In the **neighbour-joining analyses**, similar tree topologies were obtained for the three different distance measure indices (p-distance, Kimura 2-parameter model and Hasegawa-Kishino-Yano model). Consequently, only the tree based on the matrix of pairwise p-distances is presented here (Figure 1.6; the distance matrix is given in Table 1.5, see paragraph 4.1 for details).

The neighbour-joining tree based on pairwise p-distances revealed four main clusters. The first cluster consisted of the three *S. capensis* haplotypes, the second consisted of the two *S. alutus* haplotypes, the third cluster included all 23 North Atlantic *Sebastes* haplotypes and cluster four was comprised of the three haplotypes of the outgroup taxon *Helicolenus dactylopterus*. These four clusters were supported by high bootstrap values between 89 and 100 in the NJ tree. The branch separating the haplotypes of the outgroup *Helicolenus dactylopterus* from the *Sebastes* haplotypes was long and well supported (bootstrap value of 100).

Within the *Sebastes* cluster, the *S. capensis* haplotypes clustered together in a well-supported clade (bootstrap value of 100), separated from the other *Sebastes* haplotypes, whereas the *S. alutus* haplotypes grouped together with the North Atlantic haplotypes (bootstrap value of 98). Nevertheless, the *S. alutus* haplotypes and the North Atlantic *Sebastes* haplotypes each formed an independent clade, supported by high bootstrap values of 95 and 89, respectively.

Compared to the divergence from the *S. alutus* and *S. capensis* haplotype clusters, the phylogenetic structure within the North Atlantic *Sebastes* haplotype cluster was very shallow, with small branch lengths. Within this cluster the *S. viviparus* haplotypes were more clearly separated from the haplotypes of *S. fasciatus*, *S. mentella*, *S. marinus* and "giant" *S. marinus*, supported by a bootstrap value of 77.

Within the remaining haplotypes of *S. fasciatus*, *S. mentella* and *S. marinus* and "giant" *S. marinus*, only the four *S. fasciatus* haplotypes formed a separate cluster (bootstrap value of 63). In contrast, there was no clear separation of *S. mentella*, *S. marinus* and "giant" *S. marinus* haplotypes, with small branch lengths and no bootstrap support at the corresponding

nodes (bootstrap values of 50 and less). Even though a clustering of the four haplotypes ME3, ME5, ME7 and MA1, of MA2 and MA3 as well as of G1 and G3 was observed, only the grouping of the two haplotypes G1 and G3 was supported by a bootstrap value of 65.

The **maximum parsimony analysis** resulted in four equally parsimonious trees each with a total tree length of 77 steps, a relatively high consistency index (CI) of 0.792, a corresponding relatively low homoplasy index (HI) of 0.208, a high retention index (RI) of 0.909 and a rescaled consistency index (RC) of 0.720. As the strict consensus tree of these four most parsimonious reconstructions had the same topology and polytomies as the 50% majority-rule consensus tree obtained by bootstrapping (1000 replicates), only the 50% majority rule consensus tree is presented here (tree length: 79 steps; CI = 0.772; HI = 0.228; RI = 0.898; RC = 0.693; see Figure 1.7).
The 50% majority-rule consensus tree, as well as the strict consensus tree, revealed four major clusters similar to the groups in the neighbour-joining tree, one with the *S. capensis* haplotypes, one with the *S. alutus* haplotypes, one with the North Atlantic *Sebastes* haplotypes and one with the haplotypes of the outgroup taxon *Helicolenus dactylopterus*. These clusters were supported by bootstrap values between 65 and 100 in the 50% majority rule consensus tree (Figure 1.7). The *Helicolenus dactylopterus* haplotypes formed a distinct clade (bootstrap value of 100), separated from the *Sebastes* haplotypes. Within the *Sebastes* haplotypes, the *S. capensis* haplotypes grouped in a distinct well-supported cluster (bootstrap value of 100). The *S. alutus* haplotypes grouped with the North Atlantic *Sebastes* haplotypes (bootstrap value of 96) and both groups of haplotypes formed distinct clades within this cluster, both with a weaker bootstrap support (bootstrap values of 65 and 68, respectively).
Within the North Atlantic cluster, *S. viviparus* haplotypes formed one cluster and the *S. fasciatus* haplotypes formed another, both clusters supported by weaker bootstrap values between 60 (*S. fasciatus*) and 70 (*S. viviparus*). The North Atlantic *Sebastes* cluster was dominated by a large polytomy consisting of the *S. mentella*, *S. marinus* and "giant" *S. marinus* haplotypes and the *S. viviparus* and *S. fasciatus* clusters. Congruent with the neighbour-joining analysis, there was no clear clustering of the *S. mentella*, *S. marinus* and "giant" *S. marinus* haplotypes. The 50% majority rule consensus tree revealed no bootstrap support at most of the corresponding nodes (bootstrap values 50 and less). Only the two haplotypes G1 and G3 clustered together supported by a bootstrap value of 67, as well as the haplotypes ME3, ME5, ME7 and MA1. Although the latter haplotype relationship appeared in all four most parsimonious trees (not presented here), it had weak bootstrap support (bootstrap value of 53; see Figure 1.7). Both groupings were also found in the neighbour-joining tree.

The **quartet puzzling maximum likelihood analysis** of the data resulted in a tree topology congruent with that of the neighbour-joining and maximum parsimony analyses. Both models, the model of Felsenstein (F81; Felsenstein 1981) and the Hasegawa-Kishino-Yano model

(HKY85; Hasegawa et al. 1985), used in the quartet puzzling maximum likelihood analysis approach produced similar tree topologies. Therefore, only the quartet puzzling tree based on the latter model is presented here (Figure 1.8). Similar to the other phylogenetic analyses, the quartet puzzling tree revealed four major clusters, consisting of (i) the *S. capensis* haplotypes, (ii) the *S. alutus* haplotypes, (iii) the North Atlantic *Sebastes* haplotypes and (iv) the haplotypes of the outgroup taxon *Helicolenus dactylopterus* - as well as low resolution within the North Atlantic *Sebastes* cluster.

Within the *Sebastes* cluster, the *S. capensis* haplotypes grouped in a distinct cluster, supported by 80% of the puzzling replicates (= reliability value). The *S. alutus* haplotypes grouped with the North Atlantic *Sebastes* haplotypes (reliability value: 63%) and both groups of haplotypes formed a distinct clade, with a strongly supported *S. alutus* cluster (reliability value: 97%) and a weakly supported North Atlantic *Sebastes* cluster (reliability value: 54%).

Within the North Atlantic *Sebastes* cluster, the *S. fasciatus* haplotypes grouped together in a highly supported cluster (reliability value: 91%). The *S. viviparus* haplotypes also grouped together, but with a low reliability value (54%). The haplotype ME8 grouped with the *S. viviparus* cluster, but this grouping was not supported by the corresponding reliability value (less than 50%).

The grouping of G1 and G3, observed in the NJ and MP trees (Figure 1.6 and 1.7), appeared also in the quartet puzzling tree, with a high reliability value of 86%.

Also the haplotypes ME3, ME5, ME7 and MA1 formed a well-supported cluster (reliability value of 82%). This group was also observed in the NJ and MP trees, but in both cases with no or low bootstrap support for this grouping.

In addition, the haplotypes MA and MA3 grouped together, supported by a reliability value of 86%. This grouping was also observed in the neighbour-joining analysis, but without meaningful bootstrap support for this cluster. The remaining groupings of *S. mentella*, *S. marinus* and "giant" *S. marinus* haplotypes were not well supported in the quartet puzzling tree, with reliability values of less than 50%.

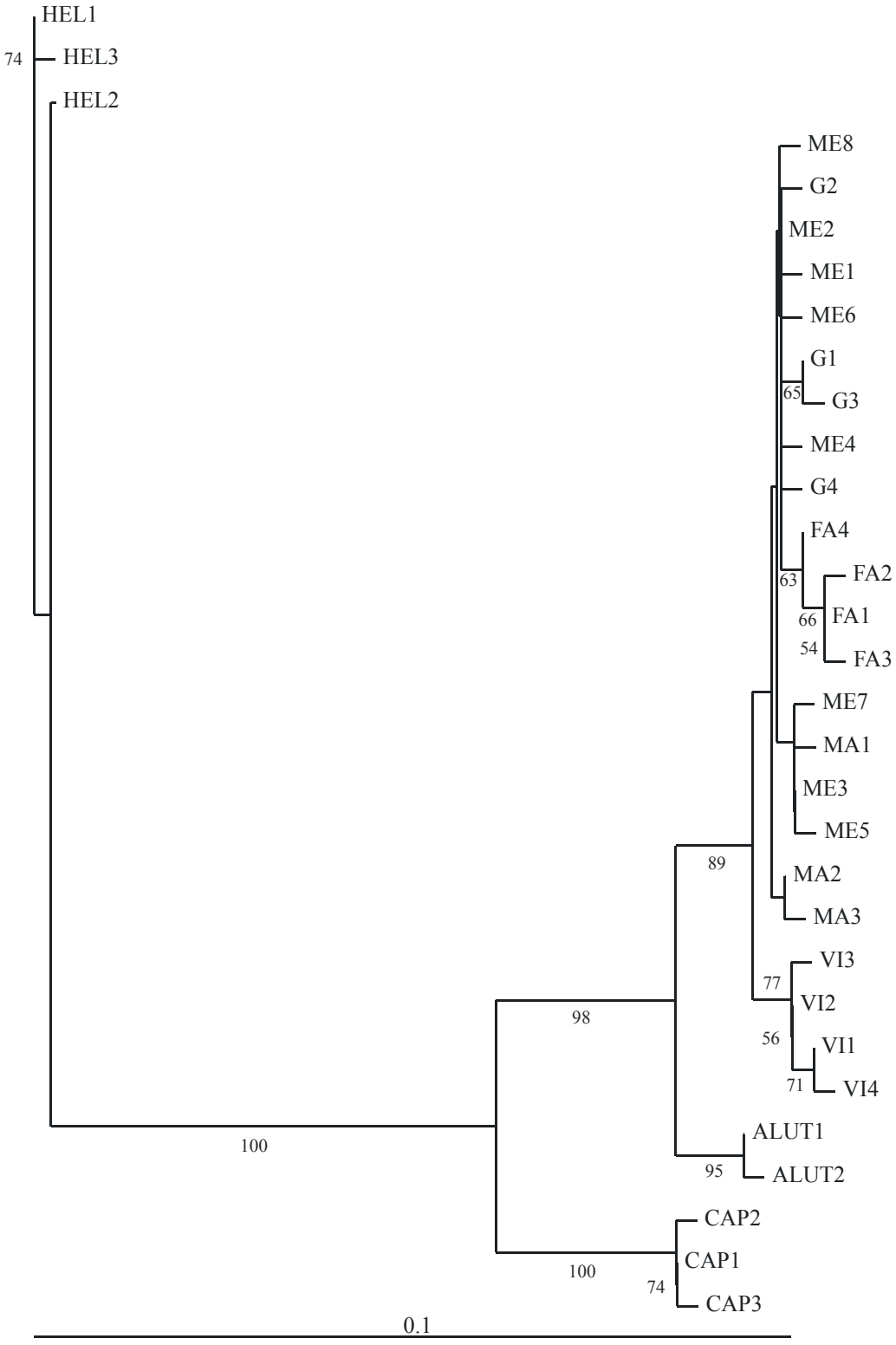


Figure 1.6. Neighbour-joining dendrogram using pairwise p-distances based on ND3 gene haplotypes of *Sebastes spp.* and *Helicolenus dactylopterus.* Numbers represent bootstrap support values of 1000 replicates. Only bootstrap values greater than 50 are shown. Haplotypes are designated by names as defined in Table 1.5.

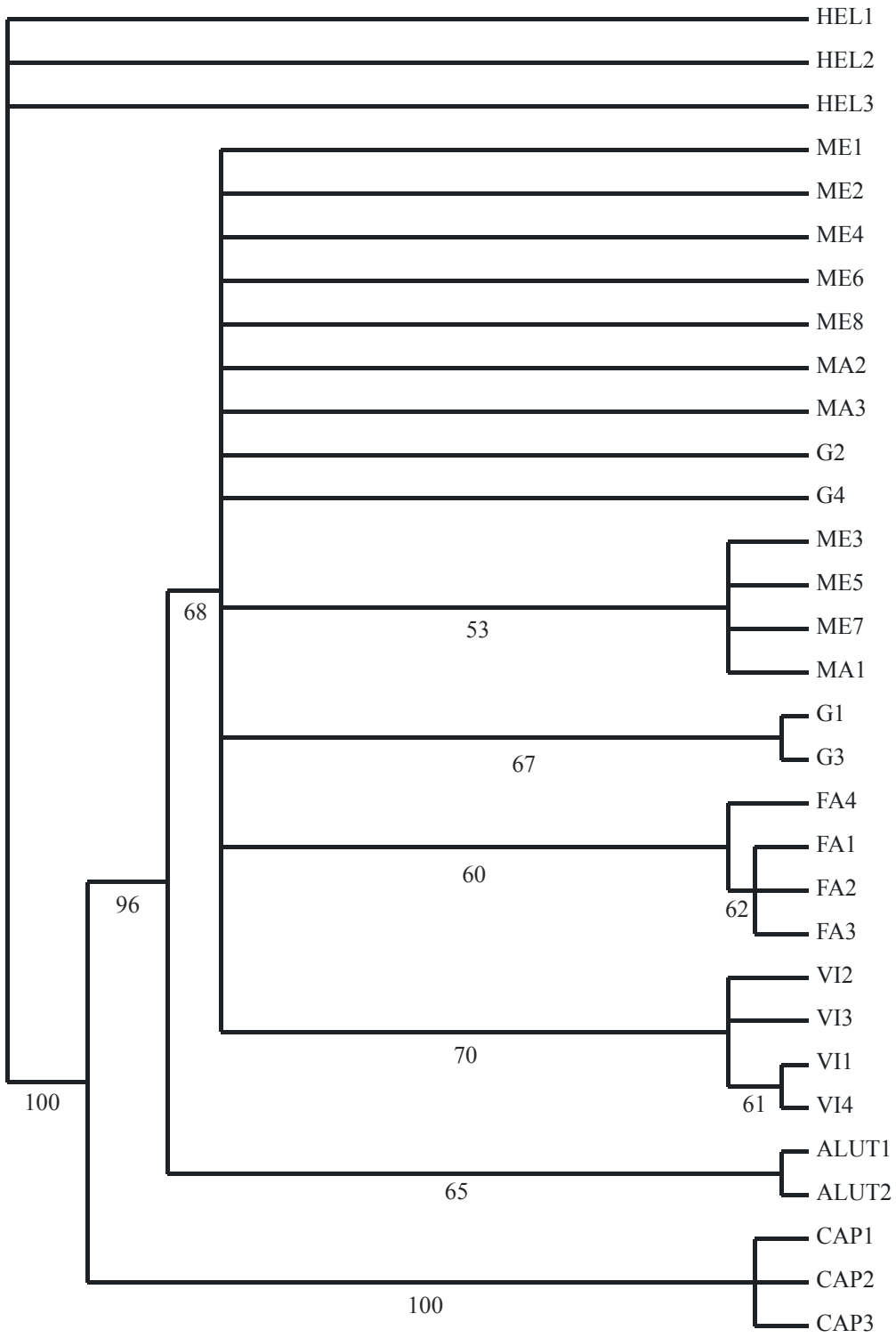


Figure 1.7. 50% majority rule consensus tree of 1000 bootstrap replicates of the data set based on ND3 gene haplotypes of *Sebastes spp.* and *Helicolenus dactylopterus*. Numbers represent bootstrap support values of 1000 replicates. Only bootstrap values greater than 50 are shown. Tree length: 79 steps; CI = 0.772; HI = 0.228; RI = 0.898; RC = 0.693. Haplotypes are designated by names as defined in Table 1.5.

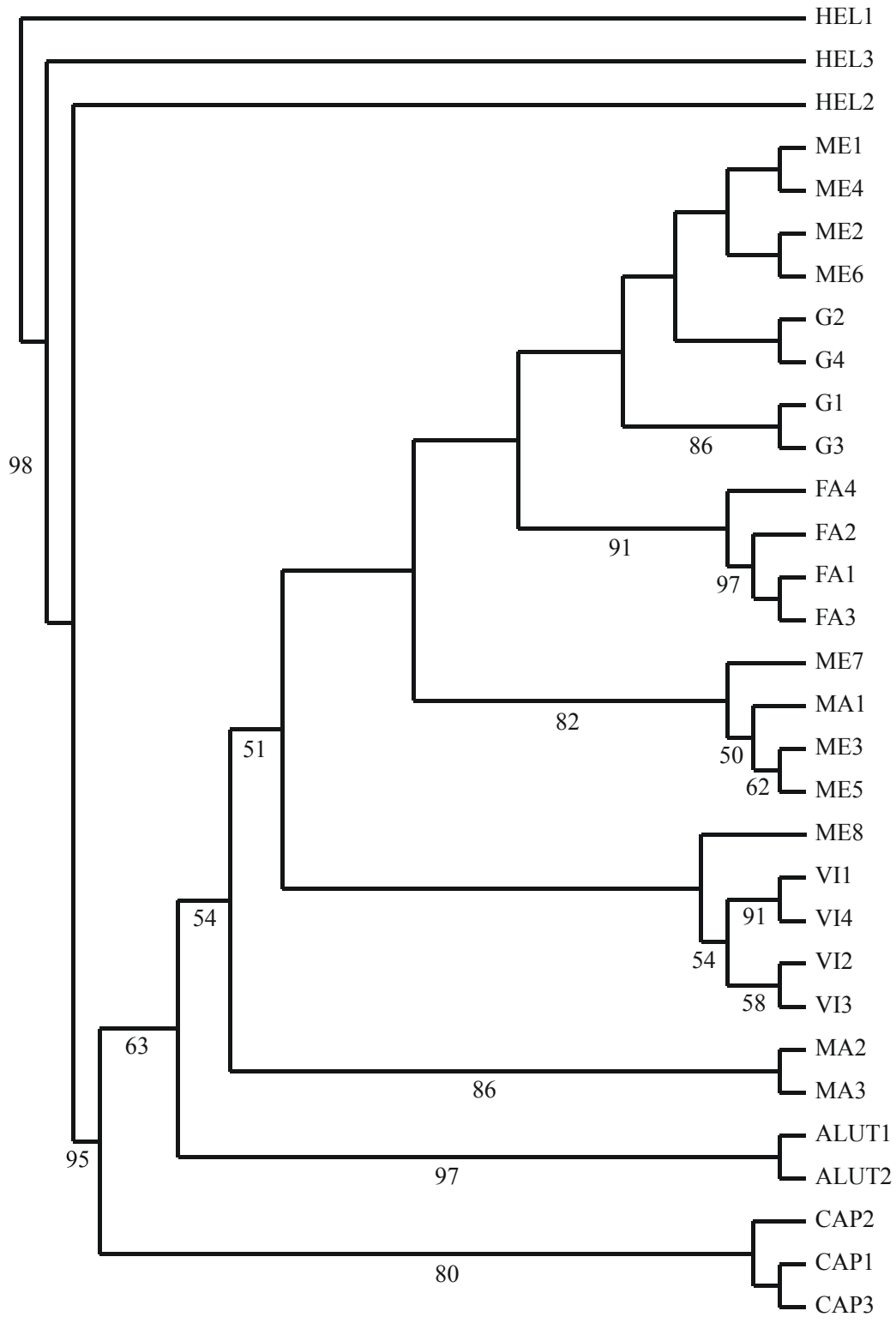


Figure 1.8. Quartet puzzling tree of ND3 gene haplotypes of *Sebastes spp.* and *Helicolenus dactylopterus*, obtained by a search procedure of 1000 puzzling steps, using the Hasegawa-Kishino-Yano substitution model (HKY85; Hasegawa et al. 1985). The numbers at the internal branches are percentages of puzzling steps, which support the corresponding branch (= reliability values). Only reliability values equal to or greater than 50% are shown. Haplotypes are designated by names as defined in Table 1.5.

As the phylogenetic structure within the North Atlantic *Sebastes* cluster could not be resolved by the different phylogenetic trees, it was represented by means of a network.

The results of the **statistical parsimony network analysis** based on statistical parsimony (Templeton et al. 1992) are presented in Figure 1.9, showing the genealogical relationships of the North Atlantic *Sebastes* haplotypes. The *S. alutus*, *S. capen*sis and *Helicolenus dactylopterus* haplotypes were not included in the TCS analysis, because the pairwise distances (= the number of mutational steps) between these haplotypes and the North Atlantic *Sebastes* haplotypes were too high for the TCS analysis to include them in the statistical parsimony network.

The relationships of the 23 North Atlantic *Sebastes* haplotypes were established within the limits of parsimony criteria.

The statistical parsimony network revealed a star-like topology with the most common haplotype ME2 in the centre and the other haplotypes radiating.

ME2 was connected to each of the haplotypes ME1, ME3, ME4, ME6, ME8, MA2, G1, G2, G4 and FA4 by one mutation.

The *S. viviparus* and *S. fasciatus* haplotypes formed two subclusters, also with the most common haplotype in the centre (VI2 and FA1, respectively) and the rare haplotypes radiating.

All the haplotypes were linked by one step except for the *S. viviparus* specific haplotypes, which were linked to the others by at least two mutations. The *S. viviparus* cluster was connected to the central haplotype ME2 by two possible pathways. In both cases, this assumes two intermediate haplotypes that were not represented within the data set, (i) between haplotype ME2 and VI2, and (ii) between haplotype ME8 and VI3.

The most distant haplotypes FA3 and VI4 differed by at least eight mutation steps (see also Table 1.5).

The most common *S. marinus* haplotype MA2 was directly linked to ME2 by one mutation step; MA3 was connected to ME2 through MA2.

MA1, which was also quite common in the *S. marinus* samples, was not directly connected with the other *S. marinus* haplotypes, MA2 and MA3, but with ME2 through haplotype ME3. The same applies to the rare haplotypes ME5 and ME7, which were also linked to ME2 by ME3.

Therefore, MA1 was separated from the other two *S. marinus* haplotypes MA2 and MA3 by three and four mutation steps, respectively.

The rare "giant" *S. marinus* haplotypes radiated directly from ME2 (except G3, which was connected to ME2 through G1), but were not directly connected to the *S. marinus* haplotypes MA1, MA2 and MA3. The most distant "giant" *S. marinus* haplotype G3 was separated from MA2 and MA3 by three and four mutation steps, respectively.

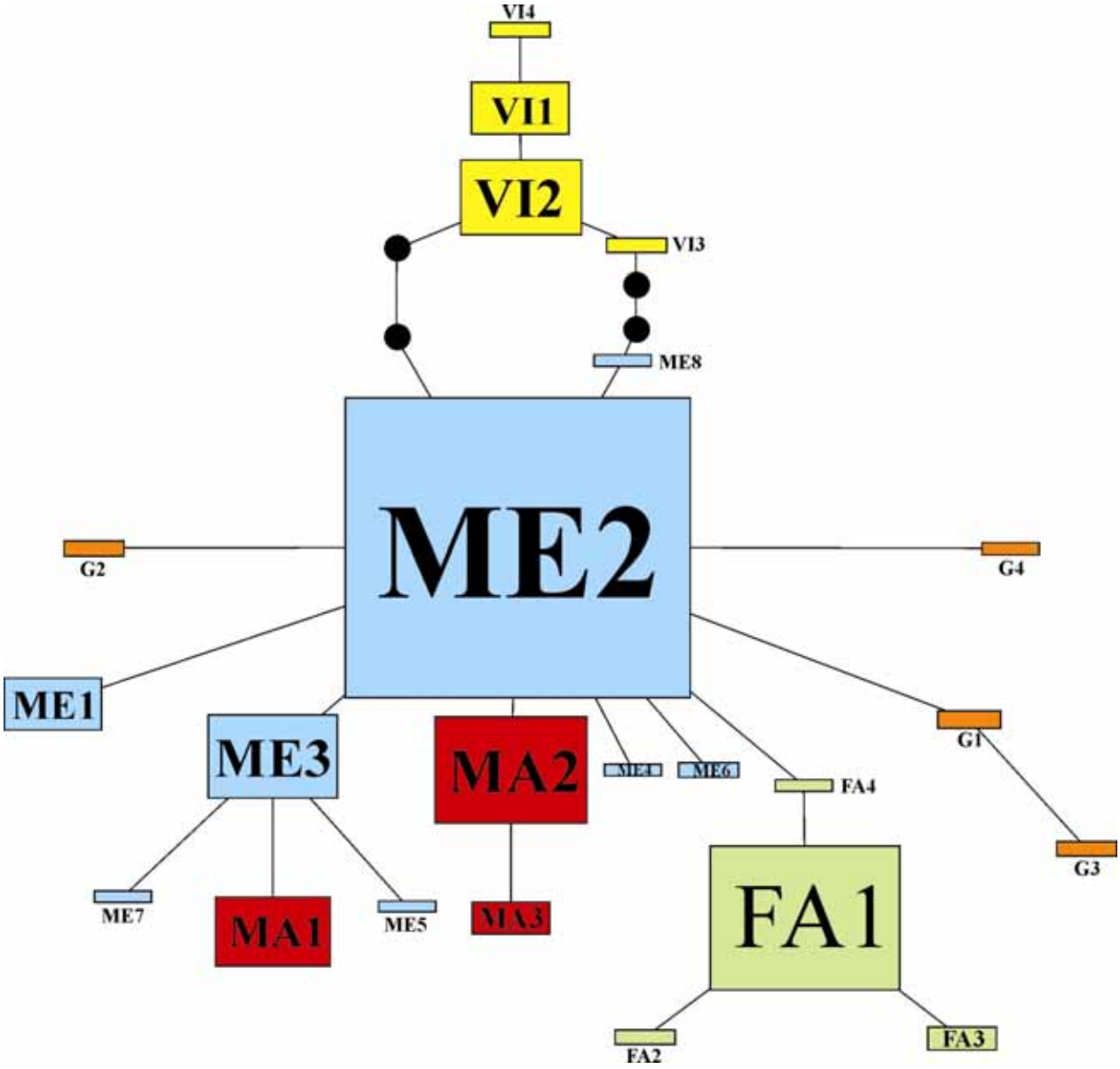


Figure 1.9. Statistical parsimony network showing the genealogical relationships of the ND3 haplotypes of the North Atlantic *Sebastes* species. Each rectangle represents one haplotype. Haplotypes are designated by names as defined in Table 1.5. Each line represents a single mutational step connecting two haplotypes and small black circles represent hypothetical (not detected) haplotypes. The size of the rectangles corresponds to the haplotype frequency.

4.4 AMOVA (analysis of molecular variance) and Φ_{ST}

The results of the analysis of molecular variance (AMOVA) are given in Table 1.9.
The AMOVA revealed that genetic variation was partitioned mainly between species, compared to the variation between samples within species and within samples: The largest part of the total genetic variation (56.5%) was due to differences between the four species samples, 25.9% was due to differences between samples within species and 17.6% of the total variation was found within samples. All variance components were highly significant (P-value < 0.001).
The pairwise comparison of the two samples of the morphological phenotypes of *S. mentella* ("oceanic" and "deep-sea") revealed a highly significant (P-value < 0.001%) difference in haplotype frequencies between the two samples, with a Φ_{ST} value of 0.44 (= 44%).

Table 1.9. Results of the analysis of molecular variance (AMOVA) with variance partitioned among species, samples within species and within samples (d. f. = degrees of freedom). *** Probability (P) values < 0.001

Source of variation	d. f.	Variance components	Percentages of variation
Among species	3	0.852	56.46***
Among samples within species	11	0.392	25.93***
Within samples	258	0.266	17.61***
Total	272	1.510	100.00

5. Discussion

5.1 Sequence evolution and substitution analysis

The molecular structure and evolution of the ND3 gene in *Sebastes* found in this study conform to patterns observed in other fish species:

Of the total number of mutations, approximately 83% among North Atlantic *Sebastes* haplotypes and 76% among all *Sebastes* haplotypes were hits at 3rd codon positions. Similar values were observed in the ND3 gene sequences of lamprey (Docker et al. 1999) and salmonids (McKay et al. 1996) as well as in cytochrome b gene sequences of Pacific *Sebastes* species (Rocha-Olivares et al. 1999a). This is in agreement with the general observation that replacements at 3rd codon positions tend to occur at faster rates than at 1st or 2nd codon positions, as they are more likely to be neutral (i.e. have no effect on fitness; Kocher and Carleton 1997). This is also reflected by the results of this study, as the majority of the mutations observed were silent (synonymous) mutations.

The high transition/transversion ratio observed in *Sebastes* is consistent with the generally accepted higher transition/transversion ratio (ti/tv) in mtDNA, congruent with the assumption that the mismatch at DNA replication, the primary mechanism of mutation in mtDNA sequences, favours transitional mutations versus transversional substitutions (McKay et al. 1996). The high transition/transversion ratio tended to be much higher for closely related species pairs, such as the North Atlantic *Sebastes* species (ti/tv approximately 10:1), and considerably lower for more distantly related pairs, e.g. when the haplotypes of all species were included (ti/tv approximately 3:1) - a pattern also observed in *Oncorhynchus* species (McKay et al. 1996). In general, the transition/transversion ratio appears to decrease with increasing sequence divergence and time since common ancestry (Meyer 1993; Meyer 1994).

The preponderance of synonymous substitutions at 3rd codon positions - as well as the high transition/transversion ratios for North Atlantic *Sebastes* and also for all *Sebastes* samples analysed - are indicative of relatively young species and low lineage times (Kocher and Carleton 1997; Cope 2004).

Also the absence of site saturation indicated a relatively recent shared ancestry among the *Sebastes* species**,** because transitions at 3rd codon positions tend to saturate within a few million years, as reported for cytochrome b sequence data (Lydeard and Roe 1997).

The saturation plots for the ND3 gene indicated site saturation only in the outgroup species *Helicolenus dactylopterus*, as a slight leveling-off in the number of transitions and transversion was observed at sequence divergences higher than 8% (Figure 1.5). Therefore, the mitochondrial sequence divergence rates observed between the outgroup *Helicolenus dactylopterus* and the other taxa are probably underestimates, as the limited number of nucleotide sites that are free to vary seem to become saturated for changes.

In general, the levels of genetic diversity in *Sebastes*, in particular within the North Atlantic *Sebastes* complex, were low. All species analysed revealed extremely low levels of nucleotide diversities, always smaller than 0.01 (nucleotide diversities of 0.1 and more would represent very deep divergences; Grant and Bowen 1998). The standard error in nucleotide diversity indices was high due to the small degree of sequence divergence.

The total sample of North Atlantic *Sebastes* was characterised by a moderate to high haplotype diversity ($\hat{h}$ = 0.85). According to Grant and Bowen (1998), low nucleotide diversities and low haplotype diversities are suggestive of a recent population bottleneck, caused for example by a founder event, whereas low nucleotide diversities and high haplotype diversities imply a rapid population expansion with an accumulation of new mutations after a period of low effective population size (founder event, population bottleneck).

Therefore, a founder effect, or a population bottleneck, followed by a subsequent rapid population expansion, is the most likely explanation for the observed genetic variation in the North Atlantic *Sebastes* species complex. This is supported by the generally assumed descendance of North Atlantic *Sebastes* from a single common ancestor (Love et al. 2002; see paragraph 5.2 and 5.4).

Even though haplotype diversity within the total North Atlantic *Sebastes* sample was on the high end of range among teleosts (haplotype diversity exceeding 0.80 have been found in various bony fishes; Grant and Bowen 1998), the observed intraspecific haplotype diversities were moderate (*S. marinus*; *S. mentella* and *S. viviparus*) to low (*S. fasciatus*). The highest intraspecific haplotype diversities were observed in the samples of *S. marinus* and "giant" *S. marinus*, most likely caused by the appearance of *S. fasciatus* and *S. mentella* haplotypes, as well as rare haplotypes (G1-G4) and two distinct *S. marinus* specific haplotype lineages (MA1 and MA2/MA3) in these samples (see paragraph 5.3.1, 5.3.2 and 5.3.3 for a detailed discussion). However, the higher levels of haplotype diversity observed in samples of *S. marinus* and "giant" *S. marinus* ($\hat{h}$ = 0.75 and 0.76, respectively) are relatively low compared to values found in Pacific *Sebastes* species ($\hat{h}$ > 0.92; e.g. Kai et al. 2002a, Cope 2004).

Marine teleost fish generally exhibit low levels of haplotype diversity, which may arise e.g. from slower rates of genome evolution, recent bottleneck or founder events (reviewed in Ovenden 1990). Furthermore, haplotypes may be not detected because the size of the sample is too small to accommodate all possible haplotypes (Ovenden 1990). Therefore, the moderate to low intraspecific haplotype diversities in North Atlantic *Sebastes* may arise from the generally assumed recent ancestry of the species. However, a non-detection of haplotypes cannot be ruled out, in particular in case of *S. viviparus* and *S. fasciatus*, as the sample sizes for both species were low compared to *S. marinus* and *S. mentella* - and the samples of *S. viviparus* and in particular the samples of *S. fasciatus* were collected from a limited number of locations.

The varying haplotype diversity indices of the samples of *S. alutus*, *S. capensis* and in particular of the outgroup *Helicolenus dactylopterus* are difficult to interpret due to the small

sample sizes. In all three cases, it is probable that only a small portion of haplotypes was sampled, as only three to ten individuals per species were analysed. This is also reflected in the high standard errors in haplotype diversity indices.

The extremely low levels of sequence divergence observed in the *Sebastes* samples, in particular within the North Atlantic species complex, further support a recent history of the species. The level of sequence divergence in the ND3 gene within the North Atlantic *Sebastes* species complex (between 0.3% and 2.3% sequence divergence between all pairs of North Atlantic *Sebastes* haplotypes) is low in comparison to other fish species: On average between 5.9% and 13.1% sequence divergence in the ND3 gene is reported from *Oncorhynchus* species (between 17 and 35 transitions and two to eight transversions; Thomas and Beckenbach 1989; Domanico and Phillips 1995) and between 16.0% to 17.4% for species of lamprey (Docker et al. 1999). In comparison, the sequence divergences between *S. alutus* and the North Atlantic *Sebastes* species, species that occur in different oceans with presumably no genetic exchange for a period of about 3.5 million years (see paragraph 5.4), were low (between 2.3% and 3.4%). Also the intergeneric differences between *Sebastes* and *Helicolenus* were low (at maximum 11.4%), comparable to differences between *Oncorhynchus* species (e.g. Domanico and Phillips 1995).

The rates of ND3 gene sequence divergences among *S. marinus*, *S. alutus*, and *S. capensis* observed in this study are remarkably similar to the rates of sequence divergences of the cytochrome b gene, determined by Rocha-Olivares et al. (1999a) and Kai et al. (2003). The cytochrome b sequence divergence between *S. marinus* and *S. alutus* was about 2.8%, between *S. capensis* and *S. alutus* 4.5% and between *S. capensis* and *S. marinus* 5.3% (Rocha-Olivares et al. 1999a; Kai et al. 2003). These values are similar to the values found in this study, implying that the substitution rate is similar in both protein-coding genes. This could indicate that within the genus *Sebastes* the rate of evolution in the ND3 gene is not higher than in the cytochrome b gene, in contrast to the higher rates of evolution in the ND3 gene observed in salmonid species (Thomas and Beckenbach 1989; Domanico and Phillips 1995). However, further ND3 and cytochrome b sequence analyses of the same individuals would be needed to confirm these results.

The observation of low levels of sequence variation and divergence among the North Atlantic *Sebastes* species is in congruence with the morphological similarities between the species (e.g. Barsukov et al. 1985) and with the general low levels of inter- and intraspecific genetic variation found in previous genetic studies. The appearance of only few fixed differences in protein markers (Nævdal 1978; Payne and Ni 1982; McGlade et al. 1983; Rehbein 1983; Nedreaas and Nævdal 1989; Trottier et al.1989; Rubec et al. 1991; Nedreaas et al. 1994), the presence of a tandem repeat unit in the mitochondrial control region, identical in length with the same two restriction sites in the three species *S. marinus*, *S. mentella* and *S. fasciatus* (Bentzen et al. 1998), and low levels of differentiation in restriction site polymorphism of mtDNA (Desrosier et al. 1999) support a close phylogenetic relationship between the four

species and again indicate a recent ancestry. The sequence analysis of Sundt and Johansen (1998), based on the mitochondrial 16S rRNA gene, revealed even lower degrees of genetic differentiation (only one transition separated *S. fasciatus* and *S. viviparus* from *S. mentella* and *S. marinus*, the latter two showed identical sequences) than the present study, further supporting a close relationship between the North Atlantic *Sebastes* species.
In conclusion, all results indicate a relatively recent shared ancestry among the members of *Sebastes* (Rocha-Olivares et al. 1999a), in particular of the North Atlantic species complex.

5.2 Phylogeny, species structure and the evolutionary history of North Atlantic Sebastes

The observed levels of sequence divergence were reflected in the results of the phylogenetic analysis. The phylogenetic relationships of the observed main *Sebastes* lineages, the *S. alutus*, the *S. capensis* and the North Atlantic *Sebastes* lineage, were consistent among all three methods of phylogenetic reconstruction (using neighbour-joining, maximum parsimony and quartet puzzling maximum likelihood methods), as well as among all different substitution models applied in the neighbour-joining and quartet puzzling maximum likelihood approaches. The congruent results among the methods, the high bootstrap (NJ and MP) and reliability values (quartet puzzling) at the corresponding nodes, as well as the relatively high consistency indices in the maximum parsimony analysis, indicated the robustness of the observed phylogenetic relationships among the three groups.
Of the three main *Sebastes* lineages observed, the *S. capensis* lineage was most distantly related to the North Atlantic *Sebastes* lineage, whereas the *S. alutus* lineage was closer related to the latter - even though both lineages, North Atlantic *Sebastes* and *S. alutus*, are separated geographically by the North American continent.
This corresponds with our knowledge about the origin and taxonomic status of *S. capensis*: *S. capensis* belongs to one of the 14 species of the monophyletic subgenus *Sebastomus* (Rocha-Olivares et al. 1999a). It has been hypothesised that the Southern Hemisphere *Sebastes* species (*S. capensis, S. oculatus* and probably a third cryptic species), all belonging to the subgenus *Sebastomus*, originated from immigration of a common ancestor from the northern Pacific through equatorial waters via isothermal submergence, during a period of cooling (Eschmeyer and Hureau 1971; Rocha-Olivares, et al. 1999c). The event took place during a glacial maximum, probably within the past two million years (Eschmeyer and Hureau 1971). This hypothesis was confirmed by genetic studies of mtDNA sequences, in which the time span since divergence was determined to be between 141.000 and 80.000 years before present, assuming a range of rates of mtDNA evolution (Rocha-Olivares et al. 1999c). Due to the monophyletic status of the subgenus as well as its evolutionary history and origin, one would expect that the South Atlantic species *S. capensis* clusters separately from the North Atlantic *Sebastes* species, as shown in the present study - even though the five species occur in the same ocean.

The clustering of the *S. alutus* lineage with the North Atlantic *Sebastes* lineage is congruent with the evolutionary history of the North Atlantic *Sebastes* species complex. It is generally believed that the North Atlantic *Sebastes* species are monophyletic and derived from a single North Pacific ancestor. Previous genetic analyses indicated that this ancestor was most likely *S. alutus* or its ancestor: Rocha-Olivares et al. (1999a) compared in their phylogenetic analysis of Pacific and South Atlantic *Sebastes* the cytochrome b gene sequences of 76 taxa, including *S. marinus*. They found that *S. marinus* consistently grouped with *S. alutus*. However, they did not analyse any of the other three North Atlantic *Sebastes* species.

The present results of the phylogenetic analysis and the relatively small genetic differences strongly support the sister-taxon relationship between *S. alutus* and North Atlantic *Sebastes*. This is also supported to some degree by morphological similarities between the species and the fact that *S. alutus* is one of the five recognised trans-Pacific species of *Sebastes*, inhabiting deep waters (Rocha-Olivares et al. 1999b; Love et al. 2002).

In contrast to the clear clustering of the three main *Sebastes* lineages, the phylogenetic trees reconstructed using neighbour-joining, maximum parsimony and quartet puzzling maximum likelihood methods showed that within the North Atlantic *Sebastes* lineage phylogenetic structure was shallow compared to the divergence from *S. alutus* and *S. capensis*. This is indicated by the short branch lengths in the neighbour-joining dendrogram (Figure 1.6), as well as by the large polytomy in the MP 50% majority rule consensus tree (Figure 1.7) and the generally low bootstrap and reliability values.

Even though all three phylogenetic approaches showed a grouping of *S. viviparus* and *S. fasciatus* haplotypes into two distinct groups, with moderately high bootstrap values (between 60 and 77) and reliability values (between 54% and 91%), differences between the methods resided in the placement of *S. marinus*, "giant" *S. marinus* and *S. mentella* haplotypes and their relationship with the *S. viviparus* and *S. fasciatus* haplotypes. No species specific clustering of *S. marinus*, "giant" *S. marinus* and *S. mentella* haplotypes was observed, but certain haplotype groupings were found in two or more phylogenetic reconstructions with sometimes high support, e.g. G1 and G3 (in all trees), MA2 and MA3 (in the NJ and quartet puzzling trees), as well as ME3, ME5, ME7 and MA1 (in all trees).

Nevertheless, none of the three approaches was able to solve the phylogenetic relationships among the haplotypes of *S. mentella*, *S. marinus* and "giant" *S. marinus* as well the relationship between these haplotypes and the *S. viviparus* and *S. fasciatus* clusters, most likely due to the low levels of genetic differentiation in the mitochondrial ND3 gene (see paragraph 4.1 and 5.1).

Because of the small genetic differences between North Atlantic *Sebastes* haplotypes, the relationships between these haplotypes were better resolved and visualised in the statistical parsimony network (Figure 1.9) than in the phylogenetic trees.

The statistical parsimony network represented an almost completely resolved, one-step network, with only one unresolved reticulation: The closed loop connecting the haplotypes

ME2 and VI2 and ME2, ME8, VI3, and VI2, respectively, via hypothetical haplotypes indicated ambiguity among two alternative, equally short pathways connecting the haplotypes. This ambiguity was also observed in the phylogenetic trees: In contrast to the NJ and MP trees, the quartet puzzling tree grouped ME8 with the *S. viviparus* lineage. The reticulation can be caused by two circumstances, (i) either VI2 arose from independent mutations in two different lineages (convergence: from ME2 or from ME2 via ME8 and VI3) or (ii) one of the lines represents a direct evolutionary path that did not actually occur.

The main feature of the statistical parsimony network was the presence of a star-like phylogeny, with a single high-frequency haplotype in the centre and radiating moderate- to low-frequency haplotypes, separated from the central haplotype by one or only few mutational steps. A star-like phylogeny represents a common pattern in widely distributed marine fish populations (Grant and Bowen 1998). Star-like phylogenies are described on the population level for haddock (*Melanogrammus aeglefinus*), Atlantic cod (*Gadus morhua*), cape hake (*Merluccius capensis*), Atlantic herring (*Clupea harengus*), orange roughy (*Haplostethus atlanticus*), Albacore tuna (*Thunnus alalunga*), sardines (*Sardina*; *Sardinops*), anchovis (*Engraulis*) and many other species (see Shields and Gust 1995; Grant and Bowen 1998 and Avise 2000 and references therein).

On species level, star-like phylogenies were observed e.g. in hamlet fish (*Hypoplectrus spp*.: Ramon et al. 2003).

A star-like phylogeny is usually regarded an indication of a rapid population expansion and speciation from a small number of ancestors, due to recent bottlenecks or other demographic effects like founder events (Slatkin and Hudson 1991; Grant and Bowen 1998; Avise 2000). This is in congruence with the general assumed colonisation of the North Atlantic by a founder event and the descendance of the North Atlantic *Sebastes* species from a single common ancestor lineage (Rocha-Olivares 1999a; reviewed in Love et al. 2002).

It is generally assumed that the single high-frequency haplotype in the centre of a star-like phylogeny represents the most ancestral lineage (Watterson and Guess 1977; Donnelly and Tavaré 1986; Crandall and Templeton 1993; Castelloe and Templeton 1994; Excoffier and Smouse 1994; Avise 2000; Posada and Crandall 2001). Therefore, the central position of the most frequent haplotype in *S. mentella*, ME2, with a high number of connections to the other haplotypes, as well as its wide geographical distribution and high frequency, suggests that this haplotype represents the most ancestral state within the North Atlantic *Sebastes* lineage.

The haplotype ME2 could therefore represent the radiation origin that gave rise to all other haplotypes, in particular as the remaining descending haplotypes found in *S. mentella*, but also the haplotypes found in *S. marinus* and *S. fasciatus,* were radiating either directly from ME2 or grouped in a subcluster that was directly connected to ME2 (except in case of *S. viviparus*, where two hypothetical haplotypes and three mutation steps connected these haplotypes with ME2).

Therefore, this study provides the hypothesis that the *S. mentella* lineage, represented by its most common haplotype ME2, is the representative of the oldest lineage of North Atlantic *Sebastes*, due to the wide distribution and the basal position of the haplotype ME2. This is also supported by the fact that ME2 was one out of three haplotypes (ME2, ME7 and MA3) that revealed the smallest genetic distance (2.3% sequence divergence and eight substitutions) to *S. alutus*. In contrast, the sequence divergence observed between *S. alutus* and all other North Atlantic *Sebastes* haplotypes was higher (at least nine mutation steps). Also the wide distribution of *S. mentella* throughout the North Atlantic - in contrast to the other species, which show a more restricted area of distribution - supports the hypothesis that the *S. mentella* lineage represents the most basal lineage.

The haplotypes found in *S. viviparus* represent the genetically most divergent lineage, and there are at least two hypothetical, not-detected haplotypes needed to connect this lineage with the other haplotypes. This indicates that the lineage leading to *S. viviparus* split off first from the common ancestral lineage (represented by ME2). This is further supported by the fact that *S. viviparus* is the only species that can be distinguished more or less clearly from the other three species using morphological, morphometrical and meristic traits (Barsukov et al. 1985; Barsukov et al. 1990). Also its restricted distribution (from the Barents Sea to the west of Iceland) and its depth preferences separate *S. vivi*parus from the other three species, as the species is found closer to the shore and in shallower waters (between 10 m and 120 m depth) than the other three species (Barsukov 1985; Whitehead et al. 1986; Magnússon and Magnússon 1995).

The *S. viviparus* lineage and the *S. fasciatus* lineage were well separated in the network and the genetic differences found between some haplotypes of the two species represented the largest interspecific differences observed among the North Atlantic haplotypes. This is in congruence with the results of a previous study by Sundt and Johansen (1998) based on the mitochondrial 16S rRNA gene. In their study, the largest genetic difference was observed between *S. fasciatus* and *S. viviparus* haplotypes (two substitutions, in contrast to one substitution between the *S. fasciatus* and *S. viviparus* haplotypes and the monomorphic 16S rRNA haplotype found in *S. marinus* as well as in *S. mentella*; Sundt and Johansen 1998).

The *S. fasciatus* lineage, even though forming a distinct subcluster, was directly related to the central *S. mentella* haplotype ME2, which indicates a direct descendance of this lineage from *S. mentella*. This applies also to the less frequent haplotypes within the *S. mentella* lineage. All haplotypes were either directly connected to haplotype ME2, or, in case of the two rare haplotypes ME5 and ME7, through the relatively frequent haplotype ME3, indicating that all the less frequent haplotypes found in *S. mentella* are closely related and descended from haplotype ME2.

In contrast to the *S. fasciatus* and *S. viviparus* lineages, which were well separated in the network, the haplotypes found in *S. marinus* did not form a single distinct subcluster. The haplotypes appearing exclusively in the samples of *S. marinus* did not cluster together: Two

of the haplotypes found in *S. marinus* (MA2 and MA3) were directly connected by only one mutation step, but the third common haplotype found in *S. marinus* - MA1 - was more closely related to haplotypes found in *S. mentella* than to the haplotypes MA2 and MA3 and formed a subcluster with ME3, ME5 and ME7 (haplotypes that were diagnostic for *S. mentella*; see paragraph 5.3). This was also reflected in the high intraspecific genetic differences between the haplotypes MA1 and MA2 as well as MA3, respectively, which was of the same order of magnitude as the interspecifc differences observed between *S. mentella* and *S. marinus*. In fact, the genetic difference between the haplotypes MA1 and MA3 represented the highest intraspecific difference observed in the data set, indicating that haplotype MA1 represents a genetically distinct lineage.

The rare haplotypes found in the sample of "giant" *S. marinus* (G1-G4) were actually closer related to ME2 than to the *S. marinus* specific haplotypes MA1, MA2 and MA3.

The observation that the individuals sampled as *S. marinus* and "giant" *S. marinus* represented a genetically heterogeneous group consisting of different mitochondrial lineages is further supported by the fact that several individuals sampled as *S. marinus* and "giant" *S. marinus* revealed *S. fasciatus* and *S. mentella* specific haplotypes, respectively. The heterogeneity observed within the *S. marinus* samples will be discussed in detail in the following paragraphs (5.3.1, 5.3.2 and 5.3.3).

In summary, however, the genetic differences - especially between the haplotypes found in *S. mentella*, *S. marinus* and *S. fasciatus* - were extremely small and indicated a direct descendance of the *S. fasciatus* lineage and the *S. marinus* lineages from the basal *S. mentella* lineage (represented by ME2), which might have occurred recently (see paragraph 5.4), whereas the *S. viviparus* lineage probably originates from an earlier split from the basal lineage.

These observations are somewhat different to the findings by Roques et al. (1999a), who concluded on the basis of shared microsatellite alleles that *S. marinus* represents the lineage from which the other species arose and that *S. fasciatus* and *S. viviparus* are more closely related to each other than to the other Atlantic species, similar to a hypothesis by Barsukov and Zakharov (1973). However, even though microsatellites are powerful genetic markers on the population level and the North Atlantic species are closely related, caution should be used when employing microsatellites in interspecific phylogenetic studies. A part of the difficulty certainly stems from restrictions to divergence imposed by range constraints, irregularities and asymmetries in the mutation process, and the degradation of microsatellites over time (Goldstein and Pollock 1997). Their high mutation rates lead to a large amount of homoplasy (convergence in the size of different alleles) over a relatively short period of time and can strongly affect the inference of phylogenetic relationships. Therefore, microsatellites make very poor markers for phylogenetic inferences, except for groups separated by no more than a few thousand generations (Jarne and Lagoda 1996).

5.3 Patterns of haplotype distribution

Even though the observed levels of genetic differentiation between the haplotypes were low - in particular between the North Atlantic *Sebastes* haplotypes - the degree of genetic variability was large enough to recognise several diagnostic mutations characteristic for the *Sebastes* species analysed. All haplotypes found in the *S. capensis*, *S. alutus* and *S. viviparus* samples were not detected in any other sample and therefore species-diagnostic, allowing an unambiguous discrimination between these species using ND3 gene sequence data.
Also the haplotypes identified in the samples of *S. mentella, S. marinus* and *S. fasciatus* were species specific in most cases, and the mitochondrial DNA analysis supported the *a priori* classification based on morphological traits and/or hemoglobin analysis (see paragraph 3.1) in most samples.
The significant differences between the *Sebastes* species were also manifested in the results of the AMOVA (Table 1.9), which indicate that the great majority of genetic variation occurred between the species.
However, in five species samples haplotypes from different species appeared. In three of these samples (MEFC01, MADB01 and MASEIc01), only one single individual per sample was involved. Most striking were the haplotype distributions in the *S. marinus* sample from the Flemish Cap and in the sample of "giant" *S. marinus*.

5.3.1 S. marinus on the Flemish Cap – a case of interspecific introgression of mitochondrial DNA?

All twenty individuals of the *S. marinus* sample from the Flemish Cap revealed haplotypes (FA1 and FA3) that were specific for *S. fasciatus*[1] and did not appear in any other sample of *S. marinus*. The individuals were taken from different stations and ten out of twenty individuals were re-analysed in the laboratory by a Spanish specialist in redfish morphology and morphometry[2] and classified as *S. marinus* morphologically and morphometrically. Therefore, a misclassification of species is unlikely, even though species discrimination of North Atlantic *Sebastes* species is difficult - in particular in areas of sympatry - as all North Atlantic *Sebastes* species show a wide overlap in morphological and morphometrical characters (Barsukov et al. 1985).
One possible explanation for the appearance of the haplotypes FA1 and FA3 in the *S. marinus* sample from the Flemish Cap is hybridisation. These individuals could be the product of an introgressive hybridisation event between *S. fasciatus* and *S. marinus*. Additional genetic analyses using nuclear markers were performed to investigate the taxonomic status of these

[1] With one exception: One individual of *S. mentella* from the Flemish Cap also displayed FA1 (for further discussion see below).
[2] Dolores Garabana, Institute of Marine Research, Vigo, Spain

individuals on the level of nuclear DNA and to test the hypothesis of hybridisation. The results of these analyses are presented in detail in Chapter 2 and 3 of this thesis. Briefly, the analysis of the individuals using nuclear microsatellite (see Chapter 2) and AFLP markers (see Chapter 3) did not indicate misclassification, as the nuclear multilocus genotypes allowed the unambiguous assignment of these individuals to *S. marinus*. Also no indication for recent hybridisation was found, as the microsatellite analysis did not reveal individuals with intermediate allelic compositions and admixed genotypes (see Chapter 2 for further discussion). Therefore, it is unlikely that these individuals are a product of a recent hybridisation event (e.g. F1 hybrids).

However, there are many examples of mitochondrial introgression unaccompanied by apparent nuclear introgression (e.g. Bernatchez et al. 1995; Ballard 2000; Sota 2002; see also Ballard and Whitlock 2004). In some cases, the mitochondrial DNA from one taxon completely replaces that of another, without any evidence of nuclear introgression or morphological signal. One example for this phenomenon is the complete interspecific replacement of the mitochondrial genome in an allopatric population of brook trout (*Salvelinus fontinalis*) with the mitochondrial genome of Arctic charr (*Salvelinus alpinus*), even though these brook trout are morphologically indistinguishable from normal brook trout and have diagnostic alleles at nuclear loci (Bernatchez et al. 1995; reviewed in Ballard and Whitlock 2004).

Alternative hypotheses to introgression might explain the distinct mtDNA composition of the *S. marinus* sample from the Flemish Cap. Ancestral polymorphic DNA may have been retained by the two species *S. marinus* and *S. fasciatus*, since there is strong indication that they diverged from a common ancestor (see paragraph 5.2). If there has not been enough time for new variant genes to become fixed between instances of speciation, the lineages have not yet sorted themselves perfectly into species, but some lineages occur in more than one taxon. This problem is known as "incomplete lineage sorting" (Kocher 2003; Ballard and Whitlock 2004). Identity with the mitochondrial genome of *S. fasciatus* could also have resulted from convergent mutation, a type of homoplasy.

It is often very difficult to distinguish which factor is responsible for the sharing of genetic variants in different species or taxa (Verspoor and Hammar 1991; Bernatchez et al. 1995). However, in the case reported here, the haplotypes FA1 and FA3 were not found in any other sample of *S. marinus* over its range of distribution, making the hypothesis of retained ancestral polymorphic DNA and incomplete lineage sorting in both species unlikely. In addition, incomplete lineage sorting should involve ancestral haplotypes and not derived haplotypes. The results of the phylogenetic analysis (see paragraph 5.2 and Figure 1.9) did not indicate that haplotypes FA1 and FA3 are of ancestral origin, from which the *S. marinus* haplotypes derived. Convergent mutation is also a rather unlikely explanation: The levels of mtDNA sequence divergence observed between the haplotypes of *S. marinus* and *S. fasciatus* (FA1 and FA3) were relatively high (three and four mutations, respectively). It is unlikely that

such convergent mutations, requiring at least three to four homoplastic events, result in a haplotype of a sympatric species group (Bernatchez et al. 1995). Also the lack of site saturation even among more distantly related taxa (e.g. between *S. capensis* and the North Atlantic *Sebastes* species) does not indicate homoplasy.

Therefore, the incorporation of the *S. fasciatus* haplotypes into *S. marinus* on the Flemish Cap through introgression is the most likely explanation, in particular as the Flemish Cap is one of the few areas on the Northeast American shelf, where *S. marinus* is noted more frequently and occurs in sympatry with *S. fasciatus* (and also with *S. mentella*; Barsukov et al. 1985). Introgressive hybridisation in the Northwest Atlantic is not an unusual observation in North Atlantic redfish. Roques et al. (2001) used microsatellite analysis to investigate hybridisation between *S. mentella* and *S. fasciatus* in the Northwest Atlantic and found evidence of broad-scale nuclear introgressive hybridisation between the two species in the Gulf of St. Lawrence and adjacent areas.

The observation that the individuals of *S. marinus* from the Flemish Cap resembled morphologically and also genetically (microsatellite and AFLP analysis, see Chapter 2 and 3) *S. marinus*, but showed mitochondrial haplotypes identical to haplotypes found in *S. fasciatus*, is somewhat similar to the observation of mitochondrial DNA introgression in *Salvelinus fontinalis* mentioned above (Bernatchez et al. 1995). According to the scenario hypothesised by Bernatchez et al. (1995), this could indicate that nuclear introgression has not occurred or has long been diluted. Several studies have found evidence that mtDNA introgresses across species boundaries more rapidly than do nuclear loci (see Avise 1994 and references therein).

Therefore, it is possible that the present situation observed in *S. marinus* on the Flemish Cap arose by unidirectional hybridisation with transfer of the maternally inherited mitochondrial DNA and subsequent repeated backcrossing of female hybrids with male *S. marinus*, until the *S. fasciatus* nuclear genome eventually disappeared.

The study by Roques et al. (2001) revealed that most introgressed individuals of *Sebastes* in the Gulf of St. Lawrence were genetically more similar to one or the other parental species (*S. mentella* and *S. fasciatus*), but also few individuals with intermediate allelic compositions between *S. fasciatus* and *S. mentella* were found. The fact that nuclear introgression has been observed in North Atlantic *Sebastes* could support the hypothesis of an ancient introgression from *S. fasciatus* to *S. marinus* on the Flemish Cap, as no intermediate or admixed genotypes (see Chapter 2) were observed in this area. However, a transfer of these observations to *S. marinus* should be treated with caution, as different taxa were studied. Nevertheless, in the sample of *S. mentella* from the Flemish Cap one single individual also revealed a *S. fasciatus* specific haplotype, but was identified as *S. mentella* due to the results of morphological, morphometrical[1] (Garabana pers. comm. 2003) and MDH analyses[2] (MDH = malate

[1] Morphometrical analyses were performed by Dolores Garabana, Institute of Marine Research, Vigo, Spain.

[2] MDH analyses were performed by Torild Johansen, University of Bergen (UIB), Norway.

dehydrogenase; Johansen pers. comm. 2003). As nuclear introgression between *S. fasciatus* and *S. mentella* is known to occur (Roques et al. 2001), this indicates that also this individual could have descended from an ancient introgressive hybridisation event.

Even though the introgressive hybridisation events on the Flemish Cap are probably ancient, they may have had significant and long-term effects on the composition of the mitochondrial DNA of *S. marinus* on the Flemish Cap. The observed complete interspecific replacements of mitochondrial genome in all twenty individuals analysed could indicate that the whole population of *S. marinus* on the Flemish Cap is introgressed, as the individuals analysed in this study have been taken from three different stations across the Flemish Cap (see Appendix I) - even though patchiness cannot be entirely ruled out. Complete interspecific replacements of mitochondrial genome in whole populations have been documented for various animal groups, including mammals, amphibians, insects and fish (Avise 2000).

The two main factors contributing to asymmetrical introgression are selection and genetic drift (e.g. by stochastic, historical demographic events; Ballard and Whitlock 2004).

Even though the ND3 gene sequenced in this analysis revealed only synonymous mutations, the mitochondrial genome is inherited as one single unit with little or no recombination. Since the mitochondrial genome is coding for several mitochondrial enzymes, differences in the mitochondrial genome may be manifested physiologically (Bernatchez et al. 1995). It has been proposed by several authors that thermal adaptation in poikilotherms could represent a type of selection that may cause repeatable patterns in mitochondrial introgression (Bernatchez et al. 1995; Ballard and Whitlock 2004 and references therein). Bernatchez et al. (1995) hypothesise that the introgression of mitochondrial genes from the more cold adapted *Salvelinus* species (*S. alpinus*) into the genome of the more temperate *Salvelinus* species (*S. fontinalis*) may provide a better physiological adaptation to cold water temperatures in high latitude lakes.

In fact, *S. fasciatus* has a more southern distribution (which ranges from the Gulf of Maine to the Scotian Shelf and the Grand Banks) and is more adapted to warmer environments than *S. marinus* and *S. mentella* (ICES 1998a; Roques et al. 2001). In addition, the temperatures of the water masses over the Flemish Cap are higher than those over the adjacent Grand Banks within similar depth ranges (e.g. at depths between 100 m and 150 m, bottom temperatures over the Grand Banks are generally lower than 0°C, compared to 3 to 4°C over the Flemish Cap; Colbourne and Sencall 1996). Therefore, it could be assumed that the introgression of the *S. fasciatus* haplotypes that partly evolved in a warmer environment could give a selective advantage to *S. marinus* living at the southern and western limit of its distribution. However, this hypothesis is difficult to test due to the large amount of abiotic and biotic factors that may affect the fitness of *S. marinus*. The fact that *S. marinus* is not very abundant in the Northwest Atlantic except on the Flemish Cap (Gascon 2003) supports the idea that the Flemish Cap is a point of contact of *S. marinus* and *S. fasciatus* at the extreme ecological limits of the

distribution range of *S. marinus* (the only other areas in the Northwest Atlantic were *S. marinus* occurs more frequently are the Northern and Great Newfoundland Banks and the Gulf of St. Lawrence; Barsukov et al. 1985). A pattern of asymmetrical introgressive hybridisation on nuclear level was also observed in *S. mentella* and *S. fasciatus* in the Gulf of St Lawrence (Roques et al. 2001). Roques et al. (2001) hypothesise that selection may act in the main zone of sympatry of *S. mentella* and *S. fasciatus* in favouring the differential survival of certain hybrids relative to either other hybrids or pure parental genotypes. The same hypothesis could be applied to the situation observed on the Flemish Cap.

However, an asymmetrical introgression can also simply be caused by chance (by genetic drift), due to the lower effective population size of mitochondrial DNA compared to nuclear DNA (Ballard and Whitlock 2004).

For example, the occurrence of relatively fewer individuals of a given species in an area of large abundance of a second species should increase the probability of introgression towards the less abundant species (Arnold et al. 1993). At present, both *S. marinus* and *S. fasciatus* show similar abundances on the Flemish Cap (e.g. Saborido-Rey 1993), but as the hybridisation event is probably ancient, this does not contradict the expectation.

Whether the interspecific replacements of the mitochondrial genome in *S. marinus* could have resulted from either stochastic demographic events (genetic drift) or from selection on certain mitochondrial genes (e.g. a selective advantage coupled with thermosensitivity or other environmental adaptations), or a combination of both, remains hypothetical, as our knowledge on the life history and ecology of the species is limited.

However, the asymmetrical introgression found in the present study as well as the findings by Roques et al. (2001) indicate that in case of hybridisation between North Atlantic *Sebastes* species, the involved taxa are not affected in the same way.

5.3.2 "Giant" S. marinus - another example of ancient introgression events within North Atlantic Sebastes?

The present sequence analysis revealed that the specimens of "giant" *S. marinus* (with a total length between 71 cm and 85 cm – see Appendix II), which were morphologically identified as *S. marinus*[1], exhibited a heterogeneous haplotype composition. The individuals displayed either (i) haplotypes found exclusively in the *S. marinus* samples (haplotype MA2 and MA3), (ii) the most frequent haplotype found in the *S. mentella* sample (ME2), or (iii) rare and unique haplotypes, appearing exclusively in the "giant" *S. marinus* sample. These unique haplotypes were closely related to ME2 (see paragraph 5.2).

[1] by scientists from the Marine Research Institute (MRI), Reykjavík, Iceland

As only very small amounts of DNA of sufficient quality from the individuals of "giant" *S. marinus* were available for further analysis, only 12 out of a total of 37 individuals could be analysed using nuclear microsatellite (see Chapter 2) and AFLP markers (see Chapter 3). However, the results of these analyses revealed that these twelve individuals could be unambiguously assigned to *S. marinus*. Of these twelve individuals, eight individuals displayed the *S. marinus* haplotypes MA2 and MA3 and four displayed the haplotype ME2, common in *S. mentella*. No indication for recent hybridisation was found, as the microsatellite analysis did not reveal admixed genotypes (see Chapter 2).

The results of the ND3 gene analysis indicate that at least 14 out of 37 "giant" *S. marinus* can be classified as *S. marinus*. This is in congruence with Barsukov (1973) and Kosswig (1974), who proposed that there is no difference between "giant" *S. marinus* and "ordinary" *S. marinus*. The morphological definition of the so-called "giants" has been discussed intensely (Kotthaus 1950 and 1961; Altukhov and Nefyodov 1968; Barsukov 1973; Kosswig 1974; Johansen et al. 2000a). Some authors defined "giants" by their average total length, being above 60 cm and therefore separated from the normal range of *S. marinus* and *S. mentella* length curves (Kotthaus 1950; Altukhov and Nefyodov 1968). Kotthaus (1961) described "giants" as being morphologically similar to *S. marinus*, but showing some differences in the average number of gill rakers and in the internal structure of the otoliths when compared with *S. marinus* and *S. mentella*. In contrast, Barsukov (1973) and Kosswig (1974) proposed that processes of intraspecific variation with age could explain these morphological differences, in particular as the morphological and meristic studies by Kotthaus (1961) did not reveal any essential differences between the "ordinary" and "giant" types of *S. marinus*.

However, not all "giants" analysed in the present study revealed a *S. marinus* specific haplotype.

The fact that several individuals resembled morphologically and also genetically (microsatellite and AFLP analysis, see above and Chapter 2 and 3) *S. marinus*, but showed a mitochondrial haplotype frequently found in *S. mentella*, could be explained by the three different hypotheses already discussed in paragraph 5.3.1: Introgression due to an ancient hybridisation event, incomplete lineage sorting and homoplasy.

In case of "giant" *S. marinus,* it is more difficult to determine which of these hypotheses is the most likely explanation for the haplotype composition, as the number of substitutions between the haplotypes is very small (in most cases only one substitution) and the basal haplotype ME2 is involved (see paragraph 5.2 and Figure 1.9). As ME2 was observed in only two other individuals of all *S. marinus* samples analysed (in MADB01and MASEIc01: Dohrn Bank and Iceland), incomplete lineage sorting is relatively unlikely. The appearance of ME2 in "giant" *S. marinus* and in the two single individuals of *S. marinus* from the Dohrn Bank and Iceland can also be explained by homoplasy: ME2 could have arisen by back–mutation, as only one mutation step separates ME2 from MA2. However, the lack of site saturation does not indicate homoplasy, even though it cannot be entirely ruled out.

Although incomplete lineage sorting and homoplasy cannot be ruled out, an ancient introgressive hybridisation event as discussed in chapter 5.3.1 could be a likely explanation for the observation of ME2 in "giant" *S. marinus*. In case of "giant" *S. marinus*, however, there is no indication of a complete replacement of the mitochondrial genome, as 14 out of 37 individuals carried haplotypes that were *S. marinus* specific and therefore morphologically as well as genetically represented *S. marinus*.

The hybrid origin of "giant" *S. marinus* was already proposed by Altukhov and Nefyodov (1968), who analysed the protein composition of blood sera. The observation of the rare ND3 gene haplotypes G1, G2, G3 and G4 further supports the hypothesis of an ancient introgression event, as the four rare haplotypes were not observed in any other sample (neither in *S. marinus* nor in *S. mentella*), but were more closely related to ME2 than to MA2 and MA3 (see paragraph 5.2). However, due to the lack of DNA material, these individuals could not be analysed using microsatellite markers. Therefore, it was not possible to determine, if these individuals resembled *S. marinus* on the nuclear level - evidence for an introgression event - or if they represented a genetically distinct lineage, as for example the individuals of *S. marinus* carrying the distinct haplotype MA1 (see paragraph 5.3.3).

The fact that the rare haplotypes exclusively occurred in the "giant" *S. marinus* could be due to non-detection of haplotypes in the other samples of *S. marinus* analysed, or a loss of rare haplotypes through random genetic drift. The fishery for large redfish in the area along the Reykjanes Ridge below 500 m - where the samples were collected - started in 1996 and decreased since then. Since 1997 there were no reports of "giant" catches (ICES 1998b). The ongoing depletion of "giant" *S. marinus* and their rare haplotypes could explain why the haplotypes G1, G2, G3 and G4 were not found in any of the other samples of *S. mentella.*

The findings of the present study contradict the results of the hemoglobin analysis by Johansen et al. (2000a), who found that most of the "giant" *S. marinus* analysed in their study revealed a hemoglobin pattern that differed from the pattern found in other *S. marinus* samples. They concluded that "giant" *S. marinus* should be considered at least to be a separate population of *S. marinus* but could also represent a separate species.

However, even though all individuals analysed in the present study revealed the "giant" specific hemoglobin pattern[1] (Johansen pers. comm. 2003) defined by Johansen et al. (2000a), the results of the mitochondrial ND3 gene analysis (as well as the microsatellite and AFLP analysis; Chapter 2 and Chapter 3) did not indicate that "giant" *S. marinus* could represent a different species.

The distinct hemoglobin pattern in "giant" *S. marinus* found by Johansen et al. (2000a) could be explained by selection. Differences at only one or two loci may be due to differential selection rather than restricted gene flow (Smith 1990). Hemoglobin is an important gene product: Many fish species have multiple Hb components which show considerable

[1] Hemoglobin analyses were performed by Torild Johansen, University of Bergen (UIB), Norway.

differences in amino acid sequences and functional properties, but no unifying theory has been proposed on the biological significance of this multiplicity (see Tamburrini et al. 2001 and references therein). The extent to which hemoglobin variation in redfish reflects selective processes or drift processes remains unknown. The fact that hemoglobin is the only marker showing a completely different genetic pattern could indicate that these hemoglobin patterns are caused by selection and adaptation to certain environmental conditions. This is supported by the fact that "giant" *S. marinus* are found only at greater depths (below 500 m) while *S. marinus* normally inhabits depths of less than 300-370 m (Barsukov et al. 1985). Selection on certain hemoglobin alleles, with temperature being the selective parameter, has already been observed in cod (Petersen and Steffensen 2003) and notothenioids (Bargelloni et al. 1998). Furthermore, Johansen et al. (2000a) did not observe similarly high levels of divergence at any of the other eighteen protein (allozyme) loci analysed (in fact most of them were monomorphic for all four species; Johansen et al. 2000a; see also Nedreaas and Nævdal 1989; 1991a; 1991b). Also the different genetic analyses performed in the present study - based on presumably neutral DNA markers - did not show this high level of genetic differentiation. This supports the hypothesis of selection acting on the distribution pattern of certain hemoglobin alleles.
In conclusion, the results indicate that a part of the individuals of "giant" *S. marinus* analysed in this study are the product of an ancient introgressive hybridisation event between *S. mentella* and *S. marinus,* but that not all individuals are introgressed and therefore represent *S. marinus* (genetically as well as morphologically).

5.3.3 S. marinus on the shelves of Greenland and Iceland comprises genetically different mitochondrial lineages

Apart from a possible introgression of mitochondrial DNA from *S. fasciatus* and *S. mentella* into *S. marinus*, another important observation was the existence of two distinct mitochondrial lineages within *S. marinus*. As mentioned in paragraph 5.2, the individuals of *S. marinus* collected on the shelves of Greenland and Iceland revealed two different haplotype lineages: MA1, MA2 and MA3, the first haplotype being genetically distinct from the latter two. This divergent MA1 haplotype was discovered in 21 individuals of *S. marinus* collected off Greenland and Iceland, but not in the *S. marinus* samples from Norway and the Flemish Cap.
The observation that the samples of *S. marinus* from Greenland and Iceland consist of two genetically distinct groups of individuals is strongly supported by the results of the microsatellite analysis (see Chapter 2) and AFLP analysis (see Chapter 3) and points to the existence of sibling species in these areas. This is supported by the results of several studies based on hemoglobin analysis, which revealed that the hemoglobin pattern is more complicated in *S. marinus* from the shelves of Icelandic and Greenland, in contrast to *S. marinus* from Norway (Nævdal 1978, Nedreaas and Nævdal 1991a; Nedreaas et al. 1994).

Sibling or cryptic species are well known to occur in many groups of organisms (see Claridge et al. 1997 and references therein). In the genus *Sebastes*, sibling species have often been the centre of taxonomic and nomenclature revisions (Rocha-Olivares 1999c). Sibling species cannot be distinguished morphologically but show genetic and reproductive differentiation (Avise 1994, Knowlton 2000). It was impossible to determine, whether the 21 individuals, revealing a distinct haplotype pattern compared to the pattern found in Norway, showed morphological differences, as even the distinction between *S. mentella* and *S. marinus* is extremely problematic using external morphological traits (McGlade et al. 1983; Barsukov et al. 1985), in particular in Greenland waters (Garabana 2005). *S. mentella* and *S. marinus* are morphologically remarkably similar and show widely overlapping meristic and morphological characters in many regions. Several authors describe individual fish that possess characters of both *S. mentella* and *S. marinus* in a varying degree (e.g. Kotthaus 1961; Nævdal 1978). Recent studies applying the analysis of the extrinsic gas bladder musculature (EGM) rib passage patterns, a method successfully used in the West Atlantic (Ni 1981; Payne and Ni 1982), did not improve the species discrimination in Greenland waters (Garabana pers. comm. 2003; Garabana 2005). Also morphometrical analyses revealed low discrimination power (Garabana 2005). Therefore, in Greenland waters, only the results of the hemoglobin analysis were used to validate the morphology-based pre-classification of individuals into species on board, as *S. mentella* shows a diagnostic hemoglobin allele (Nævdal 1978, Nedreaas and Nævdal 1989, 1991a; 1991b; see also paragraph 3.1).

However, there is indication that the individuals carrying the haplotype MA1 revealed a strong overlap of morphological characters, as at least all ten individuals from East Greenland (MAEGr01) revealing the distinct haplotype MA1 were pre-classified as *S. mentella*, as some of their morphological characters, such as colour, body proportions, a more or less well developed symphyseal knob (absent in *S. marinus*) and the direction of the lower preopercular spine were not typical for *S. marinus*, but intermediate or more similar to *S. mentella*. The hemoglobin analysis, however, revealed that they did not show the diagnostic *S. mentella* hemoglobin pattern and these individuals were therefore re-classified as *S. marinus* (Johansen pers. comm. 2001). However, hemoglobin analysis can only be used to identify *S. mentella* (Nedreaas and Nævdal 1989, 1991a; 1991b), therefore it can only be concluded from the hemoglobin analysis that these individuals were not *S. mentella*. This emphasises the problems in species identification and taxonomy of North Atlantic *Sebastes* using characters that are not based on DNA polymorphisms.

Nevertheless, the results of the sequence analysis (and also the results of the microsatellite and AFLP analyses - see Chapter 2 and 3) indicate that a group of individuals - collected on the shelves of Greenland and Iceland and pre-classified as *S. marinus* due to morphological traits and/or hemoglobin analysis - is genetically distinct from the other North Atlantic *Sebastes* lineages.

5.3.4 Haplotype distribution in the "deep-sea" and "oceanic" type of S. mentella – an indication of population structure?

The results of the AMOVA indicate that the great majority of genetic variation occurs between the species. However, there is also a degree of genetic variation at a finer scale between the subsamples, as the proportion of variation attributed to differences between samples within species was also found to be significant ($P < 0.001$). However, as the sample sizes were very low in most cases (10 individuals per location), a clear discrimination of populations within species was not possible within the current ND3 gene analysis.
Nevertheless, the haplotype frequencies differed slightly between the samples of "deep-sea" *S. mentella* from the Irminger Sea and the other samples of *S. mentella*, as the haplotype ME3 was quite abundant in "deep-sea" *S. mentella*, but occurred only in two other individuals of the whole *S. mentella* sample set (in the sample of "oceanic" *S. mentella* and *S. mentella* from the Flemish Cap). This is reflected by the significant pairwise Φ_{ST} value observed between the samples of "deep-sea" and "oceanic" type of *S. mentella*. The highly significant Φ_{ST} value indicates that the samples of "deep-sea" *S. mentella* from the Irminger Sea differ genetically from the samples of "oceanic" *S. mentella*. However, a large number of individuals of the sample of "deep-sea" *S. mentella* carried the most frequent *S. mentella* haplotype ME2, and the differences observed between "deep-sea" and "oceanic" *S. mentella* are based on differences in haplotype frequencies. As the statistical analyses are based on only 41 individuals sampled in different years (Table 1.1), these results must be interpreted with caution and need confirmation with samples from additional individuals as well as locations.
The interpretation of these results is even more equivocal, as the validity of the sorting of *S. mentella* in the Irminger Sea into a so-called "oceanic" phenotype, occurring above 500 m, and "deep-sea" phenotype, occurring below 500 m, has been controversial: Iceland has discriminated between the two types in the fisheries since 1995, but there has been a strong controversy about whether the different *S. mentella* phenotypes represent "valid" phenotypes, as the differences are subtle and only a few experts are able to distinguish between them (ICES 1998a; Johansen et al. 2000b; Saborido-Rey et al. 2005). The reported differentiation has been based on several criteria, e.g. differences in colour, length-weight relationships, length-at-maturity, rates of infestation with the parasitic copepod *Sphyrion lumpi* and differences in the size of newly extruded larvae (Magnússon et al. 1994; Magnússon and Magnússon 1995; ICES 1998a).
It has also been discussed controversially, whether the different groups of phenotypes represent different populations. A recent review of the population structure and ecology of *S. mentella* in the Irminger Sea and adjacent waters proposes that there is one population of *S. mentella* in the region Greenland, Iceland and the Irminger Sea (Saborido-Rey et al. 2005) and a recently published work by Roques et al. (2002), based on microsatellite analysis, revealed genetic homogeneity and corresponding low genetic differences within a large

"panoceanic" population of *S. mentella* over a distance of 6,000 km, from Labrador to the Faroe Islands. Roques et al. (2002), however, did not discriminate between the phenotypes.
In contrast, a genetic study by Johansen et al. (2000b) proposes that the "deep-sea" and "oceanic" types of *S. mentella* in the Irminger Sea do not share a common gene pool.
A very recent study has revealed no morphometrical differences between the two morphological types (Garabana 2005). There are strong indications that the sorting of individuals into phenotypes was based on what seems to be mainly personal experience and - at least to some extent - size (Nielsen 2004). The majority of "oceanic" S. *mentella* is smaller, between 30 and 40 cm (mean length of 35-36 cm), whereas "deep-sea" *S. mentella* is known to have a bigger size range, between 36 and 46 cm with a mean length of 42 cm (Saborido-Rey et al. 2005). Similarly, a comparison of the total lengths of all *S. mentella* specimens used in the ND3 gene analysis revealed that the total length of "deep-sea" *S. mentella* was on average higher than that of "oceanic" S. *mentella* (see Appendix II).
In addition, redfish are known to move to deeper waters as they grow (Atkinson 1986; Gauthier and Rose 2002). Even though aging of Irminger Sea *S. mentella* has been conducted only sparsely, studies have shown that the spatial and vertical distribution of the age composition fully reflects the dynamics of the length composition in fish (Saborido-Rey et al. 2005). Therefore, it can be assumed that "deep-sea" *S. mentella* (occurring below 500 m) are not only bigger, but also older than "oceanic" *S. mentella* occurring in shallower waters.
In long-lived species such as *S. mentella*, with a relatively high reproductive potential (due to the viviparous mode of reproduction, offspring are more likely to survive), a population can be sustained by few, abundant and temporally closed year-classes, which reproduce over many years (Saborido-Rey et al. 2005). Even if there is recruitment failure over a long period, the population can be sustained if such year-classes produce a series of successful new year-classes. According to Saborido-Rey et al. (2005) the *S. mentella* population could be formed by for example two sets of year-classes largely separated in time. This is supported by the fact that after a severe depletion of "deep-sea" *S. mentella* off East Greenland during the 1980's and early 1990's, a sharp increase in abundance was observed in 1995-1997, mainly due to an individual strong year-class of young fish, which later migrated into the central Irminger Sea recruiting to the adult stock there (Stransky 2000).
Under conditions of high variance in reproductive success, an entire year class may be the result of relatively few matings. Evidence for such "sweepstakes" recruitment dynamics (Hedgecock 1994) in marine fishes comes from the observation of genetic differences among individual schools of anchovies (Hedgecock et al. 1994) but also for *S. mentella* (Altukhov 1990). In these cases the effective population size for maternally inherited genes, such as the ND3 gene, may be one or two orders of magnitude smaller than the consensus size, leading to higher rates of lineage extinctions than in populations of the same size with many successful spawners (Grant and Bowen 1998). This could explain why ME3 was mainly found in the older "deep-sea" *S. mentella.*

However, it has been argued that such "cohort effects" are probably of minor importance in *S. mentella*, which has probably more than 20-30 reproductive year classes, composing stock sizes of up to one million tons (ICES 2004b).
Therefore, it cannot be ruled out that the observed differences in haplotype frequencies between "deep-sea" and "oceanic" *S. mentella* reflect genetic population structure.
It is beyond the scope of the present mitochondrial DNA sequence analysis to determine, whether the differences observed in haplotype frequencies between "deep-sea" and "oceanic" *S. mentella* are due to "true" population structure caused by reproductive isolation, cohort effects or simply patchiness due to small sample sizes.
The mitochondrial ND3 gene is not particularly well suited for analyses of *S. mentella* population structure, as it revealed relatively low levels of genetic variability. Furthermore, it represents only one locus and sequencing is too time-consuming and expensive to analyse larger sample sizes required for analysis of population structure. A study of the population structure of *S. mentella* requires more variable molecular markers and larger sample sizes, as presented in Chapter 2 and 3 of this thesis.

5.4. Divergence times in North Atlantic Sebastes evolution – an approximate time frame

The average rate of change in mtDNA can be used as a kind of "molecular clock", assuming constant rates of evolution (Avise 1994). Clocks may be calibrated based on comparisons with taxa having known divergences or using well-known geological events, or with the fossil record (Stepien and Kocher 1997). Examination of the level of DNA divergence could make it possible to estimate the rate of sequence divergence among North Atlantic *Sebastes* species, assuming a constant molecular clock. Many objections have been raised concerning the existence of constant rates of evolution (reviewed by Avise 1994). However, if we - hypothetically - assume a constant rate of evolution of the ND3 gene, the invasion of the North Atlantic by the genus *Sebastes* can be used to indirectly calibrate the molecular clock in the different North Atlantic *Sebastes* lineages, as described in Rocha-Olivares (1999b).
The invasion of the North Atlantic most likely took place as a transarctic colonisation during the "Great Transarctic Biotic Interchange" marked by a relatively abbreviated warming period and by the opening of the Bering Strait between Alaska and Siberia about 3.5 million years ago (Briggs 1970; 1995; Eschmeyer and Hureau 1971). This event allowed the marine biotas of the North Pacific and Arctic-North Atlantic to intermix (Briggs 1995). The subsequent high-latitude cooling associated with the onset of northern hemisphere glaciations about 3.0 million years ago caused the extinction of the boreal species in the Arctic area and the beginning of the development of the modern Arctic fauna (Briggs 1995). It has been hypothesised that during this period, one or more ancestor species of *Sebastes* penetrated into the Atlantic by a northern route through the Arctic Ocean, where most of the ancestors perished and only the last branch gave rise to the four Atlantic species (Barsukov and Zakharov 1973).

The minimum percentage in sequence divergence between the haplotypes of *S. alutus* and the North Atlantic *Sebastes* species is 2.3% (eight nucleotide subsitutions) ± 0.8% (standard error). Assuming a divergence time of 3 million years (MY), the probable beginning of the modern Arctic Fauna (see Briggs 1995 and Rocha-Olivares 1999b), the 2.3% ± 0.8% ND3 gene sequence divergence translates to a rate of 0.77% (or 2.7 nucleotide subsitutions) ± 0.27% per million years. The value is quite similar to ND3 divergence rates of about 0.83% per million years reported for *Oncorhynchus* species (McKay et al. 1996). The value is also similar to the assumed divergence rate of the cytochrome b gene in *Sebastes* of about 0.92% ± 0.21% per million years (Rocha-Olivares et al. 1999b). Therefore, using the level ND3 gene sequence divergence, it may be possible to apply a crude times scale to the divergence rates in North Atlantic *Sebastes* phylogeny. On the basis of this crude time scale and under the assumption that ME2 is the most basal haplotype, the divergence times, for example between the ancestral haplotype ME2 and the *S. viviparus* specific haplotype VI2 would be 1.2 million years, the split beween ME2 and the *S. marinus* specific haplotype MA2 as well as between ME2 and the *S. fasciatus* specific haplotype FA4 would be less than one million years (about 0.4 million years).

All rate estimates must be accepted with the problem that they are based on only one gene locus and are extremely vulnerable to violations of the assumption of a constant molecular clock, such as homoplasy (convergent mutations, back-mutations), selection, etc. (Avise 1994).

However, Rocha-Olivares et al. (1999b) suggested in their study based on molecular data under the assumption of a molecular clock that 14 species of *Sebastes* belonging to the subgenus *Sebastomus* could have been formed in the course of only one million years (average rate of at least 1 species every 60,000 years). Therefore, the formation of the North Atlantic *Sebastes* species within about one million years, as indicated by the rates of ND3 gene sequence evolution, is not unlikely.

Nevertheless, using molecular clocks to calculate origin times of biological lineages is still in its infancy and remains controversial (Avise 1994).

6. General conclusions and perspectives

In conclusion this study has provided the first hypothesis for the evolutionary relationship between all four North Atlantic *Sebastes* species based on quantitative phylogenetic analyses of molecular sequence data.

Highly relevant is the low level of genetic variation in the mitochondrial ND3 gene within the North Atlantic. The low level of genetic differentiation, the low resolution in the phylogenetic trees and the star-like topology of the statistical parsimony network are usually attributed to an explosive population expansion after a population bottleneck and/or a founder event. Nevertheless, the degree of genetic variability was large enough to recognise diagnostic mutations characteristic for several of the analysed species. The presence of diagnostic mutations can be interpreted as evidence of recent reproductive isolation between the lineages, even though there is indication of ancient introgressive hybridisation events, in particular between *S. marinus* and *S. fasciatus* on the Flemish Cap and - in case of "giant" *S. marinus* - between *S. marinus* and *S. mentella*.

The evolutionary scenario revealed by the phylogenetic analysis of the ND3 gene data provides the hypothesis that *S. mentella* is the representative of the most basal lineage of North Atlantic *Sebastes*, due to the wide distribution and the basal position of the haplotype ME2 in the statistical parsimony network. This is further supported by the extensive distribution of *S. mentella* throughout the North Atlantic - in contrast to the other species, which show a more restricted area of distribution. Furthermore, there is indication that the *S. viviparus* lineage originated from an earlier split, whereas the other lineages directly descended from the *S. mentella* lineage.

Several findings are largely congruent with the results of other studies, indicating that (i) the North Atlantic *Sebastes* species share a common ancestor and are therefore monophyletic, (ii) *S. alutus* is closely related to the North Atlantic *Sebastes* species, (iii) within the North Atlantic, *S. viviparus* is the most distant species and finally (iv) *S. mentella* is closely related to *S. marinus* (e.g. McGlade et al. 1983; Sundt and Johansen 1998; Rocha-Olivares 1999a).

The existence of two divergent haplotype lineages within *S. marinus* is an unexpected finding and indicates cryptic speciation within North Atlantic *Sebastes*.

Also genetic differences between the two different phenotypes of *S. mentella* were observed, even though this result should be treated with caution, as the differences are based on differences in haplotype frequencies and the sample sizes were too small for population genetic analysis.

It has to be noted that the ND3 gene is only a portion of the entire genome. The evolutionary history of a single gene locus (mitochondrial DNA is inherited as one single unit and thus has been characterised as sampling a single gene) can be different from the average history of an entire genome (Avise 1994). Furthermore, the phylogenetic relationships inferred from the current data set represent only the genealogy of the mitochondrial locus as it evolved through

the maternal lineages (Avise 1994; Stepien and Kocher 1997). Several influences can confound the phylogenetic patterns, such as introgression caused by interspecific hybridisation, homoplasy, differing rates of change of separate loci, as well as selection (McKay et al. 1996; Ballard and Whitlock 2004). This has been shown by the present study: In certain areas, there is a strong indication of introgression, maybe coupled with selection, emphasising the importance of analysing a taxon or species over its full geographical range of distribution and with additional nuclear markers.

Nevertheless, this mtDNA sequence based view on the evolution of the North Atlantic *Sebastes* species provides a framework for further genetic studies on different levels that address more specific questions:

A more extensive study applying nuclear microsatellite and AFLP markers, sampling on a larger geographical scale with larger sample sizes and temporal replication of samples is presented in Chapter 2 and 3 of this thesis and provides more information about potential sibling species, the genetic dynamics within and between the species and the potential for hybridisation and introgression.

Furthermore, a sequence analysis of a higher number of mitochondrial and nuclear genes and non-coding regions combined with the results of the present ND3 gene analysis could enhance our understanding of the phylogenetic relationships and the species structure within this species complex.

Also a further investigation of the extent of introgression of mitochondrial genomes due to ancient hybridisation events - in particular between *S. marinus* and *S. fasciatus* on the Flemish Cap - will require a study including more samples from the Flemish Cap and from the North American shelf in different years, to determine, whether the observation of a complete interspecific replacements of mitochondrial genome of *S. marinus* on Flemish Cap is temporally stable and whether introgressed *S. marinus* populations occur in other regions of the western North Atlantic or whether they are geographically limited to the Flemish Cap. Also further studies of the genetic status of "giant" *S. marinus* will require higher sample sizes and temporal replicates. This, however, will be difficult to achieve, as there have been no reported catches of "giant" *S. marinus* since 1997 (ICES 1998b).

Appendix I

Summary statistics for redfish samples collected for the ND3 gene analysis showing species names, morphological types, geographical regions, position of the sampling locations, sampling years, sampling months, average sampling depths, station numbers, coordinating institutes, cruise identification numbers, sample codes and sample sizes (*n*). Dashes indicate missing information.

			Position (mean)				Depth
Species	**Type**	**Location**	**Longitude**	**Latitude**	**Year**	**Month**	**(m)**
Sebastes mentella	deep-sea	Irminger Sea	615500N	0294900W	1997	May	750
Sebastes mentella	deep-sea	Irminger Sea	613810N	0301160W	1997	May	750
Sebastes mentella	deep-sea	Irminger Sea	622000N	0280500W	2001	June	775
Sebastes mentella	deep-sea	Irminger Sea	622400N	0280900W	2001	June	640
Sebastes mentella	deep-sea	Irminger Sea	580000N	0350400W	2001	June	850
Sebastes mentella	oceanic	Irminger Sea	604700N	0344000W	1996	October	180
Sebastes mentella	oceanic	Irminger Sea	613720N	0302090W	1997	May	725
Sebastes mentella	oceanic	Irminger Sea	622000N	0280500W	2001	June	775
Sebastes mentella	oceanic	Irminger Sea	622400N	0280900W	2001	June	640
Sebastes mentella	oceanic	Irminger Sea	580000N	0350400W	2001	June	850
Sebastes mentella		Norway	711180N	0240550E	2001	October	340
Sebastes mentella		SW-Iceland	632870N	0261700W	2001	October	784
Sebastes mentella		Greenland East	635022N	0363941W	2001	October	332
Sebastes mentella		Greenland East	634863N	0364505W	2001	October	405
Sebastes mentella		Flemish Cap	484245N	0452210W	2001	July	713
Sebastes mentella		Flemish Cap	481993N	0448093W	2001	July	525
Sebastes mentella							
Sebastes marinus		Norway	682030N	0111550E	2001	October	195
Sebastes marinus		SW-Iceland	630100N	0250800W	1997	May	630
Sebastes marinus		SE-Iceland	640143N	0131375W	2001	October	308
Sebastes marinus		Dohrn Bank	653457N	0300652W	2001	October	364
Sebastes marinus		Greenland East	641929N	0360459W	2001	October	256
Sebastes marinus		Greenland East	641664N	0355754W	2001	October	278
Sebastes marinus		Flemish Cap	468257N	0459862W	2001	July	304
Sebastes marinus		Flemish Cap	476687N	0448737W	2001	July	242
Sebastes marinus		Flemish Cap	476315N	0449568W	2001	July	244
Sebastes marinus	giant	Irminger Sea	604320N	0282540W	1996	August	695
Sebastes marinus	giant	Irminger Sea	604270N	0282520W	1996	August	677
Sebastes marinus	giant	Irminger Sea	604329N	0282550W	1996	August	735
Sebastes marinus	giant	Irminger Sea	604364N	0282330W	1996	August	594
Sebastes marinus	giant	Irminger Sea	604399N	0282268W	1996	August	-
Sebastes marinus	giant	Irminger Sea	604390N	0282800W	1996	August	658
Sebastes marinus	giant	Irminger Sea	604400N	0282700W	1996	August	-
Sebastes marinus	giant	Irminger Sea	604500N	0282700W	1996	August	-
Sebastes marinus	giant	Irminger Sea	604200N	0282000W	1996	August	-
Sebastes marinus	giant	Irminger Sea	604200N	0282000W	1996	August	-
Sebastes marinus	giant	Irminger Sea	604400N	0282000W	1996	August	732
Sebastes marinus	giant	Irminger Sea	604400N	0282000W	1996	August	786
Sebastes marinus							

Appendix I (continued)

[1] Federal Research Centre for Fisheries (BFAFi), Hamburg, Germany; [2] Institute of Marine Research, Vigo, Spain; [3] Marine Research Institute (MRI), Reykjavík, Iceland; [4] Institute for Marine Research (IMR) and the University of Bergen (UIB), Department of Fisheries and Marine Biology, Bergen, Norway; [5] Fisheries and Oceans Canada, Prince Rupert, BC; [6] Tristan Natural Resources Department, Edinburgh, Tristan da Cunha

Station	Institute	Cruise id	Sample code	*n*
13	Iceland[3]	BTH11997	MEds97	12
9	Iceland[3]	BTH11997	MEds97	14
274	Iceland[3]	A82001	MEds01	5
275	Iceland[3]	A82001	MEds01	5
304	Iceland[3]	A82001	MEds01	5
14	Iceland[3]	TH11996	MEoc96	12
11	Iceland[3]	BTH11997	MEoc97	14
274	Iceland[3]	A82001	MEoc01	5
275	Iceland[3]	A82001	MEoc01	5
304	Iceland[3]	A82001	MEoc01	5
527	Norway	JH-81516	MENO01	14
410	Iceland[3]	A122001	MESWIc01	10
1009	Germany[1]	WH233	MEEGr01	8
1010	Germany[1]	WH233	MEEGr01	2
55	Spain[2]	CAFC01	MEFC01	4
58	Spain[2]	CAFC01	MEFC01	10
			Total	**130**
560	Norway[4]	JH-81549	MANO01	10
1	Iceland[3]	BTH11997	MASWIc97	11
881	Iceland[3]	B132001	MASEIc01	10
974	Germany[1]	WH233	MADB01	9
995	Germany[1]	WH233	MAEGr01	2
996	Germany[1]	WH233	MAEGr01	8
36	Spain[2]	CAFC01	MAFC01	10
75	Spain[2]	CAFC01	MAFC01	5
76	Spain[2]	CAFC01	MAFC01	5
24	Iceland[3]	TJ11996	MAG96	1
25	Iceland[3]	TJ11996	MAG96	1
26	Iceland[3]	TJ11996	MAG96	2
27	Iceland[3]	TJ11996	MAG96	3
28	Iceland[3]	TJ11996	MAG96	3
29	Iceland[3]	TJ11996	MAG96	5
35	Iceland[3]	TJ11996	MAG96	2
36	Iceland[3]	TJ11996	MAG96	4
37	Iceland[3]	TJ11996	MAG96	3
38	Iceland[3]	TJ11996	MAG96	1
49	Iceland[3]	TJ11996	MAG96	10
50	Iceland[3]	TJ11996	MAG96	2
			Total	**107**

Appendix I (continued)

Summary statistics for redfish samples collected for the ND3 gene analysis showing species names, morphological types, geographical regions, position of the sampling locations, sampling years, sampling months, station numbers, average sampling depth, coordinating institute, cruise identification number, sample codes and sample sizes (*n*). Dashes indicate missing information.

			Position (mean)				Depth
Species	**Type**	**Location**	**Longitude**	**Latitude**	**Year**	**Month**	**(m)**
Sebastes fasciatus		Flemish Cap	461100N	0463700W	1997	July	392
Sebastes fasciatus		Flemish Cap	451300N	0462700W	1997	July	415
Sebastes fasciatus		Flemish Cap	466842N	0453872W	2001	July	240
Sebastes fasciatus		Flemish Cap	481845N	0449435W	2001	July	461
Sebastes fasciatus		Flemish Cap	479502N	0447565W	2001	July	323
Sebastes fasciatus							
Sebastes viviparus		Norway	701250N	0305630E	2001	October	150
Sebastes viviparus		Norway	704790N	0293710E	2001	October	110
Sebastes viviparus		Norway	710560N	0283650E	2001	October	170
Sebastes viviparus		Norway	711390N	0273140E	2001	October	270
Sebastes viviparus		SW-Iceland	632100N	0253900W	1996	November	370
Sebastes viviparus		SW-Iceland	635998N	0233273W	2001	March	135
Sebastes viviparus							
Sebastes alutus		Cape St. James	514900N	1303400W	2001	August	293
Sebastes capensis		Tristan da Cunha	371500S	0123000W	2002	February	-
Helicolenus dactylopterus		Shetland Islands	601250N	0045300W	2000	May	188
Total							

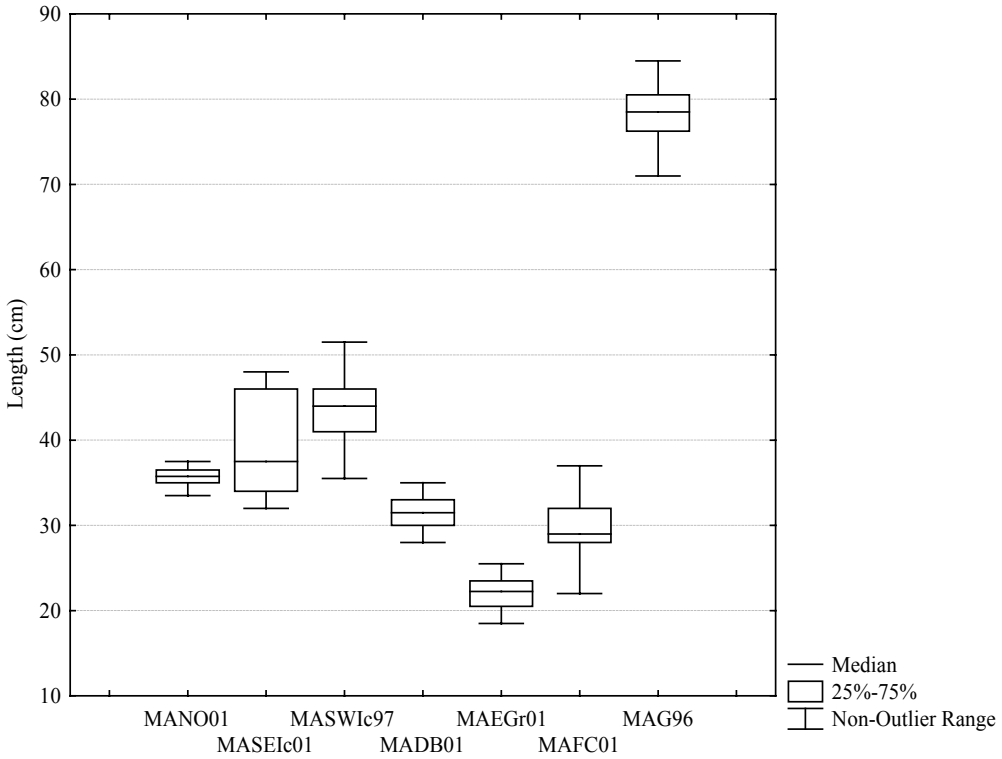
Appendix I (continued)

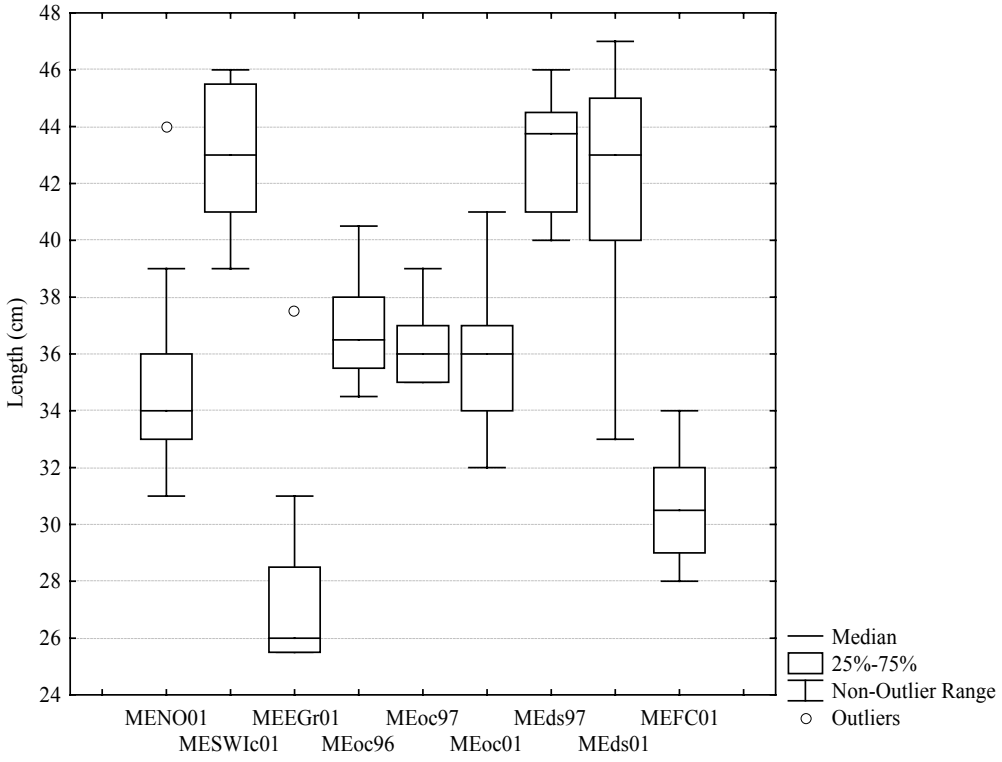
[1] Federal Research Centre for Fisheries (BFAFi), Hamburg, Germany; [2] Institute of Marine Research, Vigo, Spain; [3] Marine Research Institute (MRI), Reykjavík, Iceland; [4] Institute for Marine Research (IMR) and the University of Bergen (UIB), Department of Fisheries and Marine Biology, Bergen, Norway; [5] Fisheries and Oceans Canada, Prince Rupert, BC; [6] Tristan Natural Resources Department, Edinburgh, Tristan da Cunha

Station	Nation	Cruise id	Sample code	*n*
81	Spain[2]	CAFC97	FAFC97	3
88	Spain[2]	CAFC97	FAFC97	10
40	Spain[2]	CAFC01	FAFC01	5
59	Spain[2]	CAFC01	FAFC01	6
63	Spain[2]	CAFC01	FAFC01	10
			Total	**34**
514	Norway[4]	JH-81508	VINO01	1
516	Norway[4]	JH-81508	VINO01	1
518	Norway[4]	JH-81508	VINO01	4
519	Norway[4]	JH-81508	VINO01	8
10	Iceland[3]	TJB11996	VSWIc96	12
20	Iceland[3]	TBI-02	VSWIc01	16
			Total	**42**
-	Canada[5]	-	ALUT	10
1	Tristan da Cunha[6]	GOUGH	CAP	10
368	Germany[1]	WH216	HEL	4
				337

Appendix II

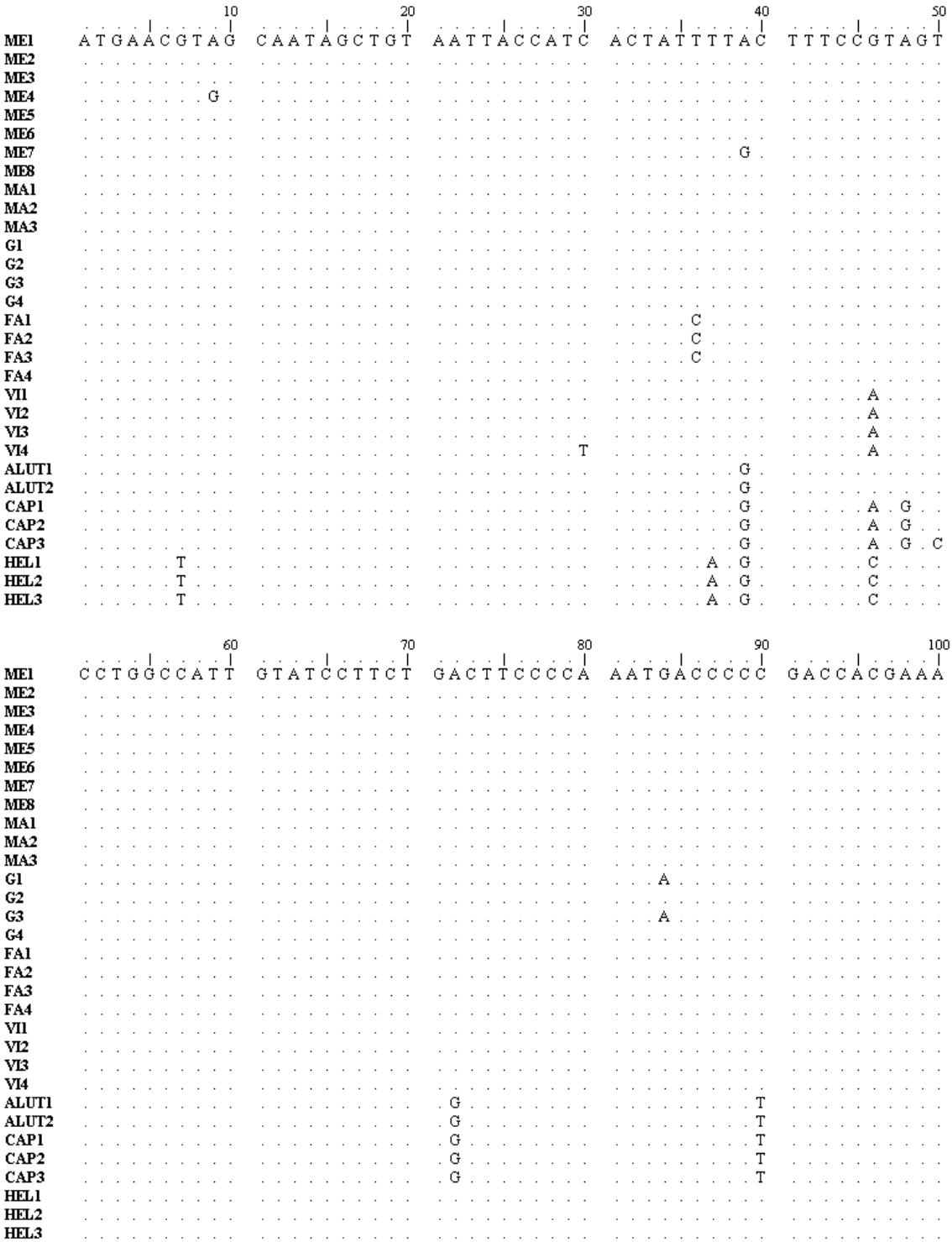
Box-and-whisker plots of the length distributions (total length) of sampled *S. marinus* (above) and *S. mentella* (below) sequenced for this study. Plots were generated in STATISTICA, version 6.1 (StatSoft, Inc. 2003). The line within the box represents the median of the distribution. The 25th and 75th percentiles are represented by the top and bottom box edges, respectively. Whiskers represent the lowest and highest values (Non-Outlier range). Sample codes are explained in Table 1.1 and Appendix I.



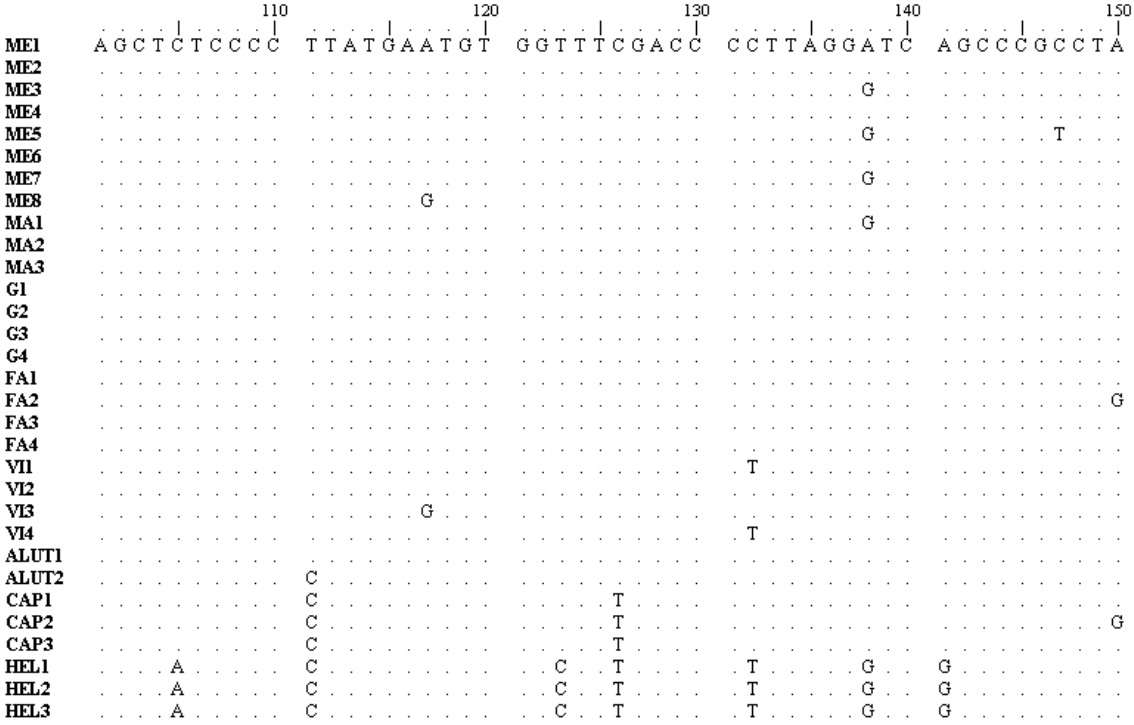


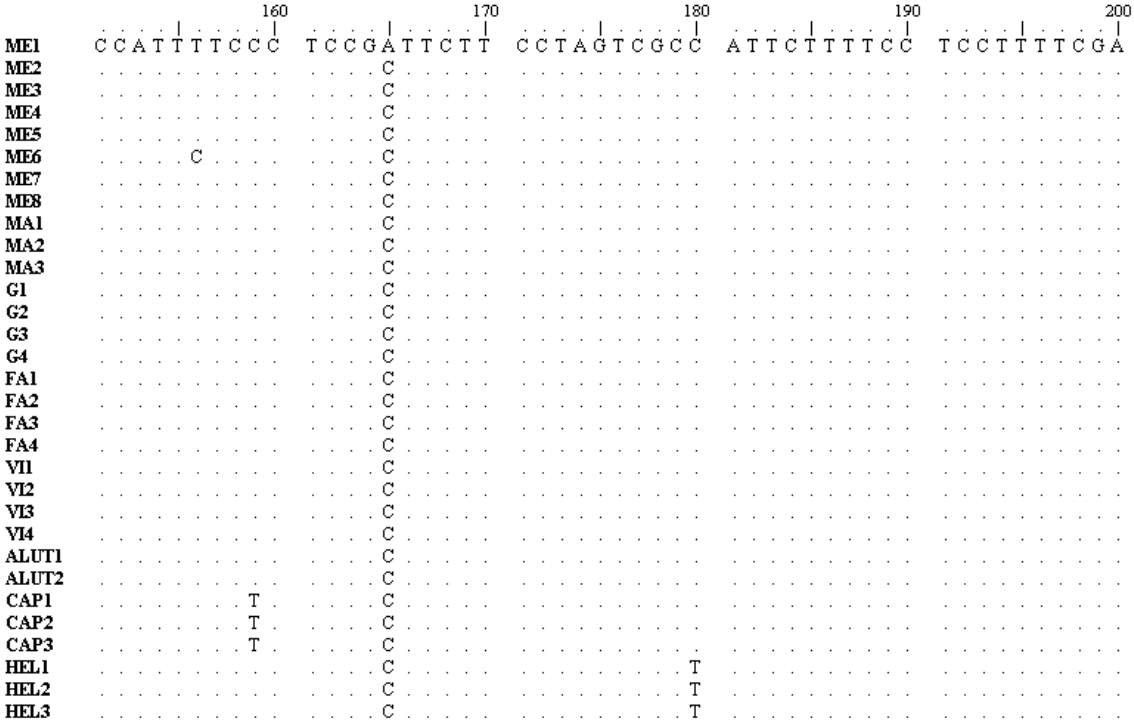
Appendix III

Alignment of the entire nucleotide sequence of the mitochondrial ND3 gene for all haplotypes found in the *Sebastes* and *Helicolenus dactylopterus* samples. Matching nucleotides are indicated with a dot. The sequences shown are reading 5' to 3' on the heavy (H) strand.

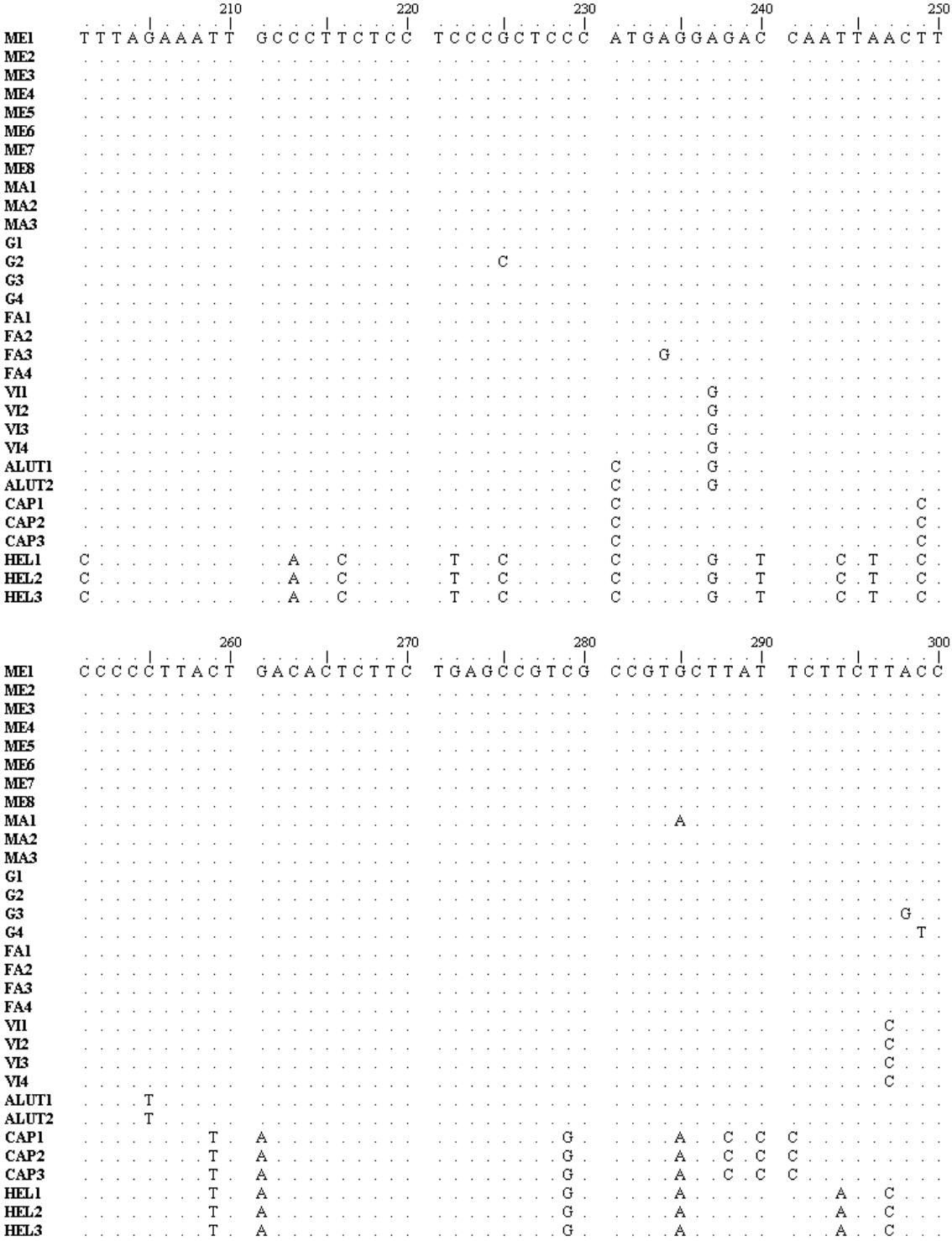


Appendix III (continued)

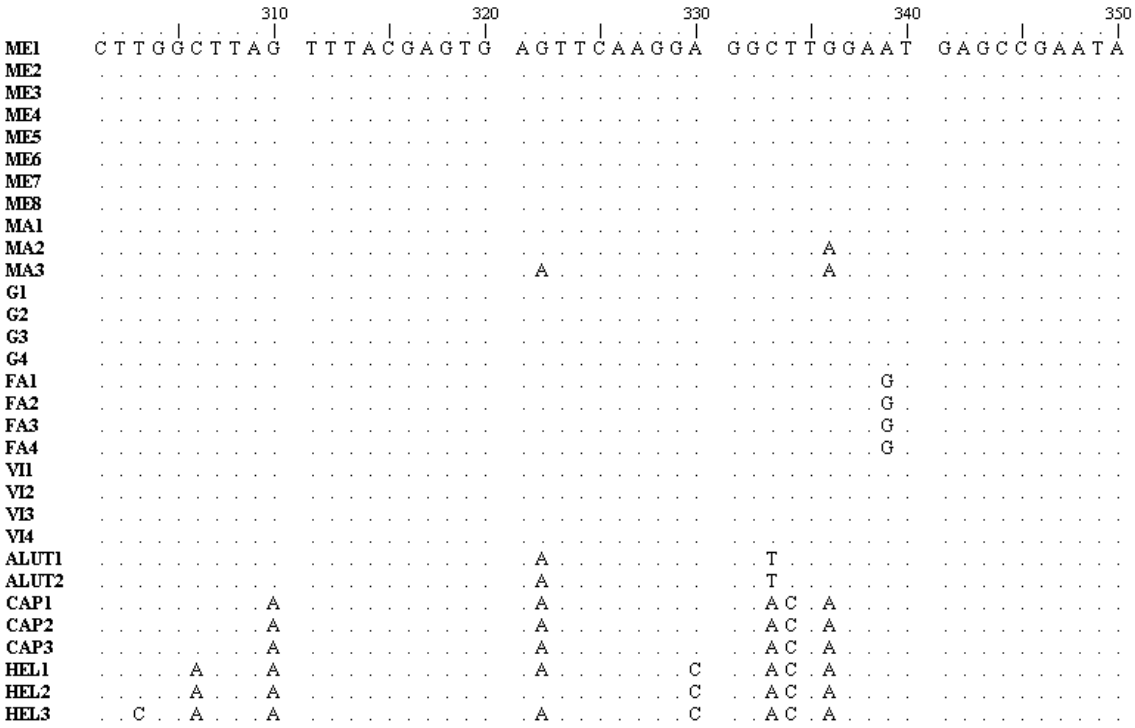


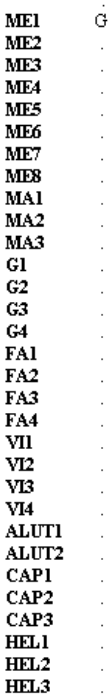


Appendix III (continued)



Appendix III (continued)





CHAPTER 2

Genetic species and population structure of North Atlantic redfish (genus *Sebastes*; Cuvier 1829) based on microsatellite polymorphism

1. Abstract

Eight highly variable microsatellite loci were used to investigate the genetic structure of species and populations of North Atlantic redfish (genus *Sebastes*; Cuvier 1829), with special emphasis on the species complex in the Irminger Sea and on the continental slopes of Greenland, Iceland and the Faroe Islands. For this purpose, patterns of genetic diversity among 16 samples of *Sebastes marinus* (n = 924), 10 samples of *S. mentella* (n = 558), and samples of *S. fasciatus* (n = 81) and *S. viviparus* (n = 42), collected from different locations throughout the North Atlantic, were determined and assessed. The results of factorial correspondence analyses (FCA) and Bayesian cluster analyses indicate the prevalence of five genetically distinct clusters among the four species studied. Individuals of *S. viviparus*, *S. fasciatus* and *S. mentella* grouped into distinct clusters, whereas samples of individuals pre-classified as *S. marinus* according to external morphological characters clustered into two genetically isolated groups. There is indication that one of the two groups represents the species *S. marinus*. The genetic differences found between these two groups were of the same order of magnitude as the genetic differences observed between the North Atlantic *Sebastes* species (F_{ST} values between 0.079 and 0.169), pointing to the existence of a fifth - apparently cryptic - species of *Sebastes* occurring on the continental slopes of Greenland, Iceland and the Faroe Islands. The potential of microsatellites to discriminate individuals from the five different clusters was high. On average, it was possible to assign 94% of the individuals to one of the five clusters using Bayesian cluster analysis. The results further indicate that there is only restricted, if any, hybridisation between *S. marinus*, *S. mentella* and the cryptic species in the areas off Iceland, the Faroe Islands, East and West Greenland, as well as between *S. fasciatus*, *S. marinus* and *S. mentella* on the Flemish Cap.

Furthermore, the results of the analyses of intraspecific genetic structure indicate that there are at least three genetically distinct population units (significant F_{ST} values between 0.009 and 0.045) of *S. marinus* in the geographic area covered in this study, represented by the samples from (i) the Flemish Cap, (ii) Greenland and (iii) Norway, Iceland and the Faroe Islands. The microsatellite analysis revealed also weak but significant genetic structure in *S. mentella*, even though the levels of genetic differentiation were lower than those observed in *S. marinus*. Pairwise F_{ST} values (between 0.007 and 0.018) indicated significant genetic differentiation of samples of *S. mentella* from the Flemish Cap, the central Irminger Sea and the samples from Greenland, Iceland, the southern Irminger Sea, NAFO area 1F and NAFO area 2J, whereby the samples collected from the latter five areas were genetically relatively homogeneous. No indication of genetic structure was found in the newly identified cryptic species of North Atlantic *Sebastes.* As the *Sebastes* species are long-lived and the distribution of potential populations can vary temporally, a more detailed temporal and geographic sampling is needed before definite conclusions about the population structure of the *Sebastes* species occurring in the Irminger Sea and adjacent waters can be drawn.

2. Introduction

In the North Atlantic, the viviparous redfishes of the genus *Sebastes* (Cuvier 1829) form a complex consisting of four closely related species, *S. mentella* (Travin 1951), *S. marinus* (Linné 1758), *S. fasciatus* (Storer 1854) and *S. viviparus* (Krøyer 1845). Except *S. viviparus*, all North Atlantic redfish species are commercially important fishery resources.

S. mentella and *S. marinus* are widely distributed throughout the North Atlantic, while *S. fasciatus* and *S. viviparus* are essentially restricted to the Northwest Atlantic and to the Northeast Atlantic, respectively (Barsukov 1973; Barsukov et al. 1990).

Due to their morphological similarity with widely overlapping meristic and morphological characters (e.g. Barsukov et al. 1985; Rubec et al. 1991) and the low levels of inter- and intraspecific genetic variation (Payne and Ni 1982; McGlade et al. 1983; Rehbein 1983; Trottier et al. 1989; Nedreaas and Nævdal 1989; 1991a and 1991b; Rubec et al. 1991; Nedreaas et al. 1994; Bentzen et al. 1998; Johansen et al. 1998; Sundt and Johansen 1998), unambiguous species identification is difficult.

In particular, the two commercially important species *S. mentella* and *S. marinus* occurring sympatrically on the continental slopes of Greenland, Iceland and the Faroe Islands are morphologically remarkably similar and some authors mention fish that possess morphological characters of both species in a varying manner (Kotthaus 1961; Nævdal 1978). Also an overlap in electrophoretic mobility patterns of hemoglobin variants, blood serum proteins and enzymes has been observed and several authors discuss the possibility of hybridisation between *S. marinus* and *S. mentella* in Icelandic waters and off Greenland (Altukhov and Nefyodov 1968; Nævdal 1978; Nedreaas and Nævdal 1991a; Nedreaas et al. 1994).

The problem of species identification and the possibility of hybridisation between the species have also hampered our understanding of the population structure of *S. marinus* and *S. mentella* in the central North Atlantic. Therefore, the division of the species into stock units for management purposes has been done more on a geographical than on a biological basis (Reinert and Lastein 1992). *S. marinus* and *S. mentella* in the central North Atlantic are presently treated as three management units: *S. marinus* and *S. mentella* on the continental slopes of Greenland, Iceland and the Faroe Islands and *S. mentella* distributed over the pelagic Irminger Sea waters.

Even though numerous studies on genetics, biological markers and distribution patterns were carried out, interpretations of the evidence on stock structure in *S. mentella* are still diverging and there has been a strong controversy about whether there are one, two or three populations of *S. mentella* present in the Irminger Sea and adjacent areas (ICES 1998a; Magnússon and Magnússon 1995; Johansen et al. 2000b; Roques et al. 2002; Saborido-Rey et al. 2005; Stransky 2005).

Furthermore, some studies suggest a more complicated population structure within *S. marinus* (Nedreaas and Nævdal 1991a, Reinert and Lastein 1992; Nedreeas 1994; Johansen et al. 2000a). Also the taxonomic status of the so-called "giant" redfish (or "giant" *S. marinus*), individuals with an average total length above 60 cm that are morphological identified as *S. marinus*, but are separated from *S. marinus* due to some morphological and hemoglobin characteristics, remains unclear (Kotthaus 1961; Altukhov and Nefyodov 1968; Johansen et al. 2000a).

Due to the serious difficulties in species and population discrimination, redfish in the North Atlantic is still exploited and managed as one species (Barsukov et al. 1985; Gascon 2003). Nevertheless, a clear understanding of the species and population structure of these exploited species is essential for an adequate analysis of population dynamics and for the conservation of genetic resources (Carvalho and Hauser 1994; Ruzzante et al. 1996) - and therefore a prerequisite to proper management of North Atlantic *Sebastes*.

Since the early to mid-1990s, microsatellites have gradually replaced allozymes and other molecular markers as the tool of choice for numerous applications in evolutionary and population genetic studies, due to their generally high level of polymorphism, codominant inheritance, assumed selective neutrality and ease and reliability of scoring (Jarne and Lagoda 1996; Luikart and England 1999). Microsatellites are of particular interest in studies requiring fine-scale resolution for which other markers may have reached their limits of applications e.g. in genetic studies of marine fish with supposedly low levels of genetic differentiation (Carvalho and Hauser 1998; Estoup and Angers 1998).

Furthermore, the codominant mode of inheritance in combination with the high levels of variability observed in microsatellites result in new, powerful applications (reviewed in Luikart and England 1999). Recently, a number of new statistical tools for individual-based analysis have been developed, which allow to infer the number of populations in a given data set, the precise assignment of individuals to a set of baseline populations or species, as well as the estimation of the genetic contribution of the baseline populations or species to the genotype of an individual (Cornuet et al. 1999; Pritchard et al. 2000; Falush et al. 2003). These methods have been successfully used to assign individual cod to their population of origin (e.g. Nielsen et al. 2001) and to detect hybrid-zones in populations of cod (Nielsen et al. 2003) and turbot (Nielsen et al. 2004).

Therefore, microsatellites are the molecular markers of choice for detecting the inter- and intraspecific genetic structure of *Sebastes* in the central North Atlantic, according to previous studies on Northwest Atlantic redfish (Roques et al. 1999a; 1999b; 2000; 2001; 2002) and Pacific *Sebastes* populations (Miller et al. 2000; Buonaccorsi et al. 2002; Rocha-Olivares et al. 2003; Matala et al. 2004a and 2004b).

In the present study, eight microsatellite loci were used to investigate the species structure of the four North Atlantic redfish species and possible hybridisation between them. The population structure within *S. mentella* and *S. marinus* in the central North Atlantic was

studied, with special emphasis on *S. marinus*. The use of microsatellites for the assignment of single individuals to species and populations was evaluated.

3. Material and Methods

3.1 Sample collection

North Atlantic redfish species were sampled by different scientists from Norwegian, Icelandic, Spanish and German institutes (see Appendix I) during research surveys across the North Atlantic from 1996 to 2001.

Individual redfish were pre-classified into species on board using morphometrical (e.g. eye diameter, length and inclination of the preopercular spines) and morphological characters (e.g. spines, body shape, symphysial tubercle, colour) usually used for redfish (genus *Sebastes*) identification (Barsukov 1973; Barsukov et al. 1985).

In addition, samples collected on the shelf of Greenland in the year 2001 were classified into species using hemoglobin electrophoresis (Nedreaas and Nævdal 1989; 1991a), as species identification - particularly the differentiation of *S. marinus* and *S. mentella* specimens - using standard characters was ambiguous. Hemoglobin analyses were performed by a Norwegian specialist[1] directly on board.

Gill samples for genetic analyses were taken either directly from the fish on board of the research vessels or afterwards in the laboratory at the different institutes (see Appendix I) from fish that had been frozen whole for subsequent morphometrical analyses[2]. The gill samples were preserved in 100% ethanol.

In total 1595 individuals of the genus *Sebastes* were collected from different locations across the North Atlantic: 15 samples of *S. marinus* (902 individuals), 10 samples of *S. mentella* (558 individuals), two samples of *S. fasciatus* (81 individuals) and three samples of *S. viviparus* (42 individuals) were collected. In addition, DNA samples of 12 extremely large *S. marinus* specimens (average length 78 cm) defined as "giant" *S. marinus* by Icelandic scientists[3] - collected in the Irminger Sea in 1996 - were provided by the cooperating Icelandic institute (see Appendix I).

Total lengths and sex were recorded as standard procedure for all samples.

Table 2.1 gives an overview of the samples collected. In Appendix I the exact sampling information (e.g. position, month, depth) is listed. The North Atlantic sampling locations of the *Sebastes* samples are pictured in detail in Figure 2.1 (*S. marinus*), 2.2 (*S. mentella*) and 2.3 (*S. viviparus* and *S. fasciatus*). Size distributions of sampled *S. marinus* and *S. mentella* are given in Appendix IIa and IIb, respectively.

[1] Hemoglobin analyses were performed by Torild Johansen, University of Bergen (UIB), Norway.

[2] Morphometrical analyses were performed by Dolores Garabana, Institute of Marine Research, Vigo, Spain.

[3] Scientists from the Marine Research Institute (MRI), Reykjavík, Iceland

Table 2.1. Summary statistics for redfish samples collected showing species names, morphological types, geographical origins, sampling years, sampling months, sample codes, sample sizes (*n*) and sex (number of males, number of females and number of individuals with unknown sex). The species names are based on the pre-classification on board using standard characters (Barsukov 1973; Barsukov et al. 1985) and, in case of the samples of *S. marinus* and *S. mentella* from Greenland collected in the year 2001, on hemoglobin electrophoretic patterns (Nedreaas and Nævdal 1989; 1991a).

						Sample size (*n*)			
Species	Type	Location	Year	Month	Sample code	Males	Females	Unknown	Total
S. marinus		Norway	2001	October	MANO01	61	17	0	78
S. marinus		Faroe Islands	2002	September	MAEF02	22	37	0	59
S. marinus		Faroe Islands	2002	September	MAWF02	33	6	0	39
S. marinus		NE-Iceland	2000	October	MANEIc00	25	25	2	52
S. marinus		SE-Iceland	2000	October	MASEIc00	44	57	0	101
S. marinus		SE-Iceland	2001	October	MASEIc01	34	20	1	55
S. marinus		SW-Iceland	2000	June	MASWIc00	32	17	11	60
S. marinus		SW-Iceland	2001	October	MASWIc01	26	26	0	52
S. marinus		Greenland East/Dohrn Bank	2001	October	MAEGrDB01	24	25	1	50
S. marinus		Greenland East	2000	October	MAEGr00	26	45	0	71
S. marinus		Greenland East	2000	October	MAEGr00b	30	17	0	47
S. marinus		Greenland East	2001	October	MAEGr01	43	17	0	60
S. marinus		Greenland West	2001	November	MAWGr01	23	25	2	50
S. marinus		Flemish Cap	2001	July	MAFC01	28	26	1	55
S. marinus		Flemish Cap	2001	July	MAFC01b	44	29	0	73
S. marinus	giant	Irminger Sea	1996	August	MAG96	1	11	0	12
S. marinus		Total				496	400	18	914
S. mentella		SE-Iceland	2001	October	MESEIc01	24	26	1	51
S. mentella		Central Irminger Sea	2001	June	MECIrm01	30	26	1	57
S. mentella		Central Irminger Sea	2003	June	MECIrm03	52	14	0	66
S. mentella		Greenland East	2000	September	MEEGr00	42	25	1	68
S. mentella		Greenland East	2001	October	MEEGr01	32	20	6	58
S. mentella		Greenland West	2001	November	MEWGr01	0	0	51	51
S. mentella		Southern Irminger Sea	2001	June	MESIrm01	35	13	0	48
S. mentella		NAFO1F	2001	July	MENAFO1F01	34	14	0	48
S. mentella		NAFO2J	2001	September	MENAFO2J01	37	13	0	50
S. mentella		Flemish Cap	2001	July	MEFC01	31	29	1	61
S. mentella		Total				317	180	61	558
S. fasciatus		Flemish Cap	1997	July	FAFC97	10	3	0	13
S. fasciatus		Flemish Cap	2001	July	FAFC01	26	42	0	68
S. fasciatus		Total				36	45	0	81
S. viviparus		SW-Iceland	1996	November	VISWIc96	7	5	0	12
S. viviparus		SW-Iceland	2001	March	VISWIc01	0	0	16	16
S. viviparus		Norway	2001	October	VINO01	5	9	0	14
S. viviparus		Total				12	14	16	42
Sebastes		Total							1595

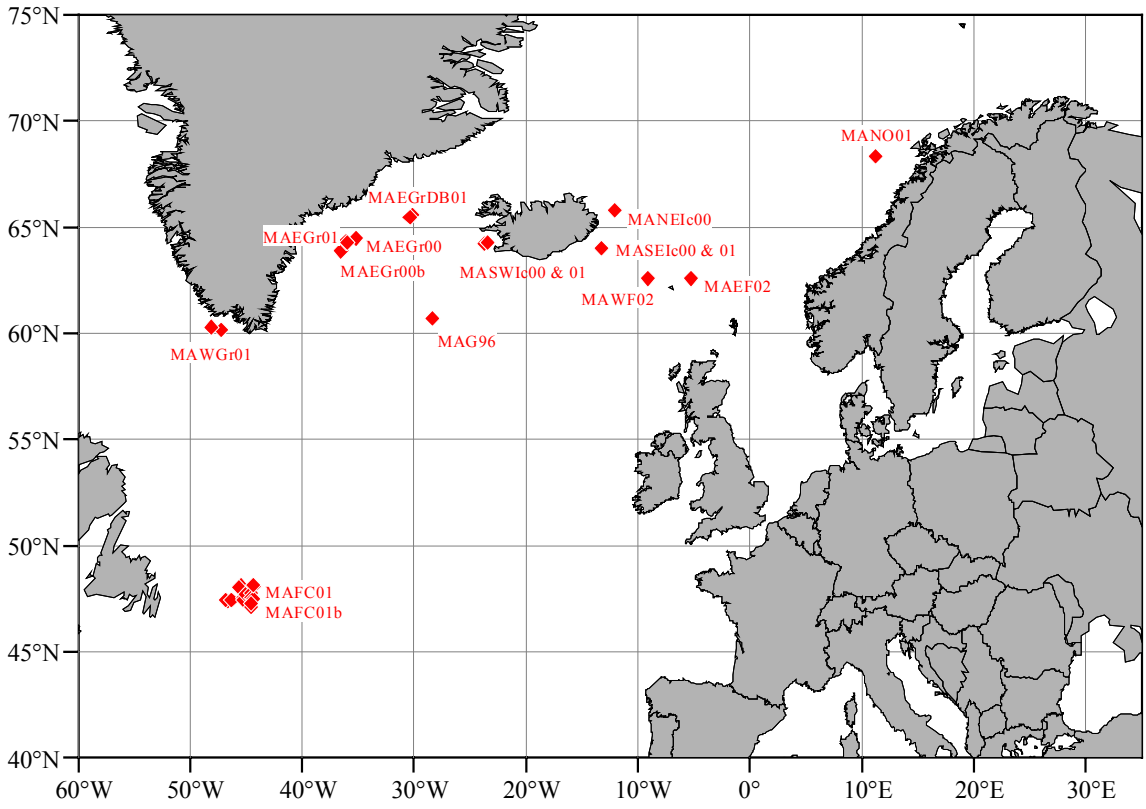


Figure 2.1. North Atlantic sampling locations for the microsatellite analysis of *S. marinus*. Sample codes are explained in Table 2.1.

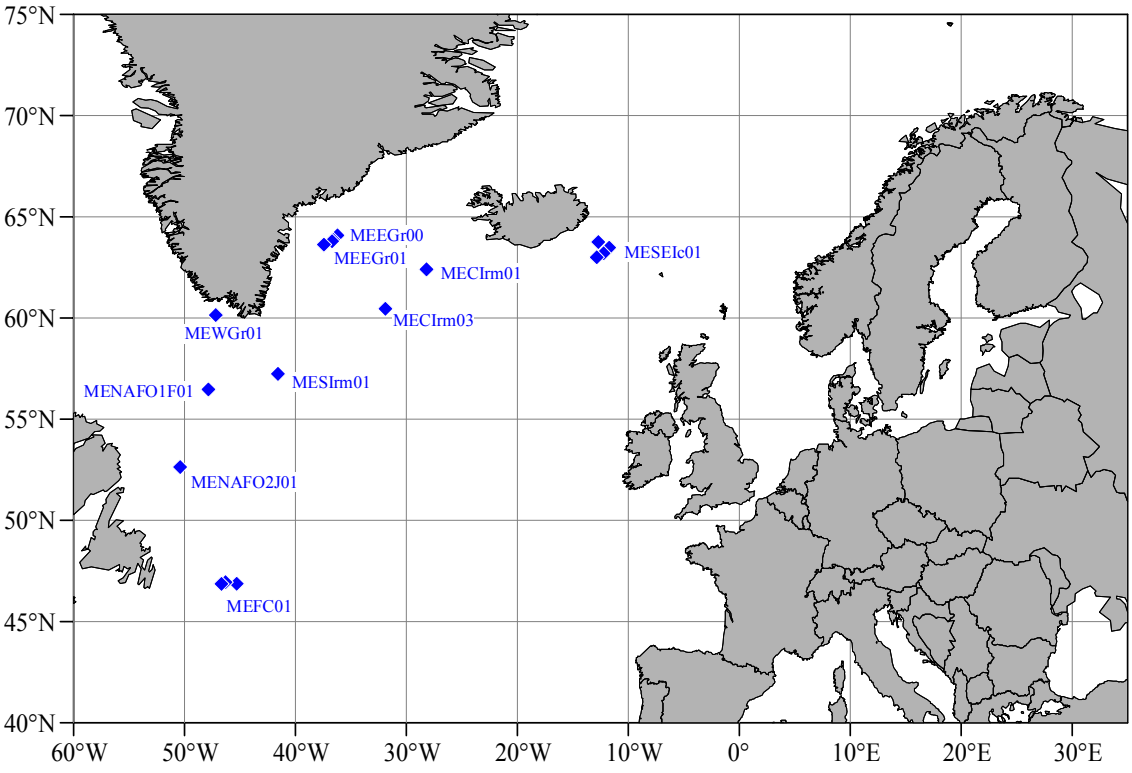


Figure 2.2. North Atlantic sampling locations for the microsatellite analysis of *S. mentella*. Sample codes are explained in Table 2.1.

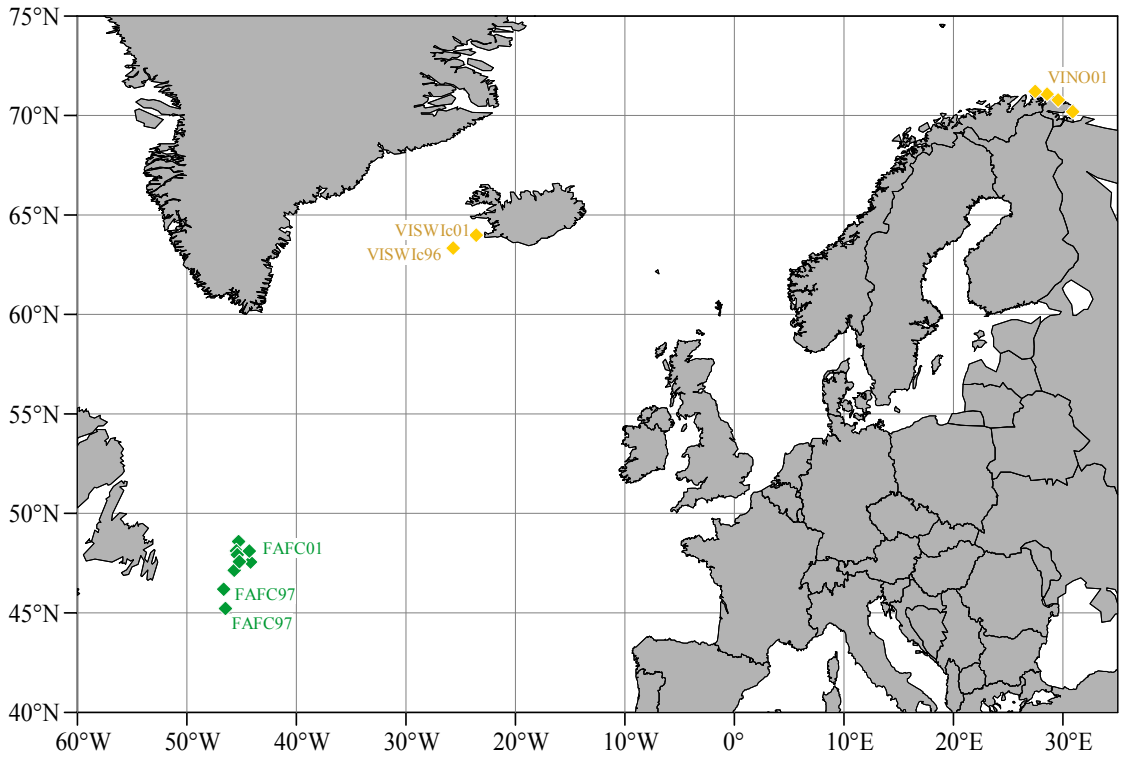


Figure 2.3. North Atlantic sampling locations for the microsatellite analysis of *S. viviparus* (yellow squares) and *S. fasciatus* (green squares). Sample codes are explained in Table 2.1.

3.2 DNA extraction

Total genomic DNA was extracted with the DNeasy™ Tissue Kit (Qiagen) following manufacturers instructions.

3.3 DNA quantitation

Quantitation of DNA was carried out using the fluorometric DNA assay with Hoechst 33258 Dye as described in Chapter 1 (paragraph 3.2.2) of this thesis.

3.4 Polymerase chain reaction (PCR)

Microsatellites can be easily amplified by polymerase chain reaction (PCR: Mullis et al. 1986; Mullis and Falcoona 1987; Saiki et al. 1988), when the DNA sequences of the flanking regions of the respective microsatellite locus are known (Estoup and Angers 1998).
Eight dinucleotide microsatellite loci (SEB9, SEB25, SEB30, SEB31, SEB33, SEB37, SEB45 and SEB46) were analysed using eight sets of primers developed by Roques et al. (1999b). SEB9, SEB25, SEB33, SEB45 and SEB46 consist of perfect dinucleotide repeats, while SEB30 and SEB31 are imperfect microsatellites, which have a few interrupting bases in an

otherwise perfect repeat array. SEB37 is a compound microsatellite, consisting of two neighbouring repeated sequences composed of two different dinucleotide repeats (Roques et al. 1999b). One primer of each primer set was end-labelled with the fluorescent dye IRD800. Repeat sequences and primer sequences are listed in Table 2.2. All microsatellite primers were purchased from MWG Biotech.

Table 2.2. Repeat sequences, primer sequences (F' = forward; R' = reverse), concentrations of the fluorescent-labelled primers (IRD800-labelled) and annealing temperatures for the eight microsatellite loci used in this study. Primary reference for their application: Roques et al. (1999b).

Locus	Sequence (5'-3')	Conc. IRD800-labelled primer (mM)	Annealing temperature (°C)
SEB9	$(GA)_{22}$	0.6	62
Primer F'	AAG GCT GAC TCT GAG TGG GA		
Primer R'	CTC TGA GTC TAT GTA TCT GGC T		
SEB25	$(CA)_{15}$	0.9	62
Primer F'	CAG CTT GAC GTG AGG GGA		
Primer R'	GTG CCT GTT TAG GGT GTT CTT		
SEB30	$(CT)_7$ TT$(CT)_{15}$TT$(CT)_{29}$	0.3	68
Primer F'	CTG TTG GAC AGA TAA AGA CGC		
Primer R'	GGT GAT ATT GCT GCT GGT AGA T		
SEB31	$(GA)_2$GGGG$(GA)_{22}$GGG$(GA)_4$	0.3	62
Primer F'	GTG AGA CCA GTA ATA AGG GCA		
Primer R'	TAC TTC TCG ACT GTG GTG		
SEB33	$(GA)_{39}$	0.6	62
Primer F'	CAG ATG TTG GTA GAC GCA AGC A		
Primer R'	AGT CCA GTG TCC ATC CTC CTT		
SEB37	$(GT)_{10}(GA)_{16}$	0.3	65
Primer F'	GTA CAG TCC ATT CAG CTT TGA		
Primer R'	AGG GTG TGT GGA AGA AAT AGT		
SEB45	$(GT)_{24}$	0.3	68
Primer F'	GAG GAG GAA AAG ACT GGA CAG A		
Primer R'	GAA AGA TGG TGA GCA GCG ATG A		
SEB46	$(GT)_{20}$	0.3	62
Primer F'	GCT GAT GTT GCT CCT AAA GAA		
Primer R'	CTC TTC ATG TCA ATC CTG CCT		

PCR amplifications were performed following a protocol by Trautner (2000) with some modifications. PCR amplifications were carried out in 5 µl reaction volumes in microtiter plates, using 1.5 µl of DNA template, 0.6 mM of the reverse (R') primers, 0.3 mM to 0.9 mM of the fluorescent-labelled (IRD800) forward (F') primers (see Table 2.2), 0.2 mM of each deoxynucleoside triphosphate (dNTPs; Gibco/Life Technologies), 1.5 mM $MgCl_2$ (MWG Biotech), 1x *Taq* polymerase specific PCR buffer (MWG Biotech) and 0.24 units *Taq* polymerase (MWG Biotech). A drop of mineral oil was used to cover the reactions to avoid evaporation.
The PCRs were performed with a T-Gradient thermocycler (Biometra), starting with an initial denaturing step of 3 minutes at 95°C, followed by 30 cycles of 30 seconds denaturation at 95°C, one minute primer annealing (annealing temperatures are listed in Table 2.2) and one minute and 30 seconds primer extension at 72°C and a final primer extension step for 10 minutes at 72°C.
With these PCR conditions, the individuals listed in Table 2.1 were analysed for all eight loci. In case of "giant" *S. marinus*, it was not possible to analyse locus SEB30, as not enough DNA material was available to screen all loci.

3.5 Analysis of microsatellite fragments on an automated sequencer

Microsatellite fragments were separated, visualised and analysed on a denaturing polyacrylamide gel in an automated sequencer (LI-COR™ 4200 GeneReadIR DNA system) according to manufacturers recommendations. For the fragment analysis, the labelled PCR products were mixed with 5 µl gel loading buffer (98% formamide, 10 mM EDTA, 0.025% bromophenol blue, 0.025% xylen cyanole), denaturated at 95°C for 5 min and immediately after denaturation placed on ice. Approximately 0.5 to 1.0 µl of the mixture was loaded quickly on a 25 cm polyacrylamide gel (SequaGel™ XR, Biozym, prepared according to manufacturers recommendations). The gel electrophoresis was run at 2500 V for two hours. The data were collected automatically and analysed on computer using the Windows®-based software Saga Version 2.0 (LI-COR™).

3.6 Determination of allele sizes

Allele sizes of 12 individuals of each species were determined by reference to a sequence reaction of the pGEM-T Easy Vector System II (Promega; see Chapter 1). For each microsatellite locus, three to four of these individuals of known length were chosen as sizing standards. These reference samples were chosen to cover as much as possible of the expected allele size range. Allele sizes for all samples analysed were determined by comparison to the reference samples of known length, which were run in every tenth lane on each gel. Scoring

of the alleles was performed using the Windows®-based software Saga Version 2.0 (LI-COR™).

3.7 Genotyping

All loci analysed in this study showed a multiple banding pattern known as stutter, shadow banding or slippage (Weber and May 1989). It is generally thought that stutter peaks are the product of slippage by the DNA polymerase during the PCR (slipped-strand mispairing: Hauge and Litt 1993). Murray et al. (1993) reported that stutter peaks are characterised by 2 bp deletions in the repeat region generated during the PCR.
Such stutters are common in dinucleotide loci, making it difficult to discriminate between homozygotes and heterozygotes. Such genotyping errors can cause deviations from Hardy-Weinberg proportions (see paragraph 3.8.1) - in particular heterozygote deficiencies - potentially biasing population genetic analyses (van Oosterhout et al. 2004). Furthermore, variable scoring of stutter peaks can affect the accuracy in reporting of allele sizes.
To avoid these problems, scoring of alleles was performed consistently, using always the following criteria:
Allele size was determined by the number of bases of the darkest band (= the band with the highest fluorescence intensity) in the stuttering pattern. In case of bands with similar fluorescence intensities, the band with the highest molecular weight was scored.
As discrimination between homozygotes and heterozygotes with only one repeat between the two alleles is sometimes difficult when a microsatellite locus reveals a stuttering pattern (Estoup and Angers 1998), homozygotes and heterozygotes were discriminated by analysing the number and relative intensity of the bands in the stutter ladder (O'Reilly and Wright 1995). If the band with the highest molecular weight was of lower or same intensity as the neighbouring (two base pairs shorter) band, then both were marked as alleles. If this was not the case, the individual was scored as homozygote.

3.8 Statistical analysis

3.8.1 Descriptive statistics

First the intra- and inter-sample genetic variation was estimated based on the allelic composition of the 28 samples (16 samples of *S. marinus*, 10 samples of *S. mentella* and one sample of each *S. fasciatus* and *S. viviparus* - due to the small sample sizes, the genetic data of the three *S. viviparus* samples were pooled and also the two samples of *S. fasciatus* were pooled).
The number of alleles (A), unbiased gene diversity (H_e) and observed heterozygosity (H_o) were calculated for each locus within each sample using the software GENETIX version 4.04

(Belkhir et al. 1996-2002). Allelic richness was calculated in FSTAT Version 2.9.3 (Goudet 1995; 2001).
Allelic richness is defined as the number of alleles in a sample, standardised for sample size. Gene diversity (H_e) is the expected level of heterozygosity in a randomly mating population, given observed allelic frequencies. Observed heterozygosity (H_o) is the observed proportion of heterozygotes in a given data set.
Subsequently, it was tested whether the allele frequencies at each locus were in Hardy-Weinberg equilibrium (HWE). The Hardy-Weinberg principle (HWP) states that in a population allelic frequencies and genotypic ratios will remain constant from generation to generation, assuming the organism is diploid, reproduction is sexual, mating is random, generations do not overlap, the population is very large (in theory: infinite) and experiences no drift, no selection, no mutation and no migration (Hartl and Clark 1997).
In the Hardy-Weinberg model, assuming the simplest case of a single locus with two alleles (*a* and *b*), the mathematical relation between allele frequencies (p_a and p_b) and the genotype frequencies is given by

$$p_a^2 + 2p_a p_b + p_b^2 = 1$$

with p_a = frequency of allele *a*, p_b = frequency of allele *b*, p_a^2= frequency of genotype *aa*, p_b^2= frequency of genotype *bb* and $2p_a p_b$= frequency of genotype *ab* (Hartl and Clark 1997).

Departures from Hardy-Weinberg equilibrium for each locus were tested both by the exact test (Probability-test or Fisher exact test) of Guo and Thompson (1992) and by tests assuming the specific alternative hypothesis of either heterozygote deficiency or heterozygote excess using the score (U) test of Raymond and Rousset (1995a) available in GENEPOP version 3.3 (Raymond and Rousset 1995b). Significance values (P-values) for the null hypothesis (H_0 = random union of gametes) were computed for each locus in each sample using the Markov chain method through 10,000 iterations (Guo and Thompson 1992) as implemented in GENEPOP. For the exact test (Probability-test), global tests across loci and samples were also computed in GENEPOP using a Markov chain (as described in Raymond and Rousset 1995a). All probability values were adjusted for multiple simultaneous table-wide tests using the sequential Bonferroni adjustment (Rice 1989).
Deviations from Hardy-Weinberg expectations may be due to a variety of causes. If an excess of heterozygotes is observed this may indicate the presence of overdominant selection (heterozygote advantage) or the occurrence of outbreeding. If a heterozygote deficiency is detected, it may be due to biological factors like inbreeding, population substructure (Wahlunds' effect), selection against heterozygotes, or artefactual factors such as stuttering (described in paragraph 3.7), null alleles, short allele dominance and linkage of loci.

Null alleles result from mutations in the primer binding site - leading to an allele that will not amplify during PCR and therefore to a false observation of homozygote excess (Estoup and Angers 1998). This may lead to incorrect inferences from genetic data, in particular in individual based analyses such as relatedness estimation and assignment tests (Hansen 2003).
Short allele dominance (or large allele dropout) is caused by preferential amplification of short alleles (Wattier et al. 1998), where the larger allele specifically fails to amplify.
Technical artefacts like null alleles and large allele dropout can be detected by testing whether particular loci mainly account for the deficit observed: Inbreeding and Wahlunds' effect should affect all loci, whereas a heterozygote deficit due to the presence of a null allele is caused only by particular loci.
Furthermore, as stuttering, short allele dominance and null alleles yield their own specific allelic "signature" (i.e. deficiencies and excesses of particular genotypes), it is possible to discriminate between deviations due to nonpanmixia and those caused by short allele dominance, null alleles and stuttering by applying the Windows®-based software package MICRO-CHECKER Version 2.2.1 (van Oosterhout et al. 2004).
MICRO-CHECKER was used to test for genotyping errors due to incorrect scoring of stutter peaks, short allele dominance and null alleles at each of the eight loci in each of the samples. The software is based on Fisher's combined probability test (see van Oosterhout et al. 2004 for details) and corrects probability values in all tests using the sequential Bonferroni method (Rice 1989).
Also a linkage of loci can cause heterozygote deficiencies: Different microsatellite loci may be located on the same chromosome. In this case, genotypes at one locus are not inherited independently from genotypes at the other locus, resulting in deviation of allele frequencies from Hardy-Weinberg expectations (linkage disequilibrium).
The null hypothesis of no linkage disequilibrium was tested in all samples using GENEPOP version 3.3 (Raymont and Rousset 1995b). Significance values were computed for each locus pair by unbiased estimates of Fisher's exact test using the Markov chain method (with 10,000 iterations) available in GENEPOP. Also a global test for each pair of loci was performed across samples.

3.8.2 Factorial correspondence analysis (FCA)

For a first exploration of the microsatellite data and to visualise the genetic relationships and patterns of genetic differentiation among the redfish individuals, a factorial correspondence analysis of individual multilocus scores (FCA; Benzécri 1973) was performed using GENETIX version 4.04. (Belkhir et al. 1996-2002). This technique was used to represent graphically the partitioning of genetic variance among samples and to project individuals into a multidimensional space on the basis of allelic information, with each allele being analysed as an independent variable.

The principle of FCA has been described in Benzécri (1973), but She et al. (1987) adopted the method to analyse genetic data (Lu et al. 2001). The method is based on the similarity of individuals in their allelic state for each allele. Crude genetic data are transformed into a contingency table (individuals x alleles), according to She et al. (1987): Each individual is described for each allele by the values 2, 1 and 0 according to whether it shows 2 (homozygote), 1 (heterozygote) or 0 copies of the considered allele. A multiple contingency table, which contains the whole crude information, is then obtained. Based on the frequencies of each allele in each individual in the contingency table, a cloud of *n* individuals located in a *k*-dimensional space (*k* = number of alleles) is associated to this table. To measure the relatedness between any two individuals in the *k*-dimensional space, the χ^2 distance centred on the marginal distribution of the contingency table (called Benzécri's χ^2) is used in correspondence analysis (Lu et al. 2001). Subsequently an inertia matrix is calculated. The FCA then identifies several independent factorial axes (or dimensions) that account for the largest part of the inertia (variation). For each dimension, each individual has an ordinate representing its position compared to the other individuals considered. Thus, it is possible to represent all individuals in the form of scatter plots. The main advantage of FCA is that each individual can be represented using each allele as an independent variable (Roques et al. 2001).
First, the FCA was calculated including genotype data of all individuals analysed (see Table 2.1) and all eight microsatellite loci.
As incomplete individual multilocus genotypes may create bias in the FCA projection (She et al. 1987), these individuals (70 out of 1595) were subsequently excluded from the analysis, to test whether this factor causes a distortion in the FCA projection.
As the specimens of "giant" *S. marinus* could not be screened for SEB30 (see paragraph 3.4) and were therefore not included in the first FCA, another FCA was performed without SEB30 and including the twelve "giant" *S. marinus* genotypes.

3.8.3 Model-based cluster analysis

The program *STRUCTURE* (Pritchard et al. 2000) Version 2.0 with the Windows® front-end (Pritchard and Wen 2003) was used to estimate the inter- and intraspecific genetic structure in the data set.
STRUCTURE is a Bayesian, Markov chain Monte Carlo-based approach that uses model-based clustering to identify the actual number of genetic clusters or gene pools (or populations) *K* in a given data set and to assign individuals probabilistically to these genetic clusters. In this way, it is possible to study hybrid zones, identify migrants and admixed individuals, and estimate population allele frequencies in situations where many individuals are migrants or admixed (Pritchard et al. 2000).

The algorithms and models are described in detail in Pritchard et al. (2000) and Falush et al. (2003). Therefore, only the basic principles will be described here:

Assuming a model in which there are *K* populations or genetic clusters (*K* may be unknown), each of which is characterised by a set of allele frequencies at each locus, *STRUCTURE* attempts to assign individuals probabilistically to these clusters on the basis of their multilocus genotypes, while simultaneously estimating population allele frequencies (Pritchard et al. 2000). The criteria for grouping individuals are to minimise Hardy-Weinberg disequilibrium and linkage disequilibrium between loci within groups, assuming that the neutral unlinked molecular markers are in Hardy-Weinberg equilibrium (HWE) and linkage equilibrium (LE) in the populations, and that recent population admixture, migration or hybridisation would probably produce departures from HWE and LE (Pierpaoli et al. 2003).

All calculations are conditional on a particular value of *K* (number of unobserved populations/genetic clusters/gene pools). Therefore, the analyses have to be performed for a range of values for *K* to find the most probable value. Each test performed in *STRUCTURE* yields a log-likelihood value of the data, $\ln\Pr(X \mid Z)$[1], the highest of which indicates, which test was closest to the actual number of genetically distinct populations or clusters (Pritchard and Wen 2003).

There are two different ancestry models available in *STRUCTURE*, the "no admixture model" (which assumes that each individual comes purely from one of the *K* populations) and the "admixture model" (which allows for the presence of admixed individuals in the sample). The latter model was chosen for the present analyses to study hybridisation between the North Atlantic *Sebastes* species and populations.

STRUCTURE assigns individuals probabilistically to one of the *K* clusters, or more than one cluster, if they are genetically admixed as a result of hybridisation, with estimates of their 90% confidence intervals. *STRUCTURE* estimates the membership coefficients (*Q*), representing an estimate of the amount of an individual's genome that is derived from one of the inferred parental populations. The tests also provide an alpha (α) value, a parameter for the degree of individual admixture. Large values of α (larger than 1) indicate that each individual has allele copies from all *K* populations in equal proportions (substantial admixture), whereas small values of α (smaller than 1) indicate that individuals originated mostly from a single population or gene pool (Pritchard et al. 2000).

Pritchard et al. (2000) recommend a number of guidelines and cautions for the choice of an appropriate value of *K*. Here, a visual inspection of estimated values of *Q* (the membership coefficient for each individual in each population) was performed in all cases, as described in Wilson et al. (2004). Where, for values of $K > 1$, estimates of *Q* showed a uniform distribution (with $Q \sim 1/K$) such that all individuals were deemed equally derived from each

[1] *X* denotes the genotypes of the sampled individuals, *Z* denotes the unknown clusters or populations of origin of the individuals

of the *K* populations, it was concluded that a value of $K = 1$ was the most appropriate choice of model for the data (Pritchard and Wen 2003; Wilson et al. 2004).
Beside the two different ancestry models, there are also two allele-frequency models available in *STRUCTURE*, differing in the way in which the allele frequencies are estimated.
The "independent allele model" assumes that allele frequencies in different populations are uncorrelated with one another (populations with very similar allele frequencies are not expected). This is a convenient approximation for populations that are not extremely closely related. This property has the result that the clustering algorithm may tend to merge subpopulations that share similar allele frequencies (Pritchard and Wen 2003).
An alternative is to permit allele frequencies to be correlated across populations (the "correlated allele frequency model" or "*F* model" by Falush et al. 2003), as allele frequencies in closely related populations are expected to be similar. This allows to extract much more information from the data and to perform accurate assignments of individuals in very closely related populations (Falush et al. 2003). One advantage of the "*F* model" is that it can sometimes detect population subdivision that is invisible when the gene frequencies of the populations are modelled without correlations (Falush et al. 2003; Pritchard and Wen 2003).
As recommended by Pritchard and Wen (2003), both models available in *STRUCTURE*, the model of uncorrelated allele frequencies and the model of correlated allele frequencies, were applied and the results were compared.
The *STRUCTURE* analysis of the data set was divided into three steps, (i) analysis of the species structure and identification of misclassified and admixed individuals, (ii) investigation of hybridisation and (iii) estimation of the number of populations within samples of *S. mentella* and *S. marinus*.

Species structure and identification of misclassified and admixed individuals:
First, it was tested whether *STRUCTURE* is able to identify the four species as genetically distinct, isolated groups. For this purpose, five samples, the *S. viviparus* sample (VI), the *S. fasciatus* sample (FA), the *S. marinus* samples from SW-Iceland 2001 (MASWIc01) and from the Faroe Islands West (MAWF02) as well as the *S. mentella* sample from the southern Irminger Sea (MESIrm01) were chosen, based on the results of the FCA and the tests for Hardy-Weinberg disequilibrium and null alleles (see paragraph 4.1 and 4.2). The number of clusters represented in these five samples was estimated by pooling these samples and calculating the posterior probability of the data, assuming that they originated from one to ten populations in the study area. Three independent simulations (to assess convergence) were run iteratively for *K* values (= number of gene pools) from 1 to 10, without prior population information, with a burnin period of 100,000, followed by 1,000,000 Markov chain Monte Carlo (MCMC) replications. The value of *K* showing the highest likelihood was chosen, and the individual membership coefficients (*Q*) of each individual were evaluated. Each sampled individual was considered as correctly assigned to one cluster if its individual proportion of

membership (*Q*) in this cluster was ≥ 0.80. Individuals were assigned to more than one cluster, if their individual proportion of membership in each cluster was < 0.80. In this case, the individuals were considered as unassigned and potentially admixed individuals.
The threshold value of 0.80, indicating that 80% of ancestry could be attributed to the respective genetic cluster, was chosen to be sure that at least 80% of an individual's genome was assigned to one or more than one inferred cluster, as described in Pierpaoli et al. (2003).
Subsequently, by using these five samples as so-called baseline samples, each individual within each of the remaining samples was assigned to these baseline samples in order to identify misclassified individuals and possible hybrids (individuals with admixed genotypes) in the data set. A five-cluster model was chosen based on the most likely number of clusters, as estimated by *STRUCTURE* (see paragraph 4.3.1).
As the sample sizes were varying, the assignment analysis was performed for each sample separately: The five baseline samples and the new, "unknown" sample were pooled, without prior information about the origin of the individuals. Simulations were run iteratively with K = 5 with a burnin period of 100,000, followed by 1,000,000 MCMC (Markov chain Monte Carlo) replications. The individual proportions of membership (*Q*) of each individual in the five inferred clusters were evaluated to identify misclassified individuals and individuals with admixed genotypes as described above (see also Pierpaoli at al. 2003 and Randi et al. 2003). Individuals were assigned to one cluster if their proportion of membership (*Q*) in that cluster was equal to or larger than 0.80. Individuals were treated as unassigned individual, when their proportion of membership in each cluster was < 0.80 (admixed individuals). If single individuals of a sample were assigned to another genetic cluster than the major part of the individuals of this sample, they were considered as misclassified specimens (pre-classified to the wrong species on board) and were removed from the sample in the subsequent population analyses using *STRUCTURE* (see below) as well as in subsequent estimation of genetic differentiation (AMOVA, F-statistics, genetic distances; see paragraph 3.8.4, 3.8.5, 3.8.6 and 3.8.7) - to reduce noise in the population genetic analysis (see Waples 1998).

Species hybridisation:
According to Hansen et al. (2001), it is advisable to include only samples relevant to a specific problem, such as interspecific hybridisation, when applying assignment methods. This strategy minimises statistical uncertainty and maximises assignment power. Therefore, after misclassified individuals were identified and the number of individuals with admixed genotypes was estimated, *STRUCTURE* analyses were performed with a reduced data set, to investigate

(i) hybridisation between *S. mentella*, *S. marinus* and *S. fasciatus* on the Flemish Cap (the only area covered in this study, where all three species occur sympatrically),

(ii) hybridisation between *S. marinus* and *S. mentella* on the shelves of the Faroe Islands, Iceland, East and West Greenland, which has been discussed by several authors (Kotthaus 1961; Altukhov and Nefyodov 1968; Nævdal 1978; Nedreaas and Nævdal 1991a; Nedreaas et al. 1994),

(iii) hybridisation in "giant" *S. marinus* (as proposed by Altukhov and Nefyodov 1968 "giant" *S. marinus* could be the product of hybridisation between *S. mentella* and *S. marinus*).

To investigate hybridisation between *S. mentella*, *S. marinus* and *S. fasciatus* on the Flemish Cap, samples of *S. mentella*, *S. marinus* and *S. fasciatus* collected on the Flemish Cap were pooled. *STRUCTURE* was run with a value of $K = 3$ to compute redfish ancestry. In this way, each individual redfish was forced to have its genotype assigned either to one of the three clusters, or, if admixed, to more than one cluster.

Hybridisation between *S. marinus* and *S. mentella* in the areas off the Faroe Islands, Iceland, East and West Greenland was investigated by pooling all samples of *S. mentella* and *S. marinus* collected in these areas. As in case of the Flemish Cap samples, *STRUCTURE* was run with a value of $K = 3$ (based on the number of clusters inferred by *STRUCTURE* in the previous analysis of species structure) to identify admixed individuals with intermediate genotypes.

In case of "giant" *S. marinus*, three of the baseline samples, the *S. marinus* samples from SW-Iceland 2001 (MASWIc01) and from the Faroe Islands West (MAWF02) and the *S. mentella* sample from the southern Irminger Sea (MESIrm01) were pooled with the twelve individuals of "giant" *S. marinus*. As locus SEB30 could not be screened for "giant" *S. marinus*, the locus was not used in this analysis. *STRUCTURE* was run with $K = 3$, as described above.

In all three analyses, simulations were run with a burnin period of 100,000, followed by 1,000,000 MCMC (Markov chain Monte Carlo) replications. The individual proportions of membership (Q) of each individual in the three pre-defined clusters were evaluated. In addition, triangle plots of Q created by *STRUCTURE* were examined visually.

Population structure within samples of *S. mentella* and *S. marinus*:

Misclassified and admixed individuals identified by *STRUCTURE* were excluded from the population analysis. The population structure within the samples of *S. marinus* and *S. mentella* was analysed for both species separately.

The number of populations represented in the samples was estimated by pooling all samples of one species (*S. marinus* and *S. mentella*, respectively). To choose an appropriate value of K for modelling the data, three independent runs at a range of values of K between 1 and 10 were run for each of the two sample sets, without prior information about the origin of the individuals. The probability $\ln\Pr(X|K)$ for $K = 1,2$ etc. was calculated. Simulations were run with a burnin period of 100,000, followed by 1,000,000 MCMC (Markov chain Monte Carlo)

replications. The individual proportions of membership (*Q*) of each individual in the putative subpopulations were estimated.

3.8.4 F-statistics and analysis of molecular variance (AMOVA)

Genetic differentiation between pairs of samples was quantified by estimates of Wright's (1951) pairwise F_{ST} (Weir and Cockerham 1984), which is based on an infinite allele model (IAM; Kimura and Crow 1964) of microsatellite mutation, and Slatkin's (1995) $R_{ST,}$ which is an F_{ST}-analogue assuming a stepwise mutation model (SMM; Kimura and Ohta 1978).

According to Weir and Cockerham (1984), F_{ST} may be interpreted as the variance in allele frequencies among populations, relative to the total variance:

$$F_{ST} = \frac{V_{between}}{V_{total}} = \frac{V_{between}}{V_{between} + V_{within}}$$

with V_{within} = variance in allele frequencies within populations, $V_{between}$ = variance in allele frequencies between populations, V_{total} = total variance in allele frequencies.

F_{ST} values range from 0 when two populations are panmictic to 1 when two populations are genetically completely isolated.
R_{ST} is analogous to the F_{ST} defined by Weir and Cockerham (1984), but takes into account the variance in allele sizes instead of the variance in allele frequencies (Slatkin 1995).
F_{ST} is similar to R_{ST}, when the differentiation is roughly independent of the mutation process and genetic drift is the driving force, that is, with large migration rates and/or recent times of divergence among populations. F_{ST} will, however, tend to underestimate the true level of genetic differentiation, when migration rates are low and/or when populations have been isolated for a longer period of time (Slatkin 1995; reviewed in Estoup and Angers 1998).
Slatkin (1995) has shown that R_{ST} performs better than F_{ST}, when microsatellites mutate in a stepwise fashion (reviewed in Estoup and Angers 1998 and Schlötterer 2000). However, as the mutational processes in microsatellites are still not fully understood (Jarne and Lagoda 1996; Goldstein and Pollock 1997), it is still unclear how departures from the mutation model would affect measure of genetic differentiation.
Therefore, it is advisable to compute and compare both statistics (Estoup and Angers 1998).
Genetic differentiation between pairs of samples was estimated as follows:
First, the degree of genetic differentiation between the four species was calculated by pooling samples of the same species - whereby species identity was based on morphological species

classification (see paragraph 3.1) - and calculating the pairwise F_{ST} and R_{ST} values between the four resulting species samples.
Subsequently, the pairwise F_{ST} and R_{ST} values between the genetic clusters inferred by the *STRUCTURE* analysis (see previous paragraph) were calculated, by pooling all individual assigned to the same cluster.
Then the pairwise F_{ST} and R_{ST} values between all pairs of samples of *S. mentella* and *S. marinus*, respectively, were calculated.
All pairwise F_{ST} and R_{ST} values were calculated using the software program Arlequin version 2.000 (Schneider et al. 2000).

The partition of genetic diversity within and among the North Atlantic *Sebastes* samples was analysed by a hierarchical analysis of molecular variance (AMOVA; Weir and Cockerham 1984; Excoffier et al. 1992) as implemented in Arlequin version 2.000 (Schneider et al. 2000). AMOVA estimates the percentages of variation for each locus among groups of samples, among samples within groups and within samples. Also the appropriate fixation indices (F-statistics) are estimated by AMOVA, namely the amount of genetic variation attributable to genetic differentiation among groups of samples (F_{CT}), among samples within groups (F_{SC}) and among samples relative to the total sample (global F_{ST}, see also above).
Four different AMOVAs were calculated. Each of these four AMOVAs was performed using first the variance in allele frequencies and subsequently the variance in allele sizes, respectively.
First, the degree of differentiation among species and among samples within species was estimated: Samples of the same species were grouped, whereby species identity was based on morphological species classification (see paragraph 3.1): All North Atlantic *Sebastes* samples were divided into four groups (one *S. viviparus* group, one *S. fasciatus* group, one *S. mentella* group and one *S. marinus* group). The proportions of the total genetic variation in allele frequencies and in allele sizes attributable to differences between species and between samples within species - as well as the proportion of genetic variation within samples - were estimated.
A second AMOVA was performed to quantify the degree of differentiation among the genetic clusters inferred by the *STRUCTURE* analysis (see previous paragraph) relative to the genetic differentiation among samples within these clusters. For this purpose, individuals originating from the same sample and assigned to the same cluster were grouped and the proportions of the total genetic variation in allele frequencies and allele sizes, respectively, that can be found among the clusters and among samples within each cluster were calculated.
A third and a fourth AMOVA were performed to quantify the degree of differentiation among temporal samples from the same location relative to the genetic differentiation among samples from different locations within each of the two species *S. mentella* and *S. marinus*.

Temporal samples of *S. mentella* and *S. marinus*, respectively, from the same location were grouped and the proportion of the total genetic variation due to temporal differentiation and the proportion of the total genetic variation due to spatial differentiation were calculated.
Locus-by-locus AMOVA was used to investigate, which loci contributed to the observed genetic structure.
Significance and resulting P-values of the AMOVA variance components, of the fixation indices and of the pairwise F_{ST} and R_{ST} values were tested in Arlequin version 2.000 (Schneider et al. 2000) using non-parametric permutation procedures as described in Excoffier et al. (1992) with 10,000 permutations.
Significance levels for simultaneous tests were adjusted by using the sequential Bonferroni approach (Rice 1989).

3.8.5 Isolation by distance

To test whether the observed genetic differences between samples of *S. mentella* and *S. marinus*, respectively, were correlated with geographic distance, isolation by distance analyses were performed. Isolation by distance (Wright 1943) should result in the increase of genetic differences with geographic distance, leading to a positive correlation among genetic and geographic distances.
Isolation by distance was examined by comparing the genetic distance as $[F_{ST}/(1-F_{ST})]$ (Rousset 1997) for each pair of samples with the geographic distance separating sampling locations - measured as shortest distance in km.
The geographic distances among sampling locations were obtained from a 1:1,000,000 scale map using the software MapViewer Version 6 (Golden Software Inc.). Significance in the isolation by distance relationship was tested using a Mantel test (Manly 1994) with a permutation procedure (10,000 iterations), as implemented in the software IBD version 2.0 (Bohonak 2002). IBD assesses whether the pairwise genetic distance matrix is correlated with the pairwise geographic distance matrix. A null distribution is generated by randomising rows and columns of one matrix while holding the other matrix constant (Bohonak 2002). IBD 2.0 calculates confidence intervals for slopes of IBD-relationships based on reduced major axis regression and bootstrapping over populations.

3.8.6 Genetic distances and multidimensional scaling (MDS)

To visualise the patterns of genetic differentiation between the species and the genetic relationships among the spatial and temporal samples of *S. mentella* and *S. marinus*, respectively, a multidimensional scaling analysis (MDSA) was performed in the program VISTA 5.6.3 (Young 1996), first for all samples and subsequently for the samples of each of the two species *S. mentella* and *S. marinus* separately, using matrices of Cavalli-Sforza and

Edwards (1967) chord distances (D_{CE}) between the samples. D_{CE} is not based on a model of evolution and assumes that divergence between populations is solely due to genetic drift. It focuses on the sum of the products of allele frequencies of those alleles shared between populations (or closely related species) and has been shown to reconstruct close genetic relationships better than stepwise mutation model (SMM) based distances (Takezaki and Nei 1996).

MDSA attempts to arrange the samples in a space with a particular number of dimensions (two-dimensional in this example) so as to reproduce the observed distances. The distances can be explained in terms of the underlying dimensions.

4. Results

4.1 Allelic composition, Hardy Weinberg and linkage equilibrium

A summary of the basic genetic data - number of individuals scored, number of alleles, number of "private" alleles (alleles only present in one sample), expected (H_e) and observed (H_o) heterozygosity, results of the exact tests for deviations from Hardy-Weinberg (HW) proportions and allelic richness (r(50)) - is given in Appendix III. Due to the small sample sizes, the genetic data of the three *S. viviparus* samples were pooled and also the two samples of *S. fasciatus* were pooled. As mentioned in paragraph 3.4, "giant" *S. marinus* could not be genotyped for locus SEB30, due to a lack of available DNA material.

In the complete data set, the total number of alleles per locus varied between 23 (locus SEB9) and 107 (locus SEB30) alleles. Private alleles were found in 20 out of 28 samples. The number of private alleles observed in the samples was low with between zero and five private alleles per sample. Five private alleles were observed in the *S. viviparus* (VI) sample and in the *S. mentella* sample from SE-Iceland 2001 (MESEIc01), in 18 of the 28 samples two or less private alleles were found (see Appendix III). In general, the frequencies of the private alleles (data not shown in Appendix III) detected were low (between 0.005 and 0.014), only in the *S. viviparus* (VI) sample three private alleles revealed slightly higher frequencies of 0.024, 0.037 and 0.049, respectively.

The observed heterozygosity was high with an average value of 0.81 and ranged from 0.44 (SEB46/MESEIc01) to 1.00 (SEB30/MAEGr01, SEB30/MEWGr01, SEB33/MAEF02, SEB33/MAEGr00b and SEB33/MAWGr01).

The high variability translated into high levels of gene diversity and allelic richness:

The expected heterozygosity (H_e) varied between 0.522 (SEB31/VI) and 0.964 (SEB33/MAEGr00), depending on locus and sample, with an expected mean heterozygosity within each sample ranging from 0.74 (MAG96) to 0.88 (MASEIc00, MASEIc01, MAEGr00, MECIrm01 and MECIrm03). Expected mean heterozygosity per locus varied between 0.79 (SEB9) and 0.95 (SEB33).

Allelic richness ranged from 3.99 (SEB31/VI) to 39.52 (SEB30/MASEIc01), with an average over loci between 8.78 (MASWIc01, MEWGr01) and 14.38 (MASEIc00).

The global tests (Probability-test; Guo and Thompson 1992) for departures from Hardy-Weinberg equilibrium revealed significant departures from expected distributions in 22 out of 28 samples (see Appendix III) after applying the sequential Bonferroni correction. These departures from Hardy-Weinberg equilibrium were caused by seven out of eight loci (SEB33 also showed significant departures, but these were not significant after applying the sequential Bonferroni correction). The results of the tests for deviation from expected Hardy-Weinberg proportions due to heterozygote deficiency and heterozygotes excess are given in Table 2.3 and 2.4, respectively. Table 2.3 shows that almost all significant departures from expected

HW proportions were due to heterozygote deficiencies. The global tests for Hardy-Weinberg proportions across loci revealed significant departures from expected distributions due to heterozygote deficiency in 20 out of 28 samples before applying the Bonferroni correction for multiple testing (Table 2.3). No heterozygote deficits were observed in the *S. marinus* samples from Norway, the Faroe Islands (East and West), SW-Iceland 2001, East Greenland 2000 (both samples), in the sample of "giant" *S. marinus* and in the *S. mentella* sample from the Flemish Cap. When the Bonferroni correction was applied, however, departures from expected distributions due to heterozygote deficiency remained significant in only 14 samples. Global tests for Hardy-Weinberg proportions due to heterozygote deficiency across samples showed that the heterozygote deficits were caused by almost all loci (except SEB33). The results of the tests for Hardy-Weinberg proportions in each sample and for each locus revealed a more detailed picture about the distribution of the heterozygote deficiencies among samples and loci (Table 2.3).

27 out of 28 samples showed significant heterozygote deficiencies caused by at least one locus. Only in the *S. marinus* sample from Southwest Iceland (MASWIc01) no significant departures from HW expectations were detected at any of the loci.

After Bonferroni correction for multiple testing, the departures from HW expectations due to heterozygote deficiency remained significant in only 15 samples (one significant heterozygote deficiency in the *S. viviparus* sample, three in the *S. marinus* samples and 11 in the *S. mentella* samples).

In general, within the *S. marinus* samples several different loci were causing the heterozygote deficits. The deficits lost statistical significance after Bonferroni adjustment for multiple tests in all except three cases (MASWIc00: SEB31 and SEB33; MAEGr01: SEB30).

Within the *S. mentella* samples mainly two particular loci, SEB37 and SEB46, were responsible for the heterozygote deficits. Some of the deficits caused by these two loci were still significant after applying the Bonferroni correction. SEB37 caused also heterozygote deficits in seven out of 15 *S. marinus* samples and in the *S. viviparus* sample, but except from the latter the deficits were not significant after Bonferroni adjustment.

However, the test for null alleles at each of the eight loci using MICRO-CHECKER (van Oosterhout et al. 2004) revealed that the heterozygotes deficits at locus SEB37 - observed in samples of both *S. marinus* and *S. mentella* and in addition at locus SEB46 in samples of *S. mentella* (see Table 2.3) - were in most cases caused by null alleles (average null allele frequency: 0.88). The *S. marinus* samples that did not reveal null alleles at locus SEB37 were MANO01, MAEF02, MAWF02, MANEIc00, MASWIc00, MASWIc01, MAFC01b and MAG96. The only *S. mentella* sample that did not show null alleles at locus SEB37 and SEB46 was MESIrm01.

As null alleles at relatively high frequency were observed at locus SEB37 in many samples of *S. marinus* and *S. mentella* and additionally at locus SEB46 in samples of *S. mentella*, SEB37 and SEB46 were subsequently either discarded completely from the subsequent analyses of

species structure, or analyses were performed with and without the two loci and the results were compared. SEB37 and SEB46 were not used in the analysis of population structure of *S. mentella*. When the population structure of *S. marinus* was analysed, only SEB37 was discarded, as SEB46 did not exhibit null alleles in samples of *S. marinus*.

Even though the global tests for departures from Hardy-Weinberg equilibrium across loci revealed no significant deviations from expected distributions due to heterozygote excess (Table 2.4), the HW test for each sample and for each locus revealed significant deviations due to heterozygote excess in four samples of *S. mentella* (MESEIc01, MEEGr01, MEWGr01 and MESIrm01), which did not remain significant after Bonferroni correction. In three of four cases the heterozygote excess was caused by locus SEB30 (this locus showed also the highest number of alleles), in one case the excess was caused by locus SEB45.

Exact tests for genotypic disequilibrium revealed no evidence of genetic linkage (data not shown). Only two out of 28 locus comparisons showed significant departures from equilibrium (SEB9 and SEB25: P = 0.007; SEB9 and SEB31: P = 0.005), which were not significant after Bonferroni correction ($\alpha = 0.05$) for multiple testing.

Table 2.3. Results of the test for deviation from expected Hardy Weinberg (HW) proportions. Probability values and level of significance for each locus and each sample when **H_1 = heterozygote deficiency**.

Probability values were adjusted for multiple tests/comparisons using the sequential Bonferroni corrections (Rice 1989): * P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001

Significant probability values before applying the sequential Bonferroni corrections are in bold type.

	SEB9	SEB25	SEB30	SEB31	SEB33	SEB37	SEB45	SEB46	Global
VI	0.0897	0.0676	**0.0371**	0.5068	0.5412	**0.0000*** **	**0.0146**	**0.0070**	**0.0000*** **
FA	0.0751	0.6574	**0.0012**	0.1525	0.2515	0.5309	0.7697	**0.0204**	**0.0703**
MANO01	**0.0043**	0.3815	0.2354	0.5550	0.6162	0.2866	0.0680	0.4710	0.1408
MAEF02	0.1158	0.3498	**0.0241**	**0.0247**	1.0000	0.1771	0.9467	0.0542	0.0750
MAWF02	0.2938	0.5704	**0.0033**	0.5247	0.6478	0.0578	0.1014	0.2076	0.1256
MANEIc00	0.1217	0.1058	**0.0031**	**0.0006**	0.2146	0.8314	0.0612	0.0597	**0.0000*** **
MASEIc00	0.0549	0.0529	0.0557	**0.0103**	0.9214	**0.0010**	**0.0092**	0.2810	**0.0000*** **
MASEIc01	0.2490	0.1649	0.0552	**0.0117**	0.4310	**0.0005**	0.3371	0.4818	**0.0000*** **
MASWIc00	0.1891	0.1334	0.2600	**0.0000*** **	**0.0000*** **	0.3444	0.3078	**0.0035**	**0.0000*** **
MASWIc01	0.7173	0.8340	0.1238	0.6925	0.0811	0.7666	0.2786	0.2956	0.1005
MAEGrDB01	**0.0009**	0.2275	**0.0345**	**0.0180**	0.6493	**0.0074**	0.0873	0.0985	**0.0000*** **
MAEGr00	**0.0361**	0.1041	0.1818	0.5048	0.8007	**0.0018**	0.3812	0.6215	0.1343
MAEGr00b	0.5646	**0.0235**	0.1332	0.8802	1.0000	**0.0155**	0.1274	0.7433	0.1025
MAEGr01	**0.0004**	0.4029	**0.0000*** **	0.4876	0.4355	**0.0003**	0.2535	0.8578	**0.0121**
MAWGr01	0.0574	0.4476	0.5006	0.4024	1.0000	**0.0019**	0.0823	0.2033	**0.0311**
MAFC01	0.9378	0.5287	0.1187	**0.0005**	0.1030	0.0987	0.2475	0.5078	**0.0057***
MAFC01b	**0.0252**	**0.0326**	**0.0090**	**0.0042**	**0.0024**	0.6072	**0.0205**	0.0690	**0.0000*** **
MAG96	0.2589	0.5348	-	0.1460	**0.0007**	0.7376	0.3331	0.5464	0.0540
MESEIc01	0.3619	0.6684	0.1486	**0.0046**	0.7013	0.5545	0.9827	**0.0000*** **	**0.0177**
MECIrm01	0.1494	0.6956	0.1555	0.4228	0.4696	**0.0000*** **	**0.0054**	**0.0000*** **	**0.0000*** **
MECIrm03	0.5531	0.1461	0.3054	0.1221	0.1290	**0.0000*** **	**0.0285**	**0.0005**	**0.0000*** **
MEEGr00	0.5163	0.1772	0.3288	0.8994	0.9538	**0.0149**	0.3893	**0.0000*** **	**0.0000*** **
MEEGr01	**0.0242**	**0.0406**	1.0000	**0.0091**	0.1518	**0.0000*** **	**0.0221**	**0.0000*** **	**0.0000*** **
MEWGr01	**0.0184**	0.1662	1.0000	0.6905	0.2854	**0.0000*** **	0.9263	**0.0024**	**0.0000*** **
MESIrm01	0.8687	0.7663	0.9958	0.3057	0.5591	0.1557	0.5739	**0.0000*** **	**0.0426**
MENAFO1F01	0.8985	0.0764	0.8489	0.3795	**0.0118**	**0.0353**	0.6532	**0.0000*** **	**0.0071**
MENAFO2J01	0.8181	0.2533	0.3280	0.3941	0.3815	**0.0069**	0.1130	**0.0000*** **	**0.0000*** **
MEFC01	0.1009	0.8505	0.3009	0.6950	0.7338	0.0544	0.5618	**0.0269**	0.1035
Global	**0.0000*** **	**0.0000*** **	**0.0000*** **	**0.0000*** **	0.2418	**0.0000*** **	**0.0058**	**0.0000*** **	

Table 2.4. Results of the test for deviation from expected Hardy Weinberg (HW) proportions. Probability values and level of significance for each locus and each sample when **H_1 = heterozygote excess**.
Probability values were adjusted for multiple tests/comparisons using the sequential Bonferroni corrections (Rice 1989): * P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001
Significant probability values before applying the sequential Bonferroni corrections are in bold type.

	SEB9	SEB25	SEB30	SEB31	SEB33	SEB37	SEB45	SEB46	Global
VI	0.9134	0.9361	0.9556	0.5864	0.4996	1.0000	0.9876	0.9928	1.0000
FA	0.9267	0.3441	0.9980	0.8823	0.7338	0.4697	0.2365	0.9792	0.9220
MANO01	0.9960	0.6084	0.7260	0.4555	0.4024	0.7266	0.9335	0.5260	0.9958
MAEF02	0.8817	0.6444	0.9794	0.9740	0.1224	0.8351	0.0553	0.9470	0.8580
MAWF02	0.7021	0.4304	0.9987	0.4933	0.3759	0.9457	0.9173	0.7856	0.9570
MANEIc00	0.8767	0.8925	0.9990	0.9994	0.7745	0.2255	0.9374	0.9442	1.0000
MASEIc00	0.9485	0.9472	0.9243	0.9907	0.0922	0.9998	0.9927	0.7271	1.0000
MASEIc01	0.7497	0.8419	0.9245	0.9880	0.6143	1.0000	0.6554	0.5162	1.0000
MASWIc00	0.8120	0.8687	0.7309	1.0000	1.0000	0.6567	0.7037	0.9971	1.0000
MASWIc01	0.3026	0.1628	0.8739	0.3137	0.9166	0.2682	0.7378	0.7031	0.4932
MAEGrDB01	0.9989	0.7748	0.9604	0.9800	0.3631	0.9939	0.9149	0.9011	1.0000
MAEGr00	0.9633	0.8931	0.8364	0.4987	0.2261	0.9969	0.6272	0.3853	0.9878
MAEGr00b	0.4486	0.9794	0.8898	0.1256	0.0938	0.9849	0.8834	0.2738	0.9378
MAEGr01	0.9996	0.6065	1.0000	0.5148	0.5582	0.9996	0.7308	0.1617	1.0000
MAWGr01	0.9447	0.5530	0.5419	0.6093	0.1730	0.9982	0.9182	0.8014	0.9926
MAFC01	0.0708	0.4679	0.8704	0.9995	0.8794	0.8997	0.7560	0.5020	1.0000
MAFC01b	0.9731	0.9672	0.9977	0.9961	0.9998	0.4385	0.9806	0.9303	1.0000
MAG96	0.8371	0.7340	-	0.8643	0.9998	0.6207	0.7661	0.6256	0.9485
MESEIc01	0.6329	0.3362	0.8512	0.9964	0.2977	0.5507	**0.0224**	1.0000	1.0000
MECIrm01	0.8572	0.3182	0.8364	0.5691	0.5095	1.0000	0.9948	1.0000	1.0000
MECIrm03	0.4472	0.8606	0.7265	0.8857	0.8916	1.0000	0.9678	0.9996	1.0000
MEEGr00	0.4864	0.8290	0.6687	0.1131	0.0726	0.9826	0.6102	1.0000	1.0000
MEEGr01	0.9783	0.9570	**0.0122**	0.9904	0.8544	1.0000	0.9789	1.0000	1.0000
MEWGr01	0.9814	0.8377	**0.0062**	0.3290	0.7431	1.0000	0.1104	0.9978	1.0000
MESIrm01	0.1545	0.2375	**0.0286**	0.6898	0.4380	0.8447	0.4460	1.0000	0.9576
MENAFO1F01	0.1141	0.9257	0.1774	0.6326	0.9839	0.9641	0.3498	1.0000	1.0000
MENAFO2J01	0.1927	0.7451	0.6752	0.6019	0.6273	0.9925	0.8901	1.0000	1.0000
MEFC01	0.8943	0.1518	0.6808	0.3129	0.2996	0.9515	0.4438	0.9737	0.9884
Global	1.0000	0.9955	1.0000	1.0000	0.7758	1.0000	1.0000	1.0000	

4.2 Factorial correspondence analysis (FCA)

Due to the high null allele frequencies observed at locus SEB37 and locus SEB46, the factorial correspondence analysis (FCA) was carried out (i) including and (ii) excluding these two loci. Figure 2.4 shows the results of the factorial correspondence analysis of microsatellite allelic composition based on 1513 individual redfish, including locus SEB37 and SEB46 (Figure 2.4; above) and excluding the two loci (Figure 2.4; below).
In both FCA plots, the 70 incomplete individual multilocus genotypes were excluded from the analysis, even though the inclusion of these individuals did not create distortion in the projection (corresponding FCA diagrams are not presented here). Also the FCA, in which locus SEB30 was excluded and the sample of "giant" *S. marinus* was included (plot not presented here), revealed a projection similar to the plots shown below.

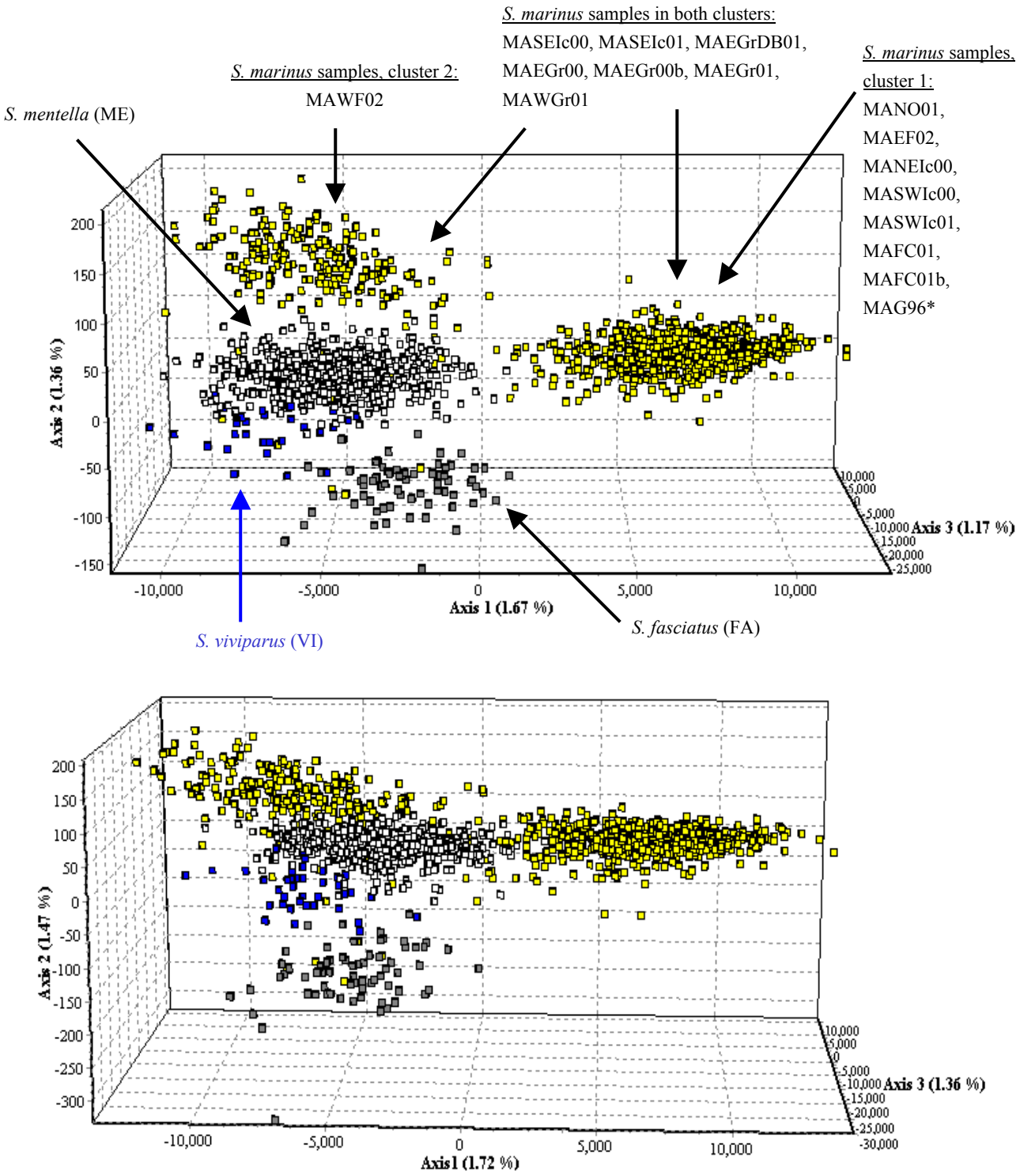


Figure 2.4. Diagrams of the factorial correspondence analyses (FCA; Benzécri 1973) showing relationships among the multilocus genotypes of 1513 individual redfish in a multidimensional space, including all eight loci (above) and excluding locus SEB37 and SEB46 (below). White squares: *S. mentella*; yellow squares: *S. marinus*; grey squares: *S. fasciatus*; blue squares: *S. viviparus*. Axis 1, axis 2 and axis 3 represent first, second and third principal factors of variability, respectively.

* "Giant" *S. marinus* (sample MAG96) were not included in these analyses due to their incomplete genotypes, but a separate analysis excluding locus SEB30 grouped these individuals with MANO01, MAEF02, MANEIc00, MASWIc00, MASWIc01, MAFC01 and MAFC01b.

Both FCA diagrams (Figure 2.4 above and below) revealed that the 1513 individual genotypes clustered into five groups, well separated by the three axes and with minimal overlap between the clusters. The clustering in the second FCA plot (without locus SEB37 and SEB46) was slightly less distinct as in the plot, in which all loci were included. In the first FCA plot, including all loci, the first axis accounted for 1.67% of the total variance, the second axis for 1.36% and the third axis for 1.17% of the total genetic variance. In the second FCA plot (without SEB37 and SEB46), the first axis accounted for 1.72% of the total variance, the second axis for 1.47% and the third axis for 1.36% of the total genetic variance.
In both FCA plots, the samples of the three species *S. viviparus* (VI), *S. mentella* (ME) and *S. fasciatus* (FA) each formed a single cluster. The samples of *S. marinus* showed a clear distinction into two distinct genetic clusters, which grouped separately (cluster 1 and cluster 2; see Figure 2.4): One cluster (cluster 1) consisted mainly of individuals from Norway (MANO01), Faroe Islands East (MAEF02), NE-Iceland (MANEIc00), SW-Iceland (MASWIc00 and MASWIc01) and the Flemish Cap (MAFC01and MAFC01b). The FCA, in which locus SEB30 was excluded and the "giant" *S. marinus* genotypes were included, revealed that also individuals of "giant" *S. marinus* (MAG96) were grouped in this cluster (plot not shown).
Individuals from the Faroe Islands West (MAWF02), SE-Iceland (MASEIc00 and MASEIc01), East Greenland/Dohrn Bank (MAEGrDB01), East Greenland (MAEGr00, MAEGr00b and MAEGr01) and West Greenland (MAWGr01) were composing the second cluster (cluster 2), but individual genotypes of the first group were also found within these samples. Only the Faroe Islands West sample consisted mainly of individuals of cluster 2, whereas the samples from SE-Iceland, East Greenland/Dohrn Bank, East Greenland and West Greenland represented an admixture of individuals of both clusters in different proportions.
Furthermore, the FCA placed several single individuals of the samples of *S. mentella*, *S. fasciatus* and *S. marinus* into another genetic cluster than the major part of the individuals of these samples (see Figure 2.4): Several individuals of the *S. fasciatus* sample were placed into the first *S. marinus* cluster (cluster 1) and a few individuals of the *S. marinus* sample from the Flemish Cap grouped with the individuals of the *S. fasciatus* sample. A few individuals of the *S. mentella* sample from East Greenland were placed into the second *S. marinus* cluster (cluster 2) and few individuals of the *S. mentella* sample from SE-Iceland (MESEIc01) grouped with the first *S. marinus* cluster (cluster 1).

4.3 Results of the model-based cluster analyses

4.3.1 Species structure and species assignment

Five samples representing the four North Atlantic *Sebastes* species were pooled to investigate the structuring of the data on the species level: The *S. viviparus* sample (VI), the *S. fasciatus*

sample (FA), the *S. marinus* samples from SW-Iceland (MASWIc01) and from the Faroe Islands West (MAWF02), as well as the *S. mentella* sample from the southern Irminger Sea (MESIrm01). The *S. mentella* sample from the southern Irminger Sea was chosen, as it was the only *S. mentella* sample that did not reveal null alleles at locus SEB37 and SEB46 (null alleles can cause an overestimation of *K* in the *STRUCTURE* analysis; Pritchard et al. 2000; Pritchard and Wen 2003). Based on the results of the FCA, which revealed five genetic clusters in the total data set with two distinct clusters in the *S. marinus* samples, two samples of *S. marinus* were selected: The samples MASWIc01 and MAWF02 were chosen, as each of them consisted mainly of individuals of one of these two clusters (MASWIc01 consisted mainly of individuals of cluster 1 and MAWF02 consisted mainly of individuals of cluster 2; see previous paragraph). These two samples did not reveal null alleles at any of the loci analysed.

The results of the *STRUCTURE* analysis revealed that the most likely number of genetic clusters/gene pools in the data set ($n = 262$) was five ($K = 5$) in each of the three independent runs (each run with a *K* ranging from 1 to 10) and for both models: The model assuming independent allele frequencies and the model assuming correlated allele frequencies. This result was also obtained when the analysis was performed without locus SEB37 and SEB46. Table 2.5 shows the results of the *STRUCTURE* analysis without locus SEB37 and SEB46 for both allele frequency models. The highest probability of the data was found with clusters set at 5 (allele frequencies independent: $\ln\Pr(X|K) = -7281.9$, $\alpha = 0.036$; allele frequencies correlated: $\ln\Pr(X|K) = -7308.8$, $\alpha = 0.033$). The alpha values were generally low.

The average proportions of membership of each sample in the five inferred clusters, as well as the number of individuals assigned to each cluster with a proportion of membership $Q \geq 0.80$ and the number of unassigned individuals (assigned to more than one cluster with a proportion of membership $Q < 0.80$), are reported in Table 2.6. The corresponding *Q*-plot is shown in Figure 2.5. Table 2.6 shows that most of the individuals (97.71%) were assigned to one of the five clusters with an admixture coefficient *Q* equal to or larger than 0.80. These individuals were considered as correctly assigned to one cluster. Only six individuals were not assigned to any single cluster with *Q* equal to or larger than 0.80, but were split into two or more clusters.

All individuals of the *S. viviparus* sample except one were assigned to cluster IV ($Q_{IV} = 0.951$). Individuals of the *S. fasciatus* sample were assigned mainly to cluster I ($Q_I = 0.923$), except one unassigned individual and four individuals, which were assigned to cluster II (two individuals) and V (two individuals), respectively. Most individuals (32 out of 39) of the *S. marinus* sample from the Faroe Islands West (MAWF02) were grouped into a separate cluster, cluster III ($Q_{III} = 0.794$). Table 2.6 shows that five individuals of this sample were assigned to cluster II, the cluster in which most individuals of *S. mentella* from the southern Irminger Sea (MESIrm01) were placed (see below), and two to cluster V - the cluster in which *S. marinus* from SW-Iceland (MASWIc01) were placed. All individuals of the *S.*

marinus sample from SW-Iceland 2001 (MASWIc01) were assigned to cluster V (Q_V = 0.962), except one unassigned individual. The individuals of the *S. mentella* sample from the southern Irminger Sea (MESIrm01) were assigned to cluster II (Q_{II} = 0.910), except three unassigned individuals and one individual that was placed in cluster III, together with most of the individuals of sample MAWF02.

Due to the clear grouping of the five samples into the five inferred clusters, these clusters were subsequently designated as follows: Cluster I was designated as "*S. fasciatus* cluster", cluster II as "*S. mentella* cluster", cluster IV was designated as "*S. viviparus* cluster" and cluster V as "*S. marinus* cluster". As most of the individuals of the *S. marinus* sample MAWF02 were grouped into cluster III, separately from MASWIc01, this cluster was named "*Sebastes sp.* cluster".

Table 2.7 shows the individual proportions of membership (Q) of each of the six-unassigned/admixed individuals in the five inferred clusters. The single unassigned individuals found in the *S. viviparus* sample, in the *S. fasciatus* sample and in the *S. marinus* sample MASWIc01, as well as two of the three unassigned individuals found in the *S. mentella* sample MESIrm01, still revealed a relatively high proportion of membership (Q) of between 0.717 and 0.771 in one of the five clusters (*S. viviparus*: cluster IV, *S. fasciatus*: cluster I, *S. marinus*: cluster V, *S. mentella*: cluster II). Only one unassigned individual of MESIrm01 showed similar proportions of membership (Q) in cluster II and V. In all six cases, the 90% confidence intervals were large.

Table 2.5. Results of the Bayesian clustering analyses performed using *STRUCTURE* (Pritchard et al. 2000) with the pooled samples VI, FA, MASWIc01, MAWF02 and MESIrm01 - using both allele frequency models, without using prior information of species origin (262 individuals, six microsatellite loci, without SEB37 and SEB46). The natural logarithm of the probability of the data lnPr($X|K$) and its variance for each K (= number of clusters) are given. Alpha (α) values indicate the admixture value. Lowest K, lnPr($X|K$) and α values are in bold type. The results were similar when locus SEB37 and locus SEB46 were included in the analysis.

	Allele frequencies independent			Allele frequencies correlated		
K	Estimated ln probability lnPr($X\|K$)	Variance lnPr($X\|K$)	Mean value of α	Estimated ln probability lnPr($X\|K$)	Variance lnPr($X\|K$)	Mean value of α
1	-9074.2	94.8	-	-9077.9	97.9	-
2	-8400.2	229.3	0.046	-8277.2	240.8	0.038
3	-7884.3	332.1	0.039	-7981.4	369.1	0.036
4	-7515.6	410.3	0.038	-7570.3	430.0	0.034
5	**-7281.9**	**492.1**	**0.036**	**-7308.8**	**504.6**	**0.033**
6	-7298.2	648.5	0.037	-7571.9	1100.5	0.033
7	-7298.8	769.0	0.038	-7801.9	1640.9	0.034
8	-7302.8	898.5	0.038	-7793.3	1626.5	0.034
9	-7408.3	1203.3	0.039	-7871.3	1793.6	0.034
10	-7400.1	1288.5	0.040	-7885.1	1787.4	0.034

Table 2.6. Results of the Bayesian clustering analysis performed using *STRUCTURE* (Pritchard et al. 2000) with the samples VI, FA, MASWIc01, MAWF02 and MESIrm01 - using the model of correlated allele frequencies, without using prior information of species origin (262 individuals, six microsatellite loci, without SEB37 and SEB46). The posterior probability of the number of genetic clusters in the sample set was maximum with $K = 5$. The table shows the average proportion of membership (Q) of each sample in each of the five inferred clusters (I, II, III, IV and V). Each individual was assigned to a single cluster if its Q_i (i = I–V) in that cluster was equal to or larger than 0.80. The numbers of individuals assigned to each cluster with a proportion of membership $Q \geq 0.80$ are indicated in parentheses. "Unassigned" individuals were assigned to more than one cluster, with a proportion of membership $Q < 0.80$.

		Inferred clusters					Number of unassigned	Percentage of assigned
Sample	*n*	I	II	III	IV	V	individuals	individuals
VI	42	0.011	0.018	0.014	0.951	0.006		
		(0)	(0)	(0)	(41)	(0)	1	97.62
FA	81	0.923	0.031	0.006	0.012	0.029		
		(76)	(2)	(0)	(0)	(2)	1	98.77
MAWF01	39	0.008	0.125	0.794	0.006	0.066		
		(0)	(5)	(32)	(0)	(2)	0	100.00
MASWIc01	52	0.008	0.010	0.012	0.008	0.962		
		(0)	(0)	(0)	(0)	(51)	1	98.08
MESIrm01	48	0.008	0.910	0.037	0.021	0.024		
		(0)	(44)	(1)	(0)	(0)	3	91.67
Individuals total	262	(76)	(51)	(33)	(41)	(55)	6	97.71

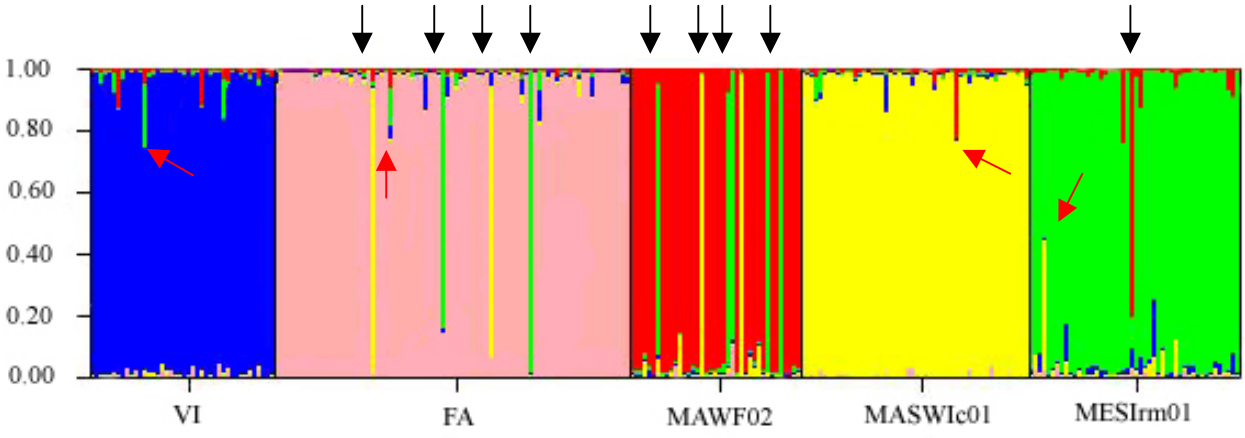


Figure 2.5. Results of the Bayesian clustering analysis performed using *STRUCTURE* (Pritchard et al. 2000) with the samples VI, FA, MASWIc01, MAWF02 and MESIrm01 - using the model of correlated allele frequencies, without using prior information of species origin (262 individuals, six microsatellite loci, without SEB37 and SEB46).
Plot of Q (estimated membership coefficient for each individual in each cluster). Each individual in the data set is represented by a single vertical line, which is partitioned into $K = 5$ (population number) coloured segments that represent the individuals' estimated membership fraction in each of the K inferred clusters. Black arrows: Individuals assigned to other genetic clusters with $Q \geq 0.80$. Red arrows: Examples for unassigned individuals.

Table 2.7. Inferred ancestry of the six unassigned (admixed) individuals in the samples VI, FA, MASWIc01, MAWF02 and MESIrm01, estimated using *STRUCTURE* (Pritchard et al. 2000) without locus SEB37 and SEB46. The posterior probability of the number of genetic clusters in the sample set was maximum with $K = 5$. The table shows the proportions of membership (Q) of each unassigned individual in each of the five inferred clusters (I, II, III, IV and V). 90% confidence intervals are given in parentheses.

	Unassigned individual	Inferred cluster				
		I	II	III	IV	V
Sample	no.	*S. fasciatus*	*S. mentella*	*Sebastes sp.*	*S. viviparus*	*S. marinus*
VI	1	0.013 (0.000-0.087)	0.211 (0.000-0.646)	0.048 (0.000-0.305)	0.717 (0.322-1.000)	0.011 (0.000-0.079)
FA	1	0.771 (0.451-1.000)	0.121 (0.000-0.453)	0.057 (0.000-0.306)	0.043 (0.000-0.319)	0.008 (0.000-0.044)
MASWIc01	1	0.003 (0.000-0.011)	0.007 (0.000-0.024)	0.277 (0.000-0.551)	0.006 (0.000-0.022)	0.707 (0.437-1.000)
MESIrm01	1	0.026 (0.000-0.195)	0.534 (0.000-1.000)	0.012 (0.000-0.068)	0.004 (0.000-0.016)	0.424 (0.000-1.000)
	2	0.004 (0.000-0.020)	0.730 (0.279-1.000)	0.246 (0.000-0.704)	0.016 (0.000-0.124)	0.004 (0.000-0.017)
	3	0.004 (0.000-0.016)	0.745 (0.064-1.000)	0.005 (0.000-0.027)	0.184 (0.000-0.629)	0.062 (0.000-0.381)

Subsequently, the five samples VI, FA, MASWIc01, MAWF02 and MESIrm01 were used as "baseline samples" as described in paragraph 3.8.3. In this way, each individual of each of the remaining samples of *S. marinus* and *S. mentella* (see Table 2.1) was assigned to one - or more than one - of the five genetic clusters inferred by *STRUCTURE*.

Table 2.8 shows, how many individuals per sample were assigned to each of the five clusters.

The global performance of *STRUCTURE* allowed the assignment of 93.54% of the individuals, with 1492 out of 1595 individuals revealing values of $Q \geq 0.80$ in one of the five clusters. In total, 103 individuals (6.46%) could not be assigned unambiguously, as they revealed values of Q less than 0.80 in each of the five clusters - and therefore admixed genotypes with a partial membership in several clusters.

In the total *S. marinus* data set, 94.97% of the individuals were assigned to one of the five clusters with a $Q \geq 0.80$. Of the 914 individuals sampled as *S. marinus*, 639 (69.9%) were assigned to cluster V (the "*S. marinus* cluster"). Also all individuals of "giant" *S. marinus* were placed in cluster V. Four individuals of *S. marinus* - all collected on the Flemish Cap - were assigned to the "*S. fasciatus* cluster" (cluster I), 13 individuals were assigned to cluster II, the "*S. mentella* cluster", one individual grouped with the "*S. viviparus* cluster" (cluster IV).

211 individuals (23.10%) sampled as *S. marinus* clustered separately into another cluster - cluster III (the "*Sebastes sp.* cluster"). These individuals were mainly found in the *S. marinus* samples from the Faroe Islands West, SE-Iceland and Greenland. A few of these individuals

also appeared in the samples from NE-Iceland and the Faroe Islands East. All these samples contained a mixture of individuals from cluster III and V in different proportions. Individuals assigned to cluster III were not found in the samples from Norway, SW-Iceland (collected in the years 2000 and 2001) and the Flemish Cap.

46 of the individuals sampled as *S. marinus* (5.03%) were assigned to more than one cluster, with a proportion of membership $Q < 0.80$.

In the total *S. mentella* data set, 90.14% of the individuals sampled as *S. mentella* were assigned to one of the five clusters with a Q equal to or larger than 0.80. Most individuals (492 out of 558; about 88.2%) were assigned to cluster II, the "*S. mentella* cluster". Nine individuals (eight collected off East Greenland in the year 2001) were assigned to cluster III, the "*Sebastes sp.* cluster", two were assigned to the "*S. marinus* cluster" (cluster V). 55 individuals (9.86%) revealed admixed genotypes.

In Appendix IV all unassigned individuals and their proportions of membership (Q) in each of five inferred clusters are listed. These individuals showed varying individual admixture proportions and were assigned in most cases to two clusters, but in some cases also to more than two clusters. Admixed genotypes were composed of almost all combinations of the five clusters. The unassigned individuals sampled as *S. marinus* revealed mainly combinations of cluster V ("*S. marinus* cluster"), cluster III ("*Sebastes sp.* cluster") and cluster II ("*S. mentella* cluster"). In some cases also cluster I ("*S. fasciatus* cluster") and cluster IV ("*S. viviparus* cluster") were involved. The unassigned individuals sampled as *S. mentella* were composed of cluster II ("*S. mentella* cluster") and the other four clusters in differing combinations and proportions. In almost all cases, the individual 90% confidence intervals of Q were large (data not shown).

Table 2.8. Results of the assignment analyses performed in *STRUCTURE* (Pritchard et al. 2000) with all samples of North Atlantic *Sebastes* (see Table 2.1). The "baseline samples" (Table 2.6) were used as reference samples, applying the model of correlated allele frequencies, without using prior information of species origin. The baseline samples are also included in the table. The table shows the average proportion of membership (Q) of each sample in each of the five inferred clusters. Each individual was assigned to a single cluster if its Q_i (i = I–V) in that cluster was equal to or larger than 0.80. The numbers of individuals assigned to each cluster with a proportion of membership $Q \geq 0.80$ are indicated in parentheses. "Unassigned" individuals were assigned to more than one cluster, with a proportion of membership $Q < 0.80$. In case of "giant" *S. marinus*, locus SEB30 was not used in the assignment analysis, as these individuals were not screened for this locus.

		Inferred cluster					Number of	Percentage
		I	II	III	IV	V	unassigned	of assigned
Sample	*n*	*S. fasciatus*	*S. mentella*	*Sebastes sp.*	*S. viviparus*	*S. marinus*	individuals	individuals
VI	42	0.011	0.018	0.014	0.951	0.006	1	97.62
		(0)	(0)	(0)	(41)	(0)		
FA	81	0.923	0.031	0.006	0.012	0.029	1	98.77
		(76)	(2)	(0)	(0)	(2)		
MANO01	78	0.009	0.021	0.013	0.008	0.949	3	96.15
		(0)	(0)	(0)	(0)	(75)		
MAEF02	59	0.011	0.032	0.041	0.009	0.907	1	98.30
		(0)	(1)	(2)	(0)	(55)		
MAWF02	39	0.008	0.125	0.794	0.006	0.066	0	100.00
		(0)	(5)	(32)	(0)	(2)		
MANEIc00	52	0.014	0.091	0.077	0.012	0.806	8	84.62
		(0)	(2)	(3)	(0)	(39)		
MASEIc00	101	0.008	0.017	0.473	0.007	0.495	2	98.02
		(0)	(0)	(49)	(0)	(50)		
MASEIc01	55	0.008	0.031	0.521	0.015	0.425	4	92.73
		(0)	(1)	(27)	(0)	(23)		
MASWIc00	60	0.013	0.012	0.010	0.037	0.928	2	96.67
		(0)	(0)	(0)	(1)	(57)		
MASWIc01	52	0.008	0.010	0.012	0.008	0.962	1	98.08
		(0)	(0)	(0)	(0)	(51)		
MAEGrDB01	50	0.010	0.057	0.213	0.038	0.681	8	84.00
		(0)	(1)	(7)	(0)	(34)		
MAEGr00	71	0.006	0.026	0.348	0.008	0.612	3	95.78
		(0)	(1)	(23)	(0)	(44)		
MAEGr00b	47	0.012	0.016	0.126	0.008	0.837	2	95.75
		(0)	(0)	(5)	(0)	(40)		
MAEGr01	60	0.008	0.048	0.763	0.008	0.173	3	95.00
		(0)	(2)	(46)	(0)	(9)		
MAWGr01	50	0.008	0.012	0.334	0.007	0.639	0	100.00
		(0)	(0)	(17)	(0)	(33)		
MAFC01	55	0.020	0.016	0.012	0.011	0.942	3	94.55
		(0)	(0)	(0)	(0)	(52)		
MAFC01b	73	0.071	0.022	0.010	0.014	0.883	6	91.78
		(4)	(0)	(0)	(0)	(63)		
MAG96	12	0.006	0.023	0.026	0.005	0.94	0	100.00
		(0)	(0)	(0)	(0)	(12)		
S. marinus total	914	(4)	(13)	(211)	(1)	(639)	46	94.97

Table 2.8 (continued). Results of the assignment analyses performed in *STRUCTURE* (Pritchard et al. 2000) with all samples of North Atlantic *Sebastes* (see Table 2.1). The "baseline samples" (Table 2.6) were used as reference samples, applying the model of correlated allele frequencies, without using prior information of species origin. The baseline samples are also included in the table. The table shows the average proportion of membership (Q) of each sample in each of the five inferred clusters. Each individual was assigned to a single cluster if its Q_i (i = I–V) in that cluster was equal to or larger than 0.80. The numbers of individuals assigned to each cluster with a proportion of membership $Q \geq 0.80$ are indicated in parentheses. "Unassigned" individuals were assigned to more than one cluster, with a proportion of membership $Q < 0.80$. In case of "giant" *S. marinus*, locus SEB30 was not used in the assignment analysis, as these individuals were not screened for this locus.

		Inferred cluster					Number of	Percentage
		I	II	III	IV	V	unassigned	of assigned
Sample	*n*	*S. fasciatus*	*S. mentella*	*Sebastes sp.*	*S. viviparus*	*S. marinus*	individuals	individuals
MESEIc01	51	0.013	0.899	0.013	0.027	0.050	1	98.04
		(0)	(48)	(0)	(0)	(2)		
MECIrm01	57	0.019	0.882	0.028	0.042	0.030	10	82.46
		(0)	(47)	(0)	(0)	(0)		
MECIrm03	66	0.016	0.895	0.037	0.025	0.027	10	84.85
		(0)	(56)	(0)	(0)	(0)		
MEEGr00	68	0.011	0.912	0.030	0.022	0.025	7	91.05
		(0)	(61)	(0)	(0)	(0)		
MEEGr01	58	0.014	0.758	0.155	0.045	0.028	10	82.76
		(0)	(40)	(8)	(0)	(0)		
MEWGr01	51	0.011	0.947	0.010	0.014	0.019	1	98.04
		(0)	(50)	(0)	(0)	(0)		
MESIrm01	48	0.008	0.910	0.037	0.021	0.024	3	91.67
		(0)	(44)	(1)	(0)	(0)		
MENAFO1F01	48	0.011	0.925	0.027	0.018	0.019	2	95.83
		(0)	(46)	(0)	(0)	(0)		
MENAFO2J01	50	0.009	0.918	0.013	0.028	0.013	5	90.00
		(0)	(45)	(0)	(0)	(0)		
MEFC01	61	0.034	0.901	0.016	0.021	0.028	6	90.16
		(0)	(55)	(0)	(0)	(0)		
S. mentella total	558	(0)	(492)	(9)	(0)	(2)	55	90.14
Total	**1595**	**(80)**	**(507)**	**(220)**	**(42)**	**(643)**	**103**	**93.54**

4.3.2 Species hybridisation

To reduce genetic noise in the subsequent analyses, the following individuals were removed from the data set: The single individual assigned to the "*S. viviparus* cluster" was removed from the *S. marinus* sample MASWIc00. Also the two individuals of the *S. fasciatus* sample that were assigned to the "*S. mentella* cluster" and the "*S. marinus* cluster", respectively, and the four individuals of the *S. marinus* sample from the Flemish Cap that were assigned to the "*S. fasciatus* cluster" were removed (see paragraph 4.3.1).

Test for hybridisation between *S. mentella* and *S. marinus* in the area Faroe Islands/ Iceland/ Greenland

To test if there is hybridisation between *S. marinus* and *S. mentella* in the areas off the Faroe Islands, Iceland, East and West Greenland, all individuals of *S. marinus* and *S. mentella* collected in these areas (n = 923; samples MAEF02, MAWF02, MANEIc00, MASEIc00, MASEIc01, MASWIc00, MASWIc01, MAEGrDB01, MAEGr00, MAEGr00b, MAEGr01, MAWGr01, MESEIc01, MEEGr00, MEEGr01 and MEWGr01) were pooled.

STRUCTURE was run with K = 3 (one cluster representing *S. mentella* and two clusters representing *S. marinus* - according to the results of the previous analyses of species structure; see paragraph 4.2 and 4.3.1), without the data from locus SEB37 and SEB46, assuming correlated allele frequencies.

Out of 923 individuals, 885 (95.88%) were assigned to one of the three clusters with a proportion of membership $Q \geq 0.80$. 38 individuals (4.12%) showed admixed genotypes and were assigned to more than one cluster with a proportion of membership $Q < 0.80$. In Figure 2.6 the corresponding triangle plot of Q is given. The triangle plot shows that most of the individuals were placed in one of the corners of the triangle and were therefore assigned completely to one of the three predefined clusters. Almost all individuals were grouped into the "*S. mentella* cluster", the "*S. marinus* cluster" or the "*Sebastes sp.* cluster", respectively, in agreement with the results of the assignment tests (see paragraph 4.3.1) - except the few unassigned/admixed individuals mentioned above. The individuals with admixed genotypes were placed somewhat outside the three predefined clusters, representing in most cases genotypes combined out of two clusters with varying individual admixture proportions. Figure 2.6 shows that most of these individuals revealed a higher proportion of membership in one cluster. Only few individuals revealed intermediate genotypes with similar proportions of membership (Q) in two clusters.

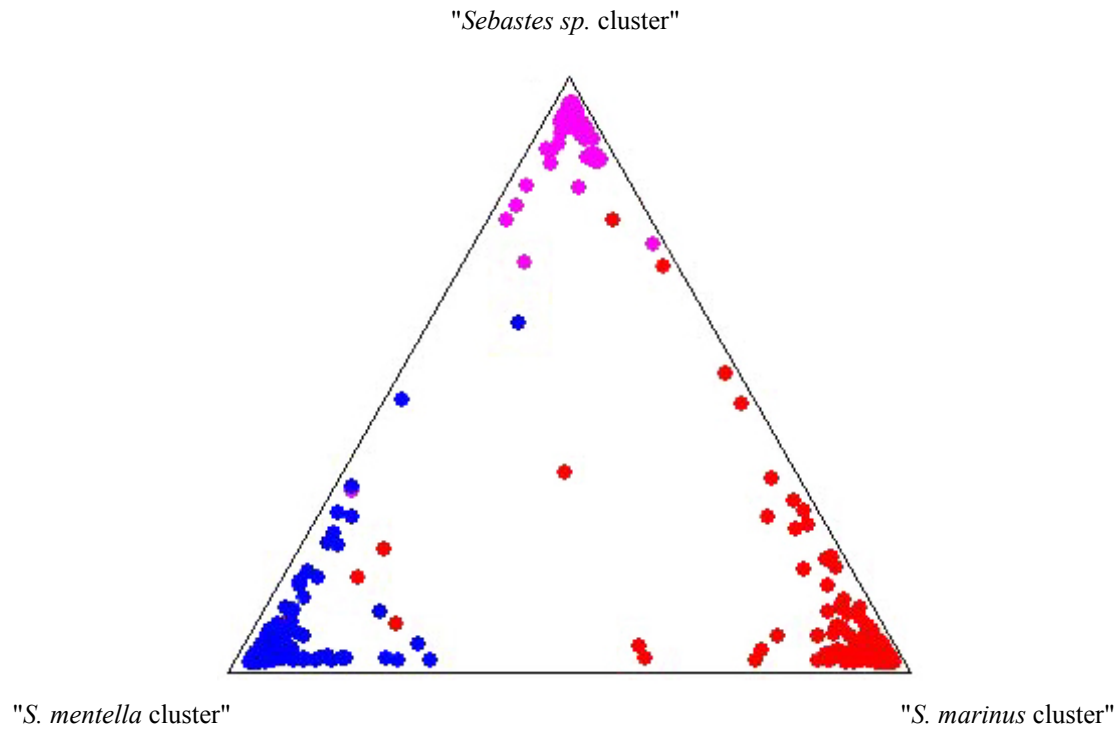


Figure 2.6. Triangle plot of *Q* representing a summary of the clustering results obtained using *STRUCTURE* for the *S. mentella* and *S. marinus* samples from the Faroe Islands, Iceland, East and West Greenland (n = 923, six loci, without SEB37 and SEB46, correlated allele frequency model) assuming three genetic clusters. Each individual is represented by a coloured point. The estimated ancestry vector for an individual consists of K components, which add up to 1. When K = 3, the ancestry vectors can be plotted onto a triangle, as shown. For a given point, each of the three components is given by the distance to one edge of the triangle. Individuals who are in one of the corners are therefore assigned completely to one cluster or another (Pritchard et al. 2000). After the clustering was performed, the points were labelled according to the results of the previous individual assignment analyses presented in paragraph 4.3.1: Blue: individuals previously assigned to the "*S. mentella* cluster"; red: individuals previously assigned to the "*S. marinus* cluster"; pink: individuals previously assigned to the "*Sebastes sp.* cluster".

Test for hybridisation between *S. mentella*, *S. marinus* and *S. fasciatus* on the Flemish Cap

The *STRUCTURE* analysis of the pooled samples of *S. mentella*, *S. marinus* and *S. fasciatus* from the Flemish Cap (262 individuals; sample MEFC01, MAFC01, MAFC01b and FA, without data from locus SEB37 and SEB46 - assuming correlated allele frequencies) assuming three genetic clusters (corresponding to the species) revealed that 95.42% of the individuals (250 individuals) could be assigned to one of the three clusters with an admixture coefficient *Q* equal to or larger than 0.80. Only 4.58% (12 individuals) revealed admixed genotypes. Also in the case of *S. mentella*, *S. marinus* and *S. fasciatus* from the Flemish Cap, the triangle plot of *Q* (Figure 2.7) shows that most of the individuals of the three species

samples were placed in one of the corners of the triangle and were therefore assigned completely to one of the three clusters.

The few unassigned individuals with admixed genotypes falling outside these clusters were genetically mainly composed of the "*S. mentella* cluster" and the "*S. marinus* cluster". Only few unassigned individuals with admixed genotypes showed a small proportion of membership in the "*S. fasciatus* cluster" (see Figure 2.7).

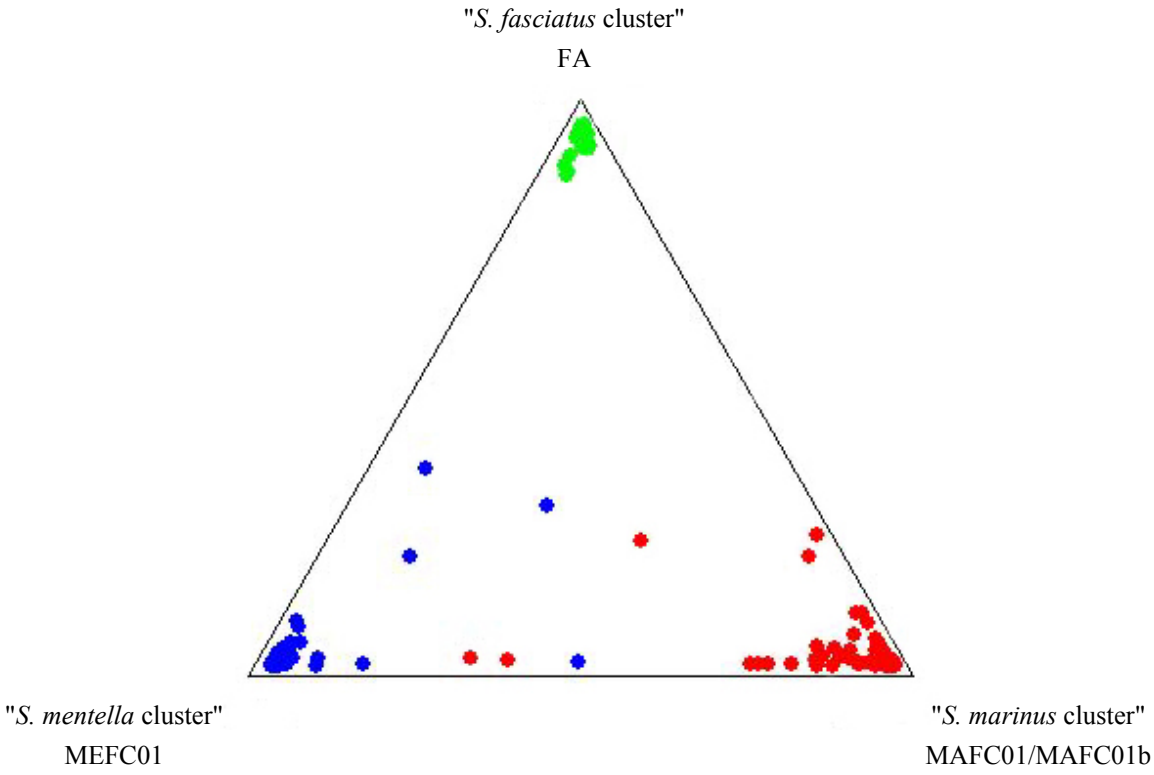


Figure 2.7. Triangle plot of *Q*, representing a summary of the clustering results obtained using *STRUCTURE* for samples of *S. mentella*, *S. marinus* and *S. fasciatus* from the Flemish Cap (FA, MAFC0, MAFC01b and MEFC01) without using prior information of species origin (n = 262, six microsatellite loci, without SEB37 and SEB46, correlated allele frequency model). Each individual is represented by a coloured point. The estimated ancestry vector for an individual consists of K components which add up to 1. When K = 3, the ancestry vectors can be plotted onto a triangle, as shown. For a given point, each of the three components is given by the distance to one edge of the triangle. Individuals who are in one of the corners are therefore assigned completely to one cluster or another (Pritchard et al. 2000). After the clustering was performed, the points were labelled according to their sample of origin: Blue: MEFC01; red: MAFC01 and MAFC01b; green: FA.

Individual ancestry of "giant" *S. marinus*

The *STRUCTURE* analysis of the pooled samples of "giant" *S. marinus*[1] (MAG96) and the baseline samples MAWF02, MASWIc01 and MESIrm01 revealed that all individuals of "giant" *S. marinus* were assigned to the same cluster as the individuals of the sample MASWIc01 (with high individual *Q* values between 0.84 and 0.99), in congruence with the results of the assignment analyses (see paragraph 4.3.1). The corresponding triangle plot (Figure 2.8) further illustrates these findings: All individuals of "giant" *S. marinus* were placed in one of the corners of the triangle - together with the individuals of the sample MASWIc01 - and are therefore assigned completely to the "*S. marinus* cluster".

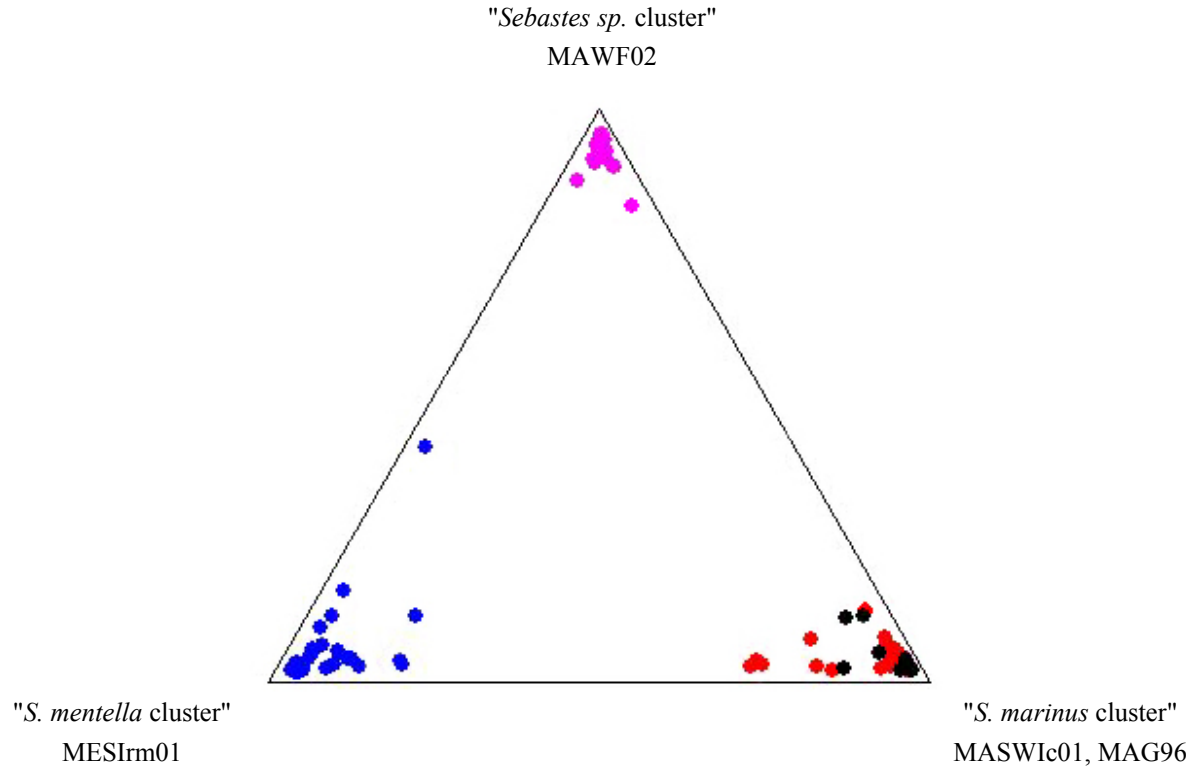


Figure 2.8. Triangle plot of *Q*, representing a summary of the clustering results obtained using *STRUCTURE* for the sample of "giant" *S. marinus* (MAG96) and the baseline samples MESIrm01, MASWIc01 and MAWF02 without using prior information of species origin (n = 151, seven microsatellite loci, without SEB30, correlated allele frequency model). Each individual is presented by a coloured point. The estimated ancestry vector for an individual consists of K components which add up to 1. When K = 3, the ancestry vectors can be plotted onto a triangle, as shown. For a given point, each of the three components is given by the distance to one edge of the triangle. Individuals who are in one of the corners are therefore assigned completely to one species cluster or another (Pritchard et al. 2000). After the clustering was performed, the points were labelled according to their sample of origin: Blue: MESIrm01; red: MASWIc01; pink: MAWF02; black: MAG96.

[1] As the individuals of "giant" *S. marinus* could not be screened for locus SEB30, the *STRUCTURE* analysis was performed without this locus. The three baseline samples (see 4.3.1) as well as the "giant" *S. marinus* sample did not reveal any null alleles at any locus, therefore all seven remaining loci were used for the *STRUCTURE* analysis.

4.3.3 Population structure of S. marinus

To investigate, whether the two main genetic clusters found in the *S. marinus* data set (see paragraph 4.3.1) show further substructuring, *STRUCTURE* was run for both groups of individuals separately. Calculations were performed without locus SEB37, as this locus revealed null alleles, which can cause an overestimation of the number of populations *K* inferred by *STRUCTURE* (Pritchard et al. 2000; Pritchard and Wen 2003).
First, all 220 individuals assigned to cluster III - the "*Sebastes sp.* cluster" (see 4.3.1 and Table 2.8) - were pooled and *STRUCTURE* was run as described in paragraph 3.8.3. Table 2.9 shows the results of the *STRUCTURE* analysis for both allele frequency models. The results of the *STRUCTURE* analysis revealed that the most likely number of populations present in the data set was one (lnPr($X|K$) = -7623.50), when the model assuming independent allele frequencies was applied. When the model assuming correlated allele frequencies was applied, the most likely number of populations was ten (lnPr($X|K$) = -7619.30; $\alpha = 5.679$). In the latter case, the visual inspection of estimated values of Q showed a uniform distribution with $Q \sim 1/K$ (= 1/10), such that all individuals were deemed equally derived from each of the ten populations inferred by *STRUCTURE*.

Table 2.9. Results of the Bayesian clustering analysis performed using *STRUCTURE* (Pritchard et al. 2000) with all individuals assigned to the "*Sebastes sp.* cluster" (220 individuals, seven microsatellite loci, without SEB37, without using prior population information). The natural logarithm of the probability of the data lnPr($X|K$) and its variance for each K (= number of populations) are given. Alpha (α) values indicate the admixture value. Lowest K, lnPr($X|K$) and α values are in bold type.

	Allele frequencies independent			Allele frequencies correlated		
K	Estimated ln probability lnPr($X|K$)	Variance lnPr($X|K$)	Mean value of α	Estimated ln probability lnPr($X|K$)	Variance lnPr($X|K$)	Mean value of α
1	**-7623.50**	**93.10**	-	-7628.10	103.70	-
2	-7705.80	359.70	6.990	-7629.90	103.90	4.799
3	-7856.40	702.50	6.875	-7721.10	297.60	5.651
4	-7817.40	627.50	7.415	-7623.90	102.70	5.780
5	-8003.00	1026.90	6.664	-7665.10	182.30	4.914
6	-8380.10	1797.00	6.405	-7648.60	144.90	4.516
7	-8171.90	1367.80	6.634	-7675.00	197.20	5.298
8	-8353.70	1746.40	5.980	-7778.00	424.10	4.068
9	-8529.10	2062.60	7.132	-7629.90	103.10	4.037
10	-8531.50	2077.60	6.415	**-7619.30**	**93.00**	**5.679**

Second, all 627 individuals assigned to cluster V - the "*S. marinus* cluster" (see Table 2.8) - were pooled and *STRUCTURE* was run as described previously. The "giant *S. marinus*" sample was not included in this analysis, due to the lack of data for locus SEB30.
The *STRUCTURE* analysis revealed that the most likely number of populations/clusters present in the data set was two for the model assuming independent allele frequencies (lnPr($X|K$) = - 23258.5; $\alpha = 1.675$) and three for the model assuming correlated allele

frequencies ($\ln \Pr(X|K) = -21526.6$; $\alpha = 0.077$). Table 2.10 shows the results of the *STRUCTURE* analysis for both allele frequency models.

Table 2.10. Results of the Bayesian clustering analysis performed using *STRUCTURE* (Pritchard et al. 2000) with all individuals assigned to the "*S. marinus* cluster" (627 individuals, seven microsatellite loci, without locus SEB37, without using prior population information). The natural logarithm of the probability of the data $\ln\Pr(X|K)$ and its variance for each K (= number of clusters) are given. Alpha (α) values indicate the admixture value. Lowest K, $\ln\Pr(X|K)$ and α values are in bold type.

	Allele frequencies independent			Allele frequencies correlated		
K	Estimated ln probability $\ln\Pr(X\|K)$	Variance $\ln\Pr(X\|K)$	Mean value of α	Estimated ln probability $\ln\Pr(X\|K)$	Variance $\ln\Pr(X\|K)$	Mean value of α
1	-23264.0	105.6	-	-21951.8	106.3	-
2	**-23258.5**	**1017.4**	**1.675**	-21794.5	714.1	0.301
3	-23382.3	1958.4	1.180	**-21526.6**	**966.1**	**0.077**
4	-23553.9	2691.2	1.104	-21727.7	1549.5	0.062
5	-23589.0	3154.5	0.991	-21936.2	2275.4	0.063
6	-23691.1	3680.0	0.893	-23117.6	4826.7	0.062
7	-23883.4	4312.7	0.809	-24670.8	8303.1	0.059
8	-24107.4	4959.6	0.746	-24110.1	7255.6	0.056
9	-24368.7	5658.1	0.690	-25904.7	10626.2	0.057
10	-24613.6	6322.4	0.629	-27685.1	14318.5	0.055

The average proportions of membership of each sample in the inferred clusters, as well as the number of individuals assigned to each cluster with a proportion of membership $Q \geq 0.80$ and the number of unassigned individuals (assigned to more than one cluster with a proportion of membership $Q < 0.80$ in each of the inferred clusters) are given in Table 2.11 (independent allele frequencies model, highest probability with $K = 2$) and 2.12 (correlated allele frequencies model, highest probability with $K = 3$), respectively.

Table 2.11 shows that all samples were rather admixed under the independent allele model. None of the samples showed a high proportion of membership in one cluster and the number of individuals assigned to one of the two clusters with a proportion of membership $Q \geq 0.80$ was very small. Only 5.42% of the 627 individuals could be assigned to one of the two clusters. The remaining 94.58% revealed admixed genotypes. The corresponding plot of Q (Figure 2.9; above) further illustrates the high proportion of admixed genotypes in the samples.

When the more sensitive F model (= correlated allele frequencies model) was applied (Table 2.12), approximately half of the individuals (52.95%) could be assigned to one of the three inferred clusters ($Q \geq 0.80$). The mean proportion of membership of the samples MANO01, MAEF02, MAWF02, MANEIc00, MASEIc00, MASEIc01, MASWIc00 and MASWIc01 was highest in cluster II in most cases - except in case of sample MASEIc00, where the mean proportion of membership in cluster III was higher than in cluster II.

The mean proportion of membership of the samples from Greenland was highest in cluster III, except in case of the samples MAEGrDB01 and MAEGr01, where Q was relatively evenly

distributed in all three clusters. The mean proportion of membership of the samples from the Flemish Cap was highest in cluster I. The corresponding *Q*-plot (Figure 2.9 below) illustrates that most of the individuals from the Flemish Cap, the individuals from Norway/Faroe Islands/Iceland and the individuals from Greenland, respectively, grouped in different clusters (cluster I, II and III, respectively) - even though many individuals revealed admixed genotypes and the general level of admixture was high.

Table 2.11. Results of the Bayesian clustering analysis of North Atlantic *S. marinus* performed using *STRUCTURE* (Pritchard et al. 2000) with the model of independent allele frequencies, including all individuals assigned previously to the "*S. marinus* cluster" (627 individuals, seven microsatellite loci, without SEB37, without using prior population information). The posterior probability of the number of populations in the sample set was maximum with $K = 2$. The table shows the mean proportion of membership (Q) of each sample in each of the two inferred clusters. Each individual was assigned to a single cluster if its Q_i (i = I–II) was equal to or larger than 0.80. The numbers of individuals assigned to each cluster with a proportion of membership $Q \geq 0.80$ are indicated in parentheses. "Unassigned" individuals were assigned to more than one cluster, with a proportion of membership $Q < 0.80$.
In this table, both samples from the Flemish Cap (MAFC01 and MAFC01b) were pooled (= sample MAFC01).

		Inferred clusters		Number of	Percentage of
Sample	*n*	I	II	unassigned individuals	assigned individuals
MANO01	75	0.595	0.405	73	2.67
		(2)	(0)		
MAEF02	55	0.636	0.364	52	5.45
		(3)	(0)		
MAWF02	2	0.642	0.358	2	0.00
		(0)	(0)		
MANEIc00	39	0.529	0.471	37	5.13
		(0)	(2)		
MASEIc00	50	0.493	0.507	47	6.00
		(2)	(1)		
MASEIc01	23	0.561	0.439	20	13.04
		(3)	(0)		
MASWIc00	57	0.653	0.347	56	1.75
		(1)	(0)		
MASWIc01	51	0.643	0.357	46	9.80
		(5)	(0)		
MAEGrDB01	34	0.383	0.617	27	20.59
		(0)	(7)		
MAEGr00	44	0.427	0.573	44	0.00
		(0)	(0)		
MAEGr00b	40	0.364	0.636	37	7.50
		(0)	(3)		
MAEGr01	9	0.515	0.485	8	11.11
		(1)	(0)		
MAWGr01	33	0.393	0.607	33	0.00
		(0)	(0)		
MAFC01	115	0.452	0.548	111	3.48
		(3)	(1)		
Total	627	(20)	(14)	593	5.42

Table 2.12. Results of the Bayesian clustering analysis of North Atlantic *S. marinus* performed using *STRUCTURE* (Pritchard et al. 2000) with the model of correlated allele frequencies, including all individuals assigned previously to the "*S. marinus* cluster" (627 individuals, seven microsatellite loci, without SEB37, without using prior population information). The posterior probability of the number of populations in the sample set was maximum with $K = 3$. The table shows the mean proportion of membership (Q) of each sample in each of the three inferred clusters. Each individual was assigned to a single cluster if its Q_i (i = I–III) was equal to or larger than 0.80. The numbers of individuals assigned to each cluster with a proportion of membership $Q \geq 0.80$ are indicated in parentheses. "Unassigned" individuals were assigned to more than one cluster, with a proportion of membership $Q < 0.80$.
In this table, both samples from the Flemish Cap (MAFC01 and MAFC01b) were pooled (= sample MAFC01).

		Inferred clusters			Number of	Percentage of
Sample	*n*	I	II	III	unassigned individuals	assigned individuals
MANO01	75	0.281	0.505	0.214	48	36.00
		(4)	(20)	(3)		
MAEF02	55	0.134	0.654	0.212	27	50.91
		(0)	(26)	(2)		
MAWF02	2	0.080	0.586	0.334	2	0.00
		(0)	(0)	(0)		
MANEIc00	39	0.289	0.468	0.243	22	43.59
		(5)	(11)	(1)		
MASEIc00	50	0.161	0.357	0.482	24	52.00
		(1)	(8)	(17)		
MASEIc01	23	0.178	0.467	0.355	9	60.87
		(1)	(7)	(6)		
MASWIc00	57	0.168	0.677	0.155	28	50.88
		(3)	(25)	(1)		
MASWIc01	51	0.163	0.651	0.185	26	49.02
		(1)	(22)	(2)		
MAEGrDB01	34	0.367	0.212	0.421	10	70.59
		(9)	(3)	(12)		
MAEGr00	44	0.093	0.193	0.714	18	59.09
		(0)	(0)	(26)		
MAEGr00b	40	0.102	0.086	0.812	13	67.50
		(0)	(0)	(27)		
MAEGr01	9	0.262	0.338	0.401	5	44.44
		(0)	(2)	(2)		
MAWGr01	33	0.184	0.206	0.610	18	45.45
		(1)	(0)	(14)		
MAFC01	115	0.684	0.186	0.130	45	60.87
		(65)	(3)	(2)		
Total	627	(90)	(127)	(115)	295	52.95

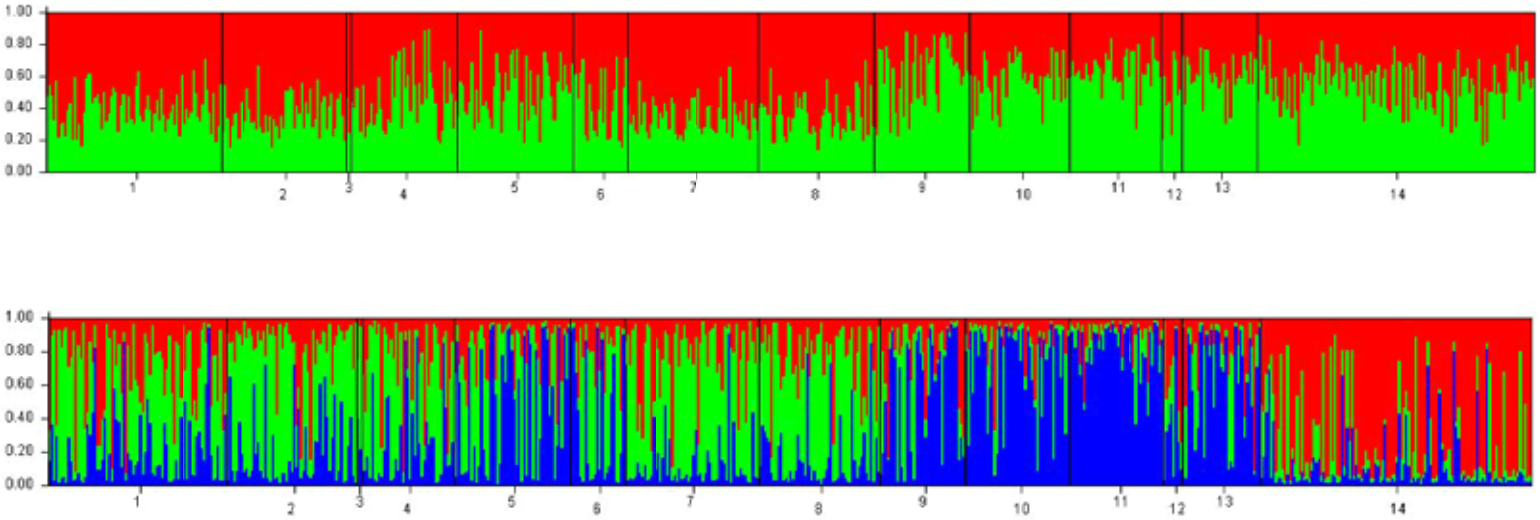


Figure 2.9. Plots of *Q* (estimated membership coefficient for each individual in each cluster) obtained by the Bayesian clustering analysis performed using *STRUCTURE* (Pritchard et al. 2000) with samples of *S. marinus*. Only the 627 individuals that were assigned to the "*S. marinus* cluster" in the previous assignment analyses (see paragraph 4.3.1) were included in this analysis. SEB37 was not used in this analysis due to null alleles. Above: *Q*-plot obtained using the model of independent allele frequencies (the most probable number of populations *K* was two); below: *Q*-plot obtained using the model of correlated allele frequencies (the most probable number of populations *K* was three). Each individual in the data set is represented by a single vertical line, which is partitioned into *K* = 2 (above) or *K* = 3 (below) coloured segments that represent that individuals' estimated membership fraction in each of the *K* inferred clusters. The corresponding mean values of *Q* are given in Table 2.11 and 2.12, respectively. Vertical black lines separate the different samples.
1 = MANO01, 2 = MAEF02, 3 = MAWF02, 4 = MANEIc00, 5 = MASEIc00, 6 = MASEIc01, 7 = MASWIc00, 8 = MASWIc01, 9 = MAEGrDB01, 10 = MAEGr00, 11 = MAEGr00b, 12 = MAEGr01, 13 = MAWGr01, 14 = MAFC01 and MAFC01b.

4.3.4 Population structure of S. mentella

The Bayesian clustering analysis performed using *STRUCTURE* (Pritchard et al. 2000) revealed that the most likely number of populations present in the data set consisting of all individuals (n = 507) assigned to the "*S. mentella* cluster" (cluster II) was one, for both models, the model assuming independent allele frequencies and the model assuming correlated allele frequencies. The *STRUCTURE* analysis was also run excluding locus SEB37 and SEB46, because the two loci showed strong deviations from Hardy-Weinberg expectations in almost all *S. mentella* samples, probably caused by null alleles (see paragraph 4.1). Leaving out these two loci did not affect the results of the *STRUCTURE* analysis. Table 2.13 shows the results of the *STRUCTURE* analysis without locus SEB37 and SEB46 for both allele frequency models. The highest probability of the data was found with K = 1 (allele frequencies independent: $\ln\Pr(X|K)$ = -16892.0; allele frequencies correlated: $\ln\Pr(X|K)$ = -16901.7).

Table 2.13. Results of the Bayesian clustering analysis of *S. mentella* performed using *STRUCTURE* (Pritchard et al. 2000) with all individuals assigned to the "*S. mentella* cluster" pooled (507 individuals, six microsatellite loci, without SEB37 and SEB46, without using prior population information). The natural logarithm of the probability of the data $\ln\Pr(X|K)$ and its variance for each K (= number of clusters) are given. Alpha (α) values indicate the admixture value. Lowest K, $\ln\Pr(X|K)$ and α values are in bold type. The results were similar when the data for locus SEB37 and locus SEB46 were included in the analysis.

	Allele frequencies independent			Allele frequencies correlated		
K	Estimated ln probability $\ln\Pr(X\|K)$	Variance $\ln\Pr(X\|K)$	Mean value of α	Estimated ln probability $\ln\Pr(X\|K)$	Variance $\ln\Pr(X\|K)$	Mean value of α
1	**-16892.0**	**91.3**	**-**	**-16901.7**	**92.6**	**-**
2	-17134.6	1093.1	2.552	-17024.7	876.8	0.832
3	-17348.8	1792.5	2.524	-17229.5	1656.4	0.419
4	-17942.3	3177.5	2.311	-18412.4	4178.6	0.261
5	-18926.5	5283.6	2.144	-17276.1	1934.3	0.152
6	-19951.4	7394.5	2.094	-24881.7	17175.8	0.272
7	-20348.8	8267.8	1.914	-34162.3	35743.3	0.349
8	-20867.8	9379.5	1.748	-24991.9	17638.2	0.086
9	-22996.8	13665.8	1.690	-17064.2	439.9	6.384
10	-25113.6	17945.9	1.596	-16897.5	116.0	4.905

4.4 AMOVA and F-statistics

Based on the results of the assignment analyses using Bayesian model-based clustering (see paragraph 4.3.1 and Table 2.8), individuals of a given sample that were assigned to another genetic cluster than the major part of the individuals of this sample were treated as misclassified specimens (= pre-classified to the wrong species, e.g. on board) and were removed from the data set in the analyses of pairwise F_{ST} and R_{ST} values and in the AMOVA.

Four individuals of the sample of *S. fasciatus,* 18 individuals of the samples of *S. marinus* and 11 individuals of the samples of *S. mentella* were excluded from the analysis (see Table 2.8). Likewise, individuals with admixed genotypes (see Table 2.8) were removed from the *S. mentella* and *S. marinus* data sets in the population analyses.
As null alleles at relatively high frequencies were observed at locus SEB37 in almost all samples of *S. marinus* and additionally at locus SEB46 in samples of *S. mentella* (see paragraph 4.1), AMOVA and F-statistics (F_{ST} and R_{ST}) were carried out (i) excluding SEB37 and SEB46, when the species structure and the population structure of *S. mentella* were analysed, and (ii) excluding SEB37, when the population structure of *S. marinus* was analysed.

4.4.1 Pairwise F_{ST} and R_{ST} values

Table 2.14 shows the pairwise F_{ST} and R_{ST} values and their corresponding Probability (P) values between pooled samples of the four species *S. viviparus, S. fasciatus, S. marinus* and *S. mentella*. The individuals of the *S. marinus* sample were not split into the two clusters inferred by the *STRUCTURE* analysis (see paragraph 4.3.1).

Table 2.14. Pairwise R_{ST} values (above diagonal) and pairwise F_{ST} values (below diagonal) between pooled samples of the four species *S. viviparus, S. fasciatus, S. marinus* and *S. mentella*, based on six microsatellite loci (without locus SEB37 and SEB46). Probability (P) values are given in parentheses. Sample sizes (*n*) are also given. Probability values were adjusted for multiple tests/comparisons using the sequential Bonferroni corrections (Rice 1989): * P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001

Sample (*n*)	*S. viviparus*	*S. fasciatus*	*S. marinus*	*S. mentella*
S. viviparus (42)		0.39056*** (0.00000)	0.11318*** (0.00000)	0.14370*** (0.00000)
S. fasciatus (77)	0.15229*** (0.00000)		0.05750*** (0.00000)	0.29935*** (0.00000)
S. marinus (896)	0.13246*** (0.00000)	0.13320*** (0.00000)		0.03876*** (0.00000)
S. mentella (547)	0.11515*** (0.00000)	0.13748*** (0.00000)	0.06100*** (0.00000)	

All pairwise F_{ST} and R_{ST} values were highly significantly different from zero (P-value < 0.001). F_{ST} values were high and varied between 0.061 (between the *S. marinus* sample and the *S. mentella* sample) and 0.152 (between the *S. fasciatus* sample and the *S. viviparus* sample). The corresponding pairwise R_{ST} values ranged from 0.038 (between the *S. marinus* sample and the *S. mentella* sample) to 0.391 (between the *S. fasciatus* sample and the

S. viviparus sample). In two pairwise comparisons, the R_{ST} was more than twice as high as the corresponding F_{ST}, e.g. between the *S. fasciatus* sample and the *S. viviparus* sample and between the *S. fasciatus* sample and the *S. mentella* sample, whereas similar values for both F_{ST} and R_{ST} were observed between the *S. viviparus* sample and the *S. mentella* sample and between the *S. viviparus* sample and the *S. marinus* sample. In two cases (between the *S. marinus* sample and the *S. mentella* sample and between the *S. marinus* sample and the *S. fasciatus* sample), R_{ST} was smaller than F_{ST}.

In Table 2.15 the pairwise F_{ST} and R_{ST} values and their P-values between the five different genetic clusters inferred by the model-based clustering analysis[1] are given.

Table 2.15. Pairwise R_{ST} values (above diagonal) and pairwise F_{ST} values (below diagonal) between the five genetic clusters inferred by the *STRUCTURE* analysis (IV: "*S. viviparus* cluster"; I: "*S. fasciatus* cluster"; V: "*S. marinus* cluster"; III: "*Sebastes sp.* cluster"; II: "*S. mentella* cluster") based on six microsatellite loci (without SEB37 and SEB46). Individuals assigned to the same cluster were pooled (according to the results of the assignment tests; see Table 2.8). Probability (P) values are given in parentheses. Sample sizes (*n*) are also given. Probability values were adjusted for multiple tests/comparisons using the sequential Bonferroni corrections (Rice 1989): * P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001

Cluster (*n*)		IV *S. viviparus*	I *S. fasciatus*	V *S. marinus*	III *Sebastes sp.*	II *S. mentella*
IV	*S. viviparus* (42)		0.39056*** (0.00000)	0.22862*** (0.00000)	0.07970*** (0.00000)	0.14370*** (0.00000)
I	*S. fasciatus* (80)	0.15229*** (0.00000)		0.07852*** (0.00000)	0.18292*** (0.00000)	0.29935*** (0.00000)
V	*S. marinus* (643)	0.15557*** (0.00000)	0.15149*** (0.00000)		0.23636*** (0.00000)	0.14465*** (0.00000)
III	*Sebastes sp.* (220)	0.15279*** (0.00000)	0.16885*** (0.00000)	0.09187*** (0.00000)		0.10730*** (0.00000)
II	*S. mentella* (507)	0.11515*** (0.00000)	0.13748*** (0.00000)	0.07936*** (0.00000)	0.08625*** (0.00000)	

Also in this case, all pairwise F_{ST} and R_{ST} values were highly significantly different from zero (P-value < 0.001). F_{ST} values were high and varied between 0.079 (between the "*S. marinus* cluster" and the "*S. mentella* cluster") and 0.169 (between the "*Sebastes sp.* cluster" and the "*S. fasciatus* cluster"). Pairwise R_{ST} values ranged from 0.079 (between the "*S. marinus* cluster" and the "*S. fasciatus* cluster") to 0.391 (between the "*S. fasciatus* cluster" and the "*S. viviparus* cluster"). In five pairwise comparisons, the R_{ST} was much higher than the F_{ST}. The

[1] Individuals assigned to the "*S. viviparus* cluster", the "*S. fasciatus* cluster", the "*S. marinus* cluster", the "*S. mentella* cluster" and the "*Sebastes sp.* cluster" (see paragraph 4.3.1), respectively, were pooled.

F_{ST} and R_{ST} values between the "*S. viviparus* cluster" and the "*S. mentella* cluster", between the "*S. fasciatus* cluster" and the "*Sebastes sp.* cluster" and between the "*S. mentella* cluster" and the "*Sebastes sp.* cluster" revealed more or less similar values. In two cases, between the "*S. viviparus* cluster" and the "*Sebastes sp.* cluster" and between the "*S. marinus* cluster" and the "*S. fasciatus* cluster", R_{ST} was smaller than F_{ST}.

Table 2.16, 2.17, 2.18 and 2.19 show matrices of pairwise F_{ST} and R_{ST} values between samples of *S. marinus* (Table 2.16, 2.17 and 2.18) and between samples of *S. mentella* (Table 2.19).

Estimations of pairwise F_{ST} and R_{ST} values between samples of *S. marinus* were carried out in two different approaches: First, pairwise F_{ST} and R_{ST} values between samples of *S. marinus* were estimated without taking into account the results of the FCA and the model-based clustering analyses (see paragraph 4.2 and 4.3.1), which revealed that the individuals pre-classified as *S. marinus* grouped into two distinct genetic clusters. Only the 18 individuals that were assigned to the "*S. viviparus* cluster", the "*S. fasciatus* cluster" and the "*S. mentella* cluster", respectively, in the assignment analyses (see Table 2.8) were excluded from the calculations.

The pairwise F_{ST} and R_{ST} values between these "unsorted" samples of *S. marinus* (Table 2.16) revealed a high degree of genetic differentiation: Most (84) of the 105 pairwise F_{ST} values between samples of *S. marinus* were significantly different from zero after applying the sequential Bonferroni correction (Table 2.16; below diagonal). The significant F_{ST} values ranged from 0.010 to 0.074, with an overall F_{ST} of 0.023. The highest F_{ST} value (0.074) was observed between the sample from the Faroe Islands West (MAWF02) and the sample from Norway (MANO01). In contrast, the pairwise F_{ST} values between the *S. marinus* samples from the geographically most distant sampling locations Norway and the Flemish Cap (MANO01 and MAFC01/MAFC01b) were much smaller (0.010 and 0.015, respectively, but highly significant).

No significant pairwise F_{ST} values were observed between the *S. marinus* samples from Norway, Faroe Islands East, NE-Iceland and SW-Iceland (collected in the years 2000 and 2001). Also pairwise comparisons of the samples from NE-Iceland, the Dohrn Bank (MAEGrDB01) and West Greenland (MAWGr01) did not reveal significant pairwise F_{ST} values. In addition, the samples from the Faroe Islands West and East Greenland 2001 did not show a significant F_{ST} value. Furthermore, no significant F_{ST} values were observed between samples from SE-Iceland (2000 and 2001), East Greenland 2000 (MAEGr00) and West Greenland (MAWGr01). Also the two samples collected on the Flemish Cap were genetically similar. All other pairwise F_{ST} values were significantly different from zero after applying the sequential Bonferroni correction. The samples from East Greenland showed some heterogeneity, as significant differences between different years (between sample MAEGr00 and MAEGr01) as well as between closely located sampling sites (between sample MAEGr00 and MAEGr00b) were observed. In contrast, the two samples from SE-Iceland (MASEIc00

and MASEIc01) as well as the two samples from SW-Iceland (MASWIc00 and MASWIc01), both sample pairs collected in the same areas in two subsequent years, were genetically similar and therefore showed temporal stability.

The number of significant pairwise R_{ST} values was lower than the number of significant pairwise F_{ST} values, with only 48 out of 105 pairwise comparisons being significantly different from zero (Table 2.16, above diagonal). Significant R_{ST} values were higher than F_{ST} values and varied between 0.067 and 0.285. The overall R_{ST} of 0.084 was almost four times as high as the corresponding overall F_{ST}. The samples MANO01 and MAEF02 revealed significant pairwise R_{ST} values when compared with MAWF02 (Faroe Islands West), the two samples from SE-Iceland and all samples from Greenland. Also the pairwise R_{ST} values between the two - genetically similar - samples from SW-Iceland and the samples MAWF02, MASEIc00, MASEIc01 and all samples from Greenland were significant. The pairwise R_{ST} value between MAWF02 and MANEIc00 was also significant. All these significant pairwise comparisons using R_{ST} were also significant when F_{ST} was used for estimating genetic differentiation.

In contrast, the two samples from the Flemish Cap, which were significantly different from all other samples in terms of F_{ST}, did not show a similar R_{ST} pattern. The pairwise R_{ST} values between these two samples and the samples from Norway, Faroe Islands East, NE-Iceland, and both samples from SW-Iceland were not significant. Furthermore, one of the two samples from the Flemish Cap (MAFC01b) did not show significant differences in terms of R_{ST} compared to one sample from Greenland (MAEGr00b).

As the results of the FCA and the model-based clustering analyses (see paragraph 4.2 and 4.3.1) grouped the individuals collected and pre-classified as *S. marinus* into two distinct genetic clusters, F_{ST} and R_{ST} values were also estimated for both clusters separately. For this purpose, these individuals were separated: First, the individuals that were assigned to the "*Sebastes sp.* cluster" (cluster III) in the *STRUCTURE* analysis were removed from the data set and F_{ST} and R_{ST} values between samples were estimated including only individuals that were assigned to the "*S. marinus* cluster" (cluster V; see paragraph 4.3.1). Two samples - MAWF01 and MAEGr01 - were excluded from the analysis, due to the too low number of individuals assigned to the "*S. marinus* cluster" in the two samples (two and nine individuals, respectively). The pairwise F_{ST} and R_{ST} values between the remaining 13 samples are given in Table 2.17.

Significant pairwise F_{ST} values ranged from 0.009 to 0.045 (overall F_{ST}: 0.012). The maximum F_{ST} value as well as the overall F_{ST} were therefore smaller (up to two times) than when individuals of both clusters were included in the analysis. 45 out of 78 pairwise comparisons were significantly different from zero. Similar to the results presented in Table 2.16, the pairwise F_{ST} values between the two (genetically similar) Flemish Cap samples and all other samples of *S. marinus* were significantly different from zero. In contrast, the

significant differences previously observed between the two samples from SE-Iceland and the samples MANO01, MAEF02, MANEIc00, MASWIc00 and MASWIc01 did not remain significant after excluding the genotypes assigned to the "*Sebastes sp.* cluster" (with one exception - the F_{ST} between MAEF02 and MASEIc00). Also the F_{ST} value between the *S. marinus* samples collected from closely located locations off Greenland, MAEGr00 and MAEGr00b, as well as the F_{ST} between the sample MAEGr00b and sample from West Greenland (MAWGr01), did not remain significant after these genotypes were excluded from the analysis. Furthermore, the samples from East and West Greenland did not show any significant differences in allele frequencies, only the sample from the Dohrn Bank revealed significant genetic differences to the samples from East Greenland (MAEGr00 and MAEGr00b). Also fewer pairwise comparisons between the samples MANEIc00, MASEIc00 and MASEIc01 and the samples from Greenland were significantly different from zero: Only three pairwise F_{ST} values (between MANEIc00 and MAEGr00b, between MASEIc00 and MAEGrDB01 and between MASEIc01 and MAEGr00b) remained significant after the genotypes were sorted into clusters, in contrast to the six significant pairwise F_{ST} values observed before the genotypes were sorted. However, MANO01, MAEF02, MASWIc00 and MASWIc01 still differed significantly from all samples from Greenland and the Flemish Cap, after the genotypes assigned to the "*Sebastes sp.* cluster" were removed.

Similar to the overall F_{ST}, the overall R_{ST} value (0.047) was approximately two times smaller after the genotypes assigned to the "*Sebastes sp.* cluster" were excluded from the analysis of *S. marinus*. Significant R_{ST} values (Table 2.17; above diagonal) ranged from 0.078 to 0.185 (the maximum value was also lower than in the data set including all individuals). Only 20 out of 78 pairwise R_{ST} values were significant. The significant R_{ST} values previously observed between the two *S. marinus* samples from SE-Iceland and the samples MANO01, MAEF02, MASWIc00 and MASWIc01 did not remain significant after excluding the genotypes assigned to the "*Sebastes sp.* cluster" - with two exceptions: The R_{ST} between MASEIc01 and MANO01 and the R_{ST} between MASEIc01 and MASWIc00. These two significant pairwise comparisons, however, were not significant when F_{ST} was applied (see Table 2.17).

Except the non significant R_{ST} between MAWGr01 and MASWIc01, the significant pairwise comparisons - in terms of R_{ST} - between the four genetically similar samples MANO01, MAEF02, MASWIc00 and MASWIc01 and the samples from the shelf of Greenland, MAEGrDB01, MAEGr00, MAEGr00b and MAWGr01 were also observed when the F_{ST} was used to estimate the degree of genetic differentiation (see Table 2.17), as well as when all individuals were included in the analysis (see Table 2.16). The same applies to the significant R_{ST} between one sample from the Flemish Cap (MAFC01) and three samples from Greenland (MAEGrDB01, MAEGr00 and MAEGr00b). The remaining pairwise comparisons between the Flemish Cap samples and all other samples, however, were not significant when R_{ST} was used in the estimation of genetic differentiation, in contrast to the pairwise comparisons using F_{ST}, which were all significantly different from zero.

Table 2.16. Pairwise R_{ST} values (above diagonal) and pairwise F_{ST} values (below diagonal) between samples of *S. marinus* based on genetic data obtained from seven microsatellite loci (without SEB37). Probability (P) values are given in parentheses. Sample sizes (*n*) are also given. Probability values were adjusted for multiple tests/comparisons using the sequential Bonferroni corrections (Rice 1989).
* P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001. Significant values are highlighted in grey.

Sample (*n*)	MANO01	MAEF02	MAWF02	MANEIc00	MASEIc00	MASEIc01	MASWIc00	MASWIc01
MANO01		0.00366 (0.18266)	0.28483*** (0.00000)	0.03381 (0.02237)	0.12897*** (0.00000)	0.22126*** (0.00000)	-0.00709 (0.95446)	-0.00725 (0.85991)
MAEF02	0.00288 -0.04079		0.23097*** (0.00000)	0.01404 (0.13395)	0.10257*** (0.00000)	0.17843*** (0.00000)	0.00856 (0.09346)	-0.00551 (0.73973)
MAWF02	0.07374*** (0.00000)	0.06093*** (0.00000)		0.12667* (0.00050)	0.00755 (0.19968)	-0.00451 (0.53975)	0.26697*** (0.00000)	0.20333*** (0.00000)
MANEIc00	0.00254 (0.08435)	0.00219 (0.16226)	0.05530*** (0.00000)		0.04273 (0.01257)	0.09208 (0.00149)	0.03528 (0.02762)	0.00882 (0.19117)
MASEIc00	0.02602*** (0.00000)	0.02252*** (0.00000)	0.01292** (0.00010)	0.01765*** (0.00000)		0.00171 (0.32611)	0.12747*** (0.00000)	0.09377** (0.00010)
MASEIc01	0.02657*** (0.00000)	0.01724*** (0.00000)	0.01148* (0.00050)	0.01726*** (0.00000)	0.00169 (0.16880)		0.21104*** (0.00000)	0.15838*** (0.00000)
MASWIc00	-0.00183 (0.89942)	0.00336 (0.04069)	0.06968*** (0.00000)	0.00347 (0.06227)	0.02666*** (0.00000)	0.02441*** (0.00000)		-0.00123 (0.39283)
MASWIc01	-0.00164 (0.80170)	-0.00226 (0.93387)	0.07165*** (0.00000)	0.00278 (0.11464)	0.02764*** (0.00000)	0.02372*** (0.00000)	0.00058 (0.37788)	
MAEGrDB01	0.01785*** (0.00000)	0.01490*** (0.00000)	0.04654*** (0.00000)	0.00655 (0.00861)	0.01862*** (0.00000)	0.01558*** (0.00000)	0.02165*** (0.00000)	0.01922*** (0.00000)
MAEGr00	0.02103*** (0.00000)	0.01592*** (0.00000)	0.02557*** (0.00000)	0.01101** (0.00010)	0.00339 (0.02752)	0.00476 (0.02020)	0.02074*** (0.00000)	0.01972*** (0.00000)
MAEGr00b	0.02274*** (0.00000)	0.01930*** (0.00000)	0.06784*** (0.00000)	0.01369*** (0.00000)	0.02444*** (0.00000)	0.02994*** (0.00000)	0.02483*** (0.00000)	0.02240*** (0.00000)
MAEGr01	0.07019*** (0.00000)	0.05779*** (0.00000)	-0.00128 (0.78408)	0.05431*** (0.00000)	0.01119*** (0.00000)	0.00971* (0.00089)	0.06677*** (0.00000)	0.06782*** (0.00000)
MAWGr01	0.01775*** (0.00000)	0.01543*** (0.00000)	0.02987*** (0.00000)	0.00640 (0.01327)	0.00441 (0.02584)	0.00606 (0.01455)	0.01898*** (0.00000)	0.01761*** (0.00000)
MAFC01	0.01022 (0.00000)	0.02075 (0.00000)	0.06587 (0.00000)	0.01314 (0.00000)	0.02785 (0.00000)	0.02914 (0.00000)	0.01141 (0.00000)	0.01695 (0.00000)
MAFC01b	0.01484 (0.00000)	0.02424 (0.00000)	0.05910 (0.00000)	0.01236 (0.00000)	0.02624 (0.00000)	0.02534 (0.00000)	0.01840 (0.00000)	0.02123 (0.00000)

Table 2.16 (continued). Pairwise R_{ST} values (above diagonal) and pairwise F_{ST} values (below diagonal) between samples of *S. marinus* based on genetic data obtained from seven microsatellite loci (without SEB37). Probability (P) values are given in parentheses. Sample sizes (*n*) are also given. Probability values were adjusted for multiple tests/comparisons using the sequential Bonferroni corrections (Rice 1989).
* P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001. Significant values are highlighted in grey.

Sample (*n*)	MAEGrDB01	MAEGr00	MAEGr00b	MAEGr01	MAWGr01	MAFC01	MAFC01b
MANO01	0.19054*** (0.00000)	0.18264*** (0.00000)	0.13365*** (0.00000)	0.20710*** (0.00000)	0.20610*** (0.00000)	-0.00364 (0.60073)	0.02176 (0.01762)
MAEF02	0.13234*** (0.00000)	0.14106*** (0.00000)	0.08827*** (0.00000)	0.16627*** (0.00000)	0.16154*** (0.00000)	0.00020 (0.32294)	0.01157 (0.07890)
MAWF02	0.04872 (0.02396)	0.0144 (0.09346)	0.0585 (0.00871)	0.00138 (0.32977)	0.00715 (0.22552)	0.19386*** (0.00000)	0.19080*** (0.00000)
MANEIc00	0.05099 (0.02069)	0.06437 (0.00129)	0.02508 (0.07692)	0.0802 (0.00198)	0.0769 (0.00406)	0.00684 (0.23720)	0.00580 (0.25760)
MASEIc00	0.00687 (0.21265)	0.00019 (0.36917)	0.00961 (0.15741)	-0.00393 (0.70340)	0.00434 (0.25889)	0.08781* (0.00020)	0.08118*** (0.00000)
MASEIc01	0.02194 (0.07643)	-0.00004 (0.35630)	0.03126 (0.04089)	0.00389 (0.28007)	-0.00261 (0.48807)	0.15179*** (0.00000)	0.14422*** (0.00000)
MASWIc00	0.18271*** (0.00000)	0.17700*** (0.00000)	0.13069*** (0.00000)	0.19689*** (0.00000)	0.19617*** (0.00000)	-0.00139 (0.42412)	0.02176 (0.02148)
MASWIc01	0.11141*** (0.00000)	0.12429*** (0.00000)	0.07141* (0.00059)	0.14616*** (0.00000)	0.14226*** (0.00000)	-0.0061 (0.74587)	0.00323 (0.22879)
MAEGrDB01		-0.00103 (0.42144)	-0.00647 (0.76359)	0.04501 (0.01515)	0.00535 (0.24968)	0.10503*** (0.00000)	0.08466* 0.00040
MAEGr00	0.00968* (0.00050)		0.00677 (0.15414)	0.01835 (0.05782)	-0.00539 (0.71686)	0.11790*** (0.00000)	0.10389*** (0.00000)
MAEGr00b	0.01749*** (0.00000)	0.01212** (0.00010)		0.04384 (0.01574)	0.01671 (0.11246)	0.06743* (0.00069)	0.05254 (0.00188)
MAEGr01	0.04552*** (0.00000)	0.02386*** (0.00000)	0.06430*** (0.00000)		0.01779 (0.09811)	0.13653*** (0.00000)	0.13577*** (0.00000)
MAWGr01	0.00819 (0.00347)	0.00068 (0.36620)	0.01256* (0.00089)	0.02633*** (0.00000)		0.13565*** (0.00000)	0.12378*** (0.00000)
MAFC01	0.02542 (0.00000)	0.02570 (0.00000)	0.03684 (0.00000)	0.06324 (0.00000)	0.02056 (0.00000)		-0.00058 (0.38897)
MAFC01b	0.02220 (0.00000)	0.02345 (0.00000)	0.03533 (0.00000)	0.05741 (0.00000)	0.01785 (0.00000)	-0.00071 (0.74765)	

Table 2.17. Pairwise R_{ST} values (above diagonal) and pairwise F_{ST} values (below diagonal) between samples of individuals assigned to the "*S. marinus* cluster" in the *STRUCTURE* analysis (cluster V; see paragraph 4.3.1). The calculations are based on genetic data obtained from seven microsatellite loci (without SEB37). Probability (P) values are given in parentheses. Sample sizes (*n*) are also given (see also Table 2.8). Probability values were adjusted for multiple tests/comparisons using the sequential Bonferroni corrections (Rice 1989).
* P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001. Significant values are highlighted in grey.

Sample (*n*)	MANO01	MAEF02	MANEIc00	MASEIc00	MASEIc01	MASWIc00	MASWIc01
MANO01 (75)		0.00035 (0.33551)	0.01029 (0.14731)	0.02955 (0.01475)	0.08955* (0.00069)	-0.00769 (0.96099)	-0.00819 (0.89942)
MAEF02 (55)	0.00377 (0.01683)		0.00219 (0.35323)	0.02273 (0.04099)	0.06798 (0.00168)	0.00735 (0.12078)	-0.0043 (0.61826)
MANEIc00 (39)	0.00130 (0.22681)	0.00271 (0.10959)		-0.00145 (0.47985)	0.02346 (0.12543)	0.01619 (0.10791)	-0.00222 (0.50233)
MASEIc00 (50)	0.00496 (0.00851)	0.00902* (0.00069)	0.00277 (0.12999)		-0.00574 (0.54074)	0.03949 (0.00980)	0.01367 (0.09851)
MASEIc01 (23)	0.00229 (0.17177)	-0.00360 (0.92674)	-0.00084 (0.64103)	0.00435 (0.11543)		0.09690* (0.00059)	0.04923 (0.00822)
MASWIc00 (57)	-0.00144 (0.8023)	0.00293 (0.06514)	0.00168 (0.23582)	0.00552 (0.01049)	0.00030 (0.45728)		-0.00248 (0.47451)
MASWIc01 (51)	-0.00139 (0.74517)	-0.00203 (0.88932)	0.00068 (0.37442)	0.00643 (0.00495)	-0.00384 (0.90823)	-0.00006 (0.49282)	
MAEGrDB01 (34)	0.01662*** (0.00000)	0.01557*** (0.00000)	0.00676 (0.01792)	0.01156** (0.00020)	0.00460 (0.11039)	0.02161*** (0.00000)	0.01767*** (0.00000)
MAEGr00 (44)	0.01718*** (0.00000)	0.01337*** (0.00000)	0.00880 (0.00158)	0.00332 (0.08474)	0.00509 (0.06514)	0.01740*** (0.00000)	0.01424*** (0.00000)
MAEGr00b (40)	0.02806*** (0.00000)	0.02603*** (0.00000)	0.01769*** (0.00000)	0.00756 (0.00634)	0.01995*** (0.00000)	0.03029*** (0.00000)	0.02776*** (0.00000)
MAWGr01 (33)	0.01465*** (0.00000)	0.0161*** (0.00000)	0.00348 (0.14692)	0.00513 (0.05326)	0.00748 (0.04445)	0.01639*** (0.00000)	0.01359** (0.00020)
MAFC01 (52)	0.01058*** (0.00000)	0.02289*** (0.00000)	0.01349*** (0.00000)	0.01694*** (0.00000)	0.01714*** (0.00000)	0.01233*** (0.00000)	0.01658*** (0.00000)
MAFC01b (63)	0.01469*** (0.00000)	0.02750*** (0.00000)	0.01316*** (0.00000)	0.01948*** (0.00000)	0.01851*** (0.00000)	0.01922*** (0.00000)	0.02031*** (0.00000)

Table 2.17 (continued). Pairwise R_{ST} values (above diagonal) and pairwise F_{ST} values (below diagonal) between samples of individuals assigned to the "*S. marinus* cluster" in the *STRUCTURE* analysis (cluster V; see paragraph 4.3.1). The calculations are based on genetic data obtained from seven microsatellite loci (without SEB37). Probability (P) values are given in parentheses. Sample sizes (*n*) are also given (see also Table 2.8). Probability values were adjusted for multiple tests/comparisons using the sequential Bonferroni corrections (Rice 1989). * P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001. Significant values are highlighted in grey.

Sample (*n*)	MAEGrDB01	MAEGr00	MAEGr00b	MAWGr01	MAFC01	MAFC01b
MANO01	0.18443***	0.17692***	0.14693***	0.13603***	-0.00606	0.02058
(75)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.79992)	(0.02307)
MAEF02	0.13956***	0.13906***	0.11040***	0.10277*	0.00298	0.01623
(55)	(0.00000)	(0.00000)	(0.00000)	(0.00030)	(0.21156)	(0.04881)
MANEIc00	0.08668	0.08596	0.06222	0.0526	-0.00527	-0.00032
(39)	(0.00485)	(0.00158)	(0.00931)	(0.03089)	(0.69152)	(0.44966)
MASEIc00	0.03131	0.03504	0.01758	0.01568	0.00798	-0.00065
(50)	(0.05277)	(0.02564)	(0.09207)	(0.13345)	(0.15583)	(0.42105)
MASEIc01	-0.00075	-0.00020	-0.00909	-0.0095	0.03613	0.01629
(23)	(0.37719)	(0.33076)	(0.63489)	(0.60301)	(0.02059)	(0.11464)
MASWIc00	0.18516***	0.17794***	0.14983***	0.13770***	-0.00092	0.02567
(57)	(0.00000)	(0.00000)	(0.00000)	(0.00020)	(0.38313)	(0.01584)
MASWIc01	0.12235***	0.12067***	0.09397*	0.08772	-0.00612	0.00595
(51)	(0.00000)	(0.00000)	(0.00030)	(0.00139)	(0.73943)	(0.16602)
MAEGrDB01		-0.00978	-0.00974	-0.01052	0.10289**	0.07289
(34)		(0.90347)	(0.87714)	(0.84407)	(0.00010)	(0.00218)
MAEGr00	0.01044*		-0.00784	-0.01044	0.10187	0.07305
(44)	(0.00069)		(0.73240)	(0.84902)	(0.00000)	(0.00089)
MAEGr00b	0.01901***	0.00355		-0.01149	0.07794*	0.05376
(40)	(0.00000)	(0.08415)		(0.90367)	(0.00059)	(0.00455)
MAWGr01	0.00636	0.00266	0.00816		0.07093	0.04609
(33)	(0.03544)	(0.17147)	(0.01079)		(0.00238)	(0.01436)
MAFC01	0.02761***	0.03004***	0.04473***	0.02358***		-0.00004
(52)	(0.00000)	(0.00000)	(0.00000)	(0.00000)		(0.35294)
MAFC01b	0.02467***	0.03041***	0.04204***	0.02292***	-0.00043	
(63)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.63934)	

Pairwise F_{ST} and R_{ST} values were also estimated between samples of individuals assigned to the "*Sebastes sp.* cluster" (cluster III) collected from different locations. The individuals that were assigned to the "*S. marinus* cluster" (cluster V) in the *STRUCTURE* analysis were removed from the *S. marinus* data set and F_{ST} and R_{ST} values between samples were estimated only including individuals that were assigned to the "*Sebastes sp.* cluster" (see paragraph 4.3.1). Only samples with a sample size consisting of more than 15 individuals were selected (see Table 2.8): The samples from the Faroe Islands West, SE-Iceland (collected in the years 2000 and 2001), East Greenland (collected in the year 2000) and West Greenland. To illustrate that the individuals were assigned to another cluster than most of the individuals collected and pre-classified as *S. marinus*, these samples were renamed: SEBWF02, SEBSEIc00, SEBSEIc01, SEBEGr00, SEBEGr01and SEBWGr01 (SEB instead of MA; see also Table 2.1).

The pairwise F_{ST} and R_{ST} values are given in Table 2.18. None of the pairwise F_{ST} and R_{ST} values between these newly arranged samples were significantly different from zero, before and after the sequential Bonferroni correction.

Table 2.18. Pairwise R_{ST} values (above diagonal) and pairwise F_{ST} values (below diagonal) between the six samples consisting of individuals assigned to the "*Sebastes sp.* cluster" (cluster III; see paragraph 4.3.1) in the *STRUCTURE* analysis. The calculations are based on genetic data obtained from seven microsatellite loci (without SEB37). Sample codes (and sampling localities) are the same as in Table 2.1, except that the abbreviation MA was replaced by SEB. Probability (P) values are given in parentheses. Sample sizes (*n*) are also given. Probability values were adjusted for multiple tests/comparisons using the sequential Bonferroni corrections (Rice 1989). * P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001.

Sample (*n*)	SEBWF02	SEBSEIc00	SEBSEIc01	SEBEGr00	SEBEGr01	SEBWGr01
SEBWF02		0.00966	0.01344	-0.00026	-0.00304	0.01550
(32)		(0.11930)	(0.08861)	(0.39560)	(0.54014)	(0.14454)
SEBSEIc00	-0.00017		-0.01290	-0.00730	-0.00369	0.00026
(49)	(0.48579)		(0.95406)	(0.79428)	(0.63895)	(0.42798)
SEBSEIc01	-0.00163	0.00329		-0.02429	-0.00691	-0.00596
(27)	(0.58044)	(0.11623)		(0.99455)	(0.72250)	(0.57341)
SEBEGr00	-0.00093	-0.00327	0.00066		-0.00540	-0.00123
(23)	(0.53024)	(0.86279)	(0.27690)		(0.67162)	(0.46302)
SEBEGr01	0.00069	0.00337	0.00258	-0.00275		-0.00534
(46)	(0.34650)	(0.09930)	(0.15513)	(0.79507)		(0.58836)
SEBWGr01	0.00336	0.00067	0.00328	-0.00643	-0.00200	
(17)	(0.20404)	(0.37917)	(0.14969)	(0.89635)	(0.61806)	

The matrices of pairwise F_{ST} and R_{ST} values between samples of *S. mentella* are presented in Table 2.19.

Overall, the degree of genetic differentiation based on allele frequencies (F_{ST}) between samples of *S. mentella* from different locations and years was small (Table 2.19; below diagonal). Significant F_{ST} values varied between 0.007 and 0.018, with 24 out of 45 pairwise F_{ST} values being significantly different from zero after applying the sequential Bonferroni correction (overall $F_{ST} = 0.008$).

The pairwise F_{ST} values revealed a heterogeneous pattern of temporal and spatial genetic structure in the samples of *S. mentella*: Genetic differences between temporal samples from the same location were in one case significant (F_{ST} between MEEGr00 and MEEGr01: 0.009) and in another case not significant (F_{ST} between MECIrm01 and MECIrm03: 0.001).

Small but significant F_{ST} values between 0.009 and 0.018 were observed between the *S. mentella* sample from the central Irminger Sea 2003 and all other samples of *S. mentella*, except the sample collected in the central Irminger Sea in the year 2001. The sample from the central Irminger Sea 2001 was significantly different (F_{ST} values between 0.010 and 0.015) from the samples from East Greenland collected in the year 2000 (but not 2001), West Greenland, the southern Irminger Sea, NAFO area 1F, NAFO area 2J and the Flemish Cap.

Furthermore, small but significant genetic differences (F_{ST} values between 0.007 and 0.018) were observed between the *S. mentella* sample from the Flemish Cap (MEFC01) and all other North Atlantic *S. mentella* samples, except sample MEWGr01 and sample MENAFO1F01, respectively. Also the sample from East Greenland collected in the year 2000 differed significantly (F_{ST} values between 0.007 and 0.014) from almost all other samples - except from the *S. mentella* sample from SE-Iceland. No significant F_{ST} values were found between the *S. mentella* samples from East Greenland, West Greenland, the southern Irminger Sea, NAFO area 1F and NAFO area 2J, all collected in the same year (2001). The sample from SE-Iceland (MESEIc01) differed significantly from the sample from the central Irminger Sea collected in 2003 ($F_{ST} = 0.009$) and the sample from the Flemish Cap ($F_{ST} = 0.007$), but did not reveal significant F_{ST} values in all other pairwise comparisons.

In contrast to the 24 significant pairwise F_{ST} values, only 10 out of 45 pairwise R_{ST} values were significantly different from zero (Table 2.19; above diagonal). Similar to the observation in *S. marinus*, significant R_{ST} values were higher than F_{ST} values and varied between 0.031 and 0.111. The overall R_{ST} of 0.026 was approximately three times as high as the corresponding overall F_{ST}.

Significant R_{ST} values were observed between the sample from the central Irminger Sea 2001 and all other samples (R_{ST} between 0.047 and 0.111) - except the sample from the central Irminger Sea 2003 - and between the latter sample and the samples from West Greenland, NAFO area 2J and the Flemish Cap (R_{ST} values between 0.031 and 0.055). These comparisons were also significant, when F_{ST} was used for estimating genetic differentiation (see above). The remaining pairwise comparisons were not significant when R_{ST} was used.

Table 2.19. Pairwise R_{ST} values (above diagonal) and pairwise F_{ST} values (below diagonal) between samples of *S. mentella* based on genetic data obtained from on six microsatellite loci (without SEB37 and SEB46). Probability (P) values are given in parentheses. Sample sizes (*n*) are also given. Probability values were adjusted for multiple tests/comparisons using the sequential Bonferroni corrections (Rice 1989). * P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001. Significant values are highlighted in grey.

Sample (*n*)	MESEIc01	MECIrm01	MECIrm03	MEEGr00	MEEGr01
MESEIc01 (48)		0.04709 (0.00158)	0.01668 (0.02683)	-0.00317 (0.68181)	0.01550 (0.06554)
MECIrm01 (47)	0.00474 (0.01366)		0.00745 (0.11019)	0.06609* (0.00030)	0.07578** (0.00010)
MECIrm03 (56)	0.00887** (0.00000)	0.00109 (0.24641)		0.02335 (0.00465)	0.01664 (0.02505)
MEEGr00 (61)	0.00099 (0.26324)	0.01456*** (0.00000)	0.01571*** (0.00000)		0.00169 (0.35046)
MEEGr01 (40)	0.00283 (0.11712)	0.00512 (0.01416)	0.00774* (0.00208)	0.00886** (0.00030)	
MEWGr01 (50)	0.00689 (0.00426)	0.01198*** (0.00000)	0.01637*** (0.00000)	0.01248*** (0.00000)	-0.00030 (0.55876)
MESIrm01 (44)	0.00264 (0.08960)	0.01000*** (0.00000)	0.01176** (0.00010)	0.00722* (0.00069)	0.00015 (0.44867)
MENAFO1F01 (46)	0.00723 (0.00317)	0.01247*** (0.00000)	0.01536*** (0.00000)	0.01427*** (0.00000)	0.0024 (0.16177)
MENAFO2J01 (45)	0.00539 (0.01535)	0.01023** (0.00010)	0.01266*** (0.00000)	0.0135*** (0.00000)	0.00069 (0.36620)
MEFC01 (55)	0.00717* (0.00069)	0.01430*** (0.00000)	0.01839*** (0.00000)	0.00900*** (0.00000)	0.00926** (0.00030)

Table 2.19 (continued). Pairwise R_{ST} values (above diagonal) and pairwise F_{ST} values (below diagonal) between samples of *S. mentella* based on genetic data obtained from on six microsatellite loci (without SEB37 and SEB46). Probability (P) values are given in parentheses. Sample sizes (*n*) are also given. Probability values were adjusted for multiple tests/comparisons using the sequential Bonferroni corrections (Rice 1989). * P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001. Significant values are highlighted in grey.

Sample (*n*)	MEWGr01	MESIrm01	MENAFO1F01	MENAFO2J01	MEFC01
MESEIc01 (48)	0.01084 (0.11078)	-0.00415 (0.74557)	0.01147 (0.09643)	0.02875 (0.00802)	0.01520 (0.03881)
MECIrm01 (47)	0.08993*** (0.00000)	0.04709* (0.00089)	0.06801*** (0.00000)	0.11121*** (0.00000)	0.10948*** (0.00000)
MECIrm03 (56)	0.03082* (0.00119)	0.01005 (0.07247)	0.01505 (0.02584)	0.03931* (0.00030)	0.05494*** (0.00000)
MEEGr00 (61)	-0.00089 (0.49361)	-0.00697 (0.95406)	0.00256 (0.28987)	0.00493 (0.19879)	0.00847 (0.08603)
MEEGr01 (40)	-0.00087 (0.51094)	-0.00100 (0.51876)	-0.00884 (0.96931)	-0.00228 (0.60242)	0.02218 (0.00683)
MEWGr01 (50)		0.00117 (0.38828)	-0.0053 (0.79804)	-0.00378 (0.68023)	0.00471 (0.1784)
MESIrm01 (44)	-0.00018 (0.45817)		-0.00055 (0.47421)	0.00811 (0.14523)	0.01780 (0.0197)
MENAFO1F01 (46)	0.00094 (0.28819)	0.00467 (0.02010)		-0.00195 (0.53846)	0.01840 (0.01267)
MENAFO2J01 (45)	0.00128 (0.23295)	0.00202 (0.13840)	0.0004 (0.37887)		0.01660 (0.01812)
MEFC01 (55)	0.00521 (0.01109)	0.00657* (0.00099)	0.00654 (0.00317)	0.00989*** (0.00000)	

4.4.2 Analysis of molecular variance (AMOVA)

The results of the AMOVA analysing the degree of differentiation among species and among samples within species[1] (see Table 2.20) revealed that 90.51% of the total variance in allele frequencies was found within the samples. The divergence among the species significantly ($P < 0.001$) accounted for 7.77% of the total genetic variance in allele frequencies. The smallest - but still highly significant - part (1.72%) of the total variance in allele frequencies was due to differences between samples within species.
When the variance in allele sizes was analysed, 88.02% of the total variance was found within the samples. The divergence among the four *Sebastes* species significantly ($P < 0.05$) accounted for 4.50% of the total genetic variance. An even higher part of the total variance in allele sizes (7.48%; $P < 0.001$) was due to differences between samples within species.

Table 2.20. Results of the analysis of molecular variance (AMOVA) with variance in allele frequencies and allele sizes partitioned among species, samples within species and within samples[1] (d.f. = degrees of freedom). The calculations were based on genetic data obtained from six microsatellite loci (without SEB37 and SEB46).
* P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001

		Variance in allele frequencies		Variance in allele sizes	
	d.f.	Variance components	% of variation	Variance components	% of variation
Among species	3	0.21867	7.77***	113.13329	4.50**
Among samples within species	23	0.04841	1.72***	187.99339	7.48***
Within samples	3095	2.54889	90.51***	2212.50959	88.02***
Total	3121	2.81597	100.00	2513.63627	100.00

Table 2.21 shows the results of the second AMOVA, in which the degree of differentiation among the five genetic clusters inferred by *STRUCTURE*[2] (Pritchard et al. 2000; see 4.3.1), among the samples within these five clusters and within the samples was analysed. Individuals originating from the same sample and assigned to the same cluster were grouped. In case of *S. marinus*, samples were divided into two clusters and individuals were sorted into samples as described in paragraph 4.4.1 (see Table 2.17 and 2.18).
The results of this AMOVA revealed that the largest part of the total variance in allele frequencies was found within the samples (89.47%). The genetic divergence among the five

[1] without splitting the *S. marinus* samples into the two groups inferred by the *STRUCTURE* analysis (see paragraph 4.3.1) and excluding locus SEB37 and locus SEB46
[2] the "*S. viviparus* cluster", the "*S. fasciatus* cluster", the "*S. marinus* cluster", the "*S. mentella* cluster" and the "*Sebastes sp.* cluster" (see paragraph 4.3.1)

clusters significantly (P < 0.001) accounted for 9.65% of the total genetic variance in allele frequencies. The smallest - but still highly significant - part (0.88%) of the total variance in allele frequencies was due to genetic differences between samples within clusters.
When the variance in allele sizes was analysed, 81.21% of the total genetic variance was found within the samples. The divergence among the five clusters significantly (P < 0.001) accounted for 16.04% of the total variance in allele sizes. 2.74% of the total variance in allele sizes (highly significant; P < 0.001) was due to differences between samples within clusters.

Table 2.21. Results of the analysis of molecular variance (AMOVA) of the North Atlantic *Sebastes* microsatellite data set with variance in allele frequencies and allele sizes partitioned among clusters inferred by *STRUCTURE*[1] (Pritchard et al. 2000), samples within clusters and within samples (d.f. = degrees of freedom). The calculations were based on genetic data obtained from six microsatellite loci (without SEB37 and SEB46).
* P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001

		Variance in allele frequencies		Variance in allele sizes	
Source of variation	d.f.	Variance components	% of variation	Variance components	% of variation
Among genetic clusters	4	0.27045	9.65***	416.01631	16.04***
Among samples within genetic clusters	26	0.02479	0.88***	71.11065	2.74***
Within samples	2989	2.50694	89.47***	2105.93039	81.21***
Total	3019	2.80219	100.00	2593.05735	100.00

Table 2.22. Results of the locus-by-locus analysis of molecular variance of North Atlantic *Sebastes*. The percentages of variation in allele frequencies among clusters inferred by *STRUCTURE*[1] (Pritchard et al. 2000), samples within clusters and within samples are given for each locus, along with the appropriate fixation indices (F_{CT}, F_{SC}, and F_{ST}, respectively[2]). All indices were significantly different from zero (P < 0.001).

	Among genetic clusters		Among samples within genetic clusters		Within samples	
Locus	% of variation	F_{CT}	% of variation	F_{SC}	% of variation	F_{ST}
Seb 9	16.87	0.1687	1.56	0.0187	81.57	0.1843
Seb 25	6.94	0.0694	1.32	0.0142	91.74	0.0826
Seb 30	8.15	0.0815	0.87	0.0094	90.98	0.0902
Seb 31	15.81	0.1581	0.76	0.0091	83.43	0.1657
Seb 33	1.51	0.0151	0.31	0.0032	98.17	0.0183
Seb 45	8.46	0.0846	0.75	0.0082	90.79	0.0921

[1] the "*S. viviparus* cluster", the "*S. fasciatus* cluster", the "*S. marinus* cluster", the "*S. mentella* cluster" and the "*Sebastes sp.* cluster" (see paragraph 4.3.1)

[2] F_{CT} = the amount of genetic variation attributable to genetic differentiation among groups of samples; F_{SC} = the amount of genetic variation attributable genetic differentiation among samples within groups; F_{ST} = the amount of genetic variation attributable to genetic differentiation among samples relative to the total sample

The corresponding results of the locus-by-locus-AMOVA based on differences in allele frequencies (Table 2.22) revealed that all six loci - in particular locus SEB9 and SEB31, but also SEB25, SEB30, and SEB45 - contributed to a high extent to the genetic variation. SEB33 showed less variation than the other five loci, but the variance components attributed to this locus were nevertheless highly significant. All three fixation indices (F_{CT}, F_{SC}, and F_{ST}, respectively[1]) were significantly different from zero ($P < 0.001$) for each locus.

Table 2.23 and 2.24 show the results of the AMOVAs, in which the variance distributions among temporal samples of *S. marinus* from the same location relative to the variance distributions among *S. marinus* samples from different locations were calculated.
In the first AMOVA (Table 2.23), the two clusters revealed by the FCA and model-based clustering analyses (see paragraph 4.2 and 4.3.1) were not taken into account. Only the 18 individuals that were assigned to another species in the assignment test (*S. fasciatus*, *S. mentella* and *S. viviparus*; see Table 2.8) were excluded. All other genotype data were left in the data set and spatial samples as well as temporal samples within spatial samples were grouped.

Table 2.23. Results of the analysis of molecular variance (AMOVA) of the *S. marinus* samples with variance in allele frequencies and allele sizes partitioned among locations, among temporal samples from the same locations and within samples (d.f. = degrees of freedom). The calculations were based on genetic data obtained from seven

		Variance in allele frequencies		Variance in allele sizes	
	d.f.	Variance components	% of variation	Variance components	% of variation
Among locations	9	0.03376	1.09	265.77397	7.45*
Among temporal samples from the same location	5	0.03868	1.25***	34.48070	0.97
Within samples	1789	3.01686	97.66***	3263.87276	91.58***
Total	1803	3.08930	100.00	3564.12743	100.00

The results (Table 2.23) show that the largest part of the total variance in allele frequencies (97.66%) was found within samples. Divergence among temporal samples from the same location significantly accounted for 1.25% of the total variance in allele frequencies. In contrast, the amount of genetic variance in allele frequencies due to genetic differentiation between samples from different locations did not contribute significantly to the total variance, although it was still high (1.09%). When the variance in allele sizes was analysed, 91.58% (highly significant; $P < 0.001$) of the total variance was found within the samples. The genetic

[1] F_{CT} = the amount of genetic variation attributable to genetic differentiation among groups of samples; F_{SC} = the amount of genetic variation attributable genetic differentiation among samples within groups; F_{ST} = the amount of genetic variation attributable to genetic differentiation among samples relative to the total sample

variation distributed among samples from different locations significantly accounted for 7.45% of the total variance in allele sizes. The amount of genetic variance in allele sizes due to genetic differentiation between temporal samples from the same location accounted for only 0.97% of the total variance and was not significant.

When individuals that were assigned to the "*Sebastes sp.* cluster" (cluster III) in the *STRUCTURE* analysis (see paragraph 4.3.1) were removed from the *S. marinus* data set[1], the AMOVA (Table 2.24) showed a different pattern. Still the largest part of the total variance in allele frequencies (98.63%) was found within samples. In contrast, the part of the total variance in allele frequencies that was due to genetic differentiation among temporal samples from the same location was very small (0.15%) and not significant, when the individuals assigned to the "*Sebastes sp.* cluster" were not included in the analysis. Furthermore, the amount of genetic variance in allele frequencies that was due to genetic differentiation between samples from different locations did contribute significantly (1.22%; $P < 0.01$) to the total variance, when these individuals were excluded.

When the variance in allele sizes was analysed, however, the results of the AMOVA were similar to the results of the AMOVA, in which the individuals of the "*Sebastes sp.* cluster" were included (see above).

Table 2.24. Results of the analysis of molecular variance (AMOVA) of the *S. marinus* samples with variance in allele frequencies and allele sizes partitioned among locations, among temporal samples from the same locations and within samples (d.f. = degrees of freedom). All individuals assigned to the "*Sebastes sp.* cluster" (cluster III) in the *STRUCTURE* analysis (see paragraph 4.3.1) were not included in this analysis. The calculations were based on genetic data obtained from seven microsatellite loci (without SEB37).
* P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001.

		Variance in allele frequencies		Variance in allele sizes	
	d.f.	Variance components	% of variation	Variance components	% of variation
Among locations	8	0.03648	1.22**	133.1812	5.30*
Among temporal samples from the same location	4	0.00449	0.15	-7.24667	-0.29
Within samples	1275	2.94418	98.63***	2384.97481	94.99***
Total	1287	2.98515	100.00	2510.90934	100.00

The corresponding results of the locus-by-locus-AMOVA based on differences in allele frequencies (Table 2.25) revealed that mainly six loci (SEB9, SEB25, SEB30, SEB31, SEB45 and SEB46) contributed to a similar high extent to the variation among samples from different locations. Mainly one locus (SEB9) contributed to the variation among temporal samples

[1] MAWF01 and MAEGr01 were also excluded from the analysis, due to the too low number of individuals assigned to the "*S. marinus* cluster" in the two samples (two and nine individuals, respectively).

from the same location, whereas all loci contributed to a similar high extent to the variation within samples. Therefore, only F_{ST} (the amount of genetic variation attributable to genetic differentiation among samples relative to the total sample) was significant for all loci, whereas F_{CT} (the amount of genetic variation attributable to genetic differentiation among groups of samples collected from the same location) was significant for locus SEB9, SEB25, SEB30, SEB31, SEB45 and SEB46. F_{SC} (the amount of genetic variation attributable to genetic differentiation among temporal samples from the same location) was significant for only one locus (SEB9).

Table 2.25. Results of the locus-by-locus analysis of molecular variance of the *S. marinus* samples. The percentages of genetic variation in allele frequencies for each locus among temporal samples of *S. marinus* from the same location relative to the percentages of variation among *S. marinus* samples from different locations and within samples are given, along with the appropriate fixation indices (F_{CT}, F_{SC}, and F_{ST}, respectively). All individuals assigned to the "*Sebastes sp.* cluster" (cluster III) in the *STRUCTURE* analysis (see paragraph 4.3.1)

	Among locations		Among temporal samples from the same location		Within samples	
Locus	% of variation	F_{CT}	% of variation	F_{SC}	% of variation	F_{ST}
Seb 9	2.1840	0.0218*	0.6464	0.0066*	97.1696	0.0283***
Seb 25	1.4079	0.0141*	0.3640	0.0037	98.2281	0.0177***
Seb 30	1.2455	0.0125*	0.1190	0.0012	98.6355	0.0136***
Seb 31	1.1807	0.0118*	-0.0103	-0.0001	98.8296	0.0117***
Seb 33	0.2466	0.0025	0.1788	0.0018	99.5747	0.0043***
Seb 45	1.2804	0.0128*	-0.0468	-0.0005	98.7663	0.0123***
Seb 46	1.5890	0.0159***	-0.2716	-0.0028	98.6826	0.0132***

The AMOVA partitioning the variance among temporal samples of *S. mentella* from the same location relative to the variance among *S. mentella* samples from different locations (Table 2.26) revealed that - similar to the results of the previous AMOVAs - the largest part of the total variance in allele frequencies (99.13%) was found within the samples. A smaller - but still significant - part (0.50%; $P > 0.01$) of the total genetic variance in allele frequencies was distributed among temporal samples from the same location. Only 0.36% of the total genetic variation contributed to the genetic differences between samples from different locations. This "among location" variance component was not significant. When the variance in allele sizes was analysed, 97.99% (highly significant; $P < 0.001$) of the total variance in allele sizes was found within the samples. The genetic variation distributed among samples from different locations as well as among temporal samples from the same location accounted for only 1.52% and 0.49% of the total variance in allele sizes, respectively. Both variance components were not significant.

The results of the locus-by-locus-AMOVA based on differences in allele frequencies (Table 2.27) revealed that four loci (SEB25, SEB30, SEB31 and SEB45) contributed to an

approximately similar high extent to the genetic variation among samples from different locations. Mainly one locus (SEB9) contributed to the variation among temporal samples from the same location, whereas all loci contributed to a similar high extent to the variation within samples. Only F_{ST} (the amount of genetic variation attributable to genetic differentiation among samples relative to the total sample) was significant for all loci, whereas F_{CT} (the amount of genetic variation attributable to genetic differentiation among groups of samples collected from the same location) was significant only for locus SEB25 and SEB31. F_{SC} (the amount of genetic variation attributable to genetic differentiation among temporal samples from the same location) was significant only for locus SEB9.

Table 2.26. Results of the analysis of molecular variance (AMOVA) of the *S. mentella* samples with variance in allele frequencies and allele sizes partitioned among locations, among temporal samples from the same locations and within samples (d.f. = degrees of freedom). The calculations were based on genetic data obtained from six

		Variance in allele frequencies		Variance in allele sizes	
	d.f.	Variance components	% of variation	Variance components	% of variation
Among locations	7	0.00883	0.34	0.57457	1.52
Among temporal samples from the same location	2	0.01286	0.50***	0.18565	0.49
Within samples	1078	2.57160	99.16***	36.97070	97.99***
Total	1087	2.59329	100.00	37.73092	100.00

Table 2.27. Results of the locus-by-locus analysis of molecular variance of the samples of *S. mentella*. The percentages of variation in allele frequencies for each locus among temporal samples of *S. mentella* from the same location relative to the percentages of variation among *S. mentella* samples from different locations and within samples are given, along with the appropriate fixation indices (F_{CT}, F_{SC}, and F_{ST}, respectively[1]).
* P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001

	Among locations		Among temporal samples from the same location		Within samples	
Locus	% of variation	F_{CT}	% of variation	F_{SC}	% of variation	F_{ST}
Seb 9	-1.2907	-0.0129	2.36827	0.0234**	98.9224	0.0108***
Seb 25	1.2278	0.0123*	0.19588	0.0020	98.5763	0.0143***
Seb 30	0.7467	0.0075	0.19455	0.0020	99.0588	0.0094***
Seb 31	0.9298	0.0093*	-0.11211	-0.0011	99.1823	0.0082***
Seb 33	-0.0215	-0.0002	0.33293	0.0033	99.6886	0.0031**
Seb 45	0.5057	0.0051	0.18188	0.0018	99.3124	0.0069***

[1] F_{CT} = the amount of genetic variation attributable to genetic differentiation among groups of samples; F_{SC} = the amount of genetic variation attributable genetic differentiation among samples within groups; F_{ST} = the amount of genetic variation attributable to genetic differentiation among samples relative to the total sample

4.5 Isolation by distance

Mantel tests revealed no significant correlation (r = 0.278, P < 0.092; plot not shown) between geographic and genetic distances for samples of *S. marinus*, when pairwise F_{ST} values between samples of *S. marinus* were estimated without taking into account the results of the FCA and the model-based clustering analyses (see paragraph 4.2 and 4.3.1).
However, when all individuals assigned to the "*Sebastes sp.* cluster" in the model-based clustering analysis (*STRUCTURE*; see paragraph 4.3.1) were excluded from the IBD analysis, Mantel tests revealed a significant (though perhaps low) correlation (r = 0.454, P < 0.034) between geographic and genetic distances as measured by [$F_{ST}/(1-F_{ST})$] for samples of *S. marinus* (Figure 2.10) at the spatial scale of this study (from Norway to the Flemish Cap).
No significant correlation (r =0.1767, P < 0.16) between geographic and genetic distances as measured by $F_{ST}/(1-F_{ST})$ was observed for samples of *S. mentella* (Figure 2.11) at the spatial scale of this study (from SE-Iceland to the Flemish Cap).

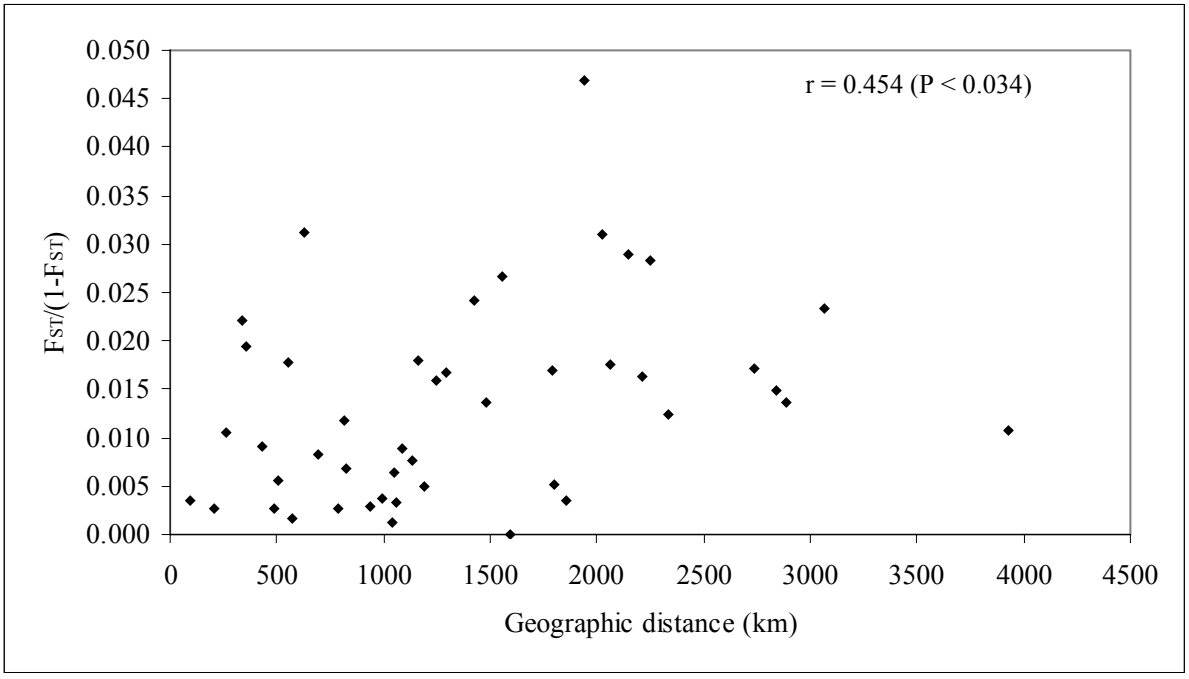


Figure 2.10. Scatterplot of pairwise genetic distances quantified as [$F_{ST}/(1-F_{ST})$] versus pairwise geographic distances (in km) for 13 samples of *S. marinus*. All individuals assigned to the "*Sebastes sp.* cluster" using model-based clustering (see paragraph 4.3.1) were excluded from the analysis. Significance of the correlation was estimated with Mantel's test (r = standardised Mantel test relationship coefficient).

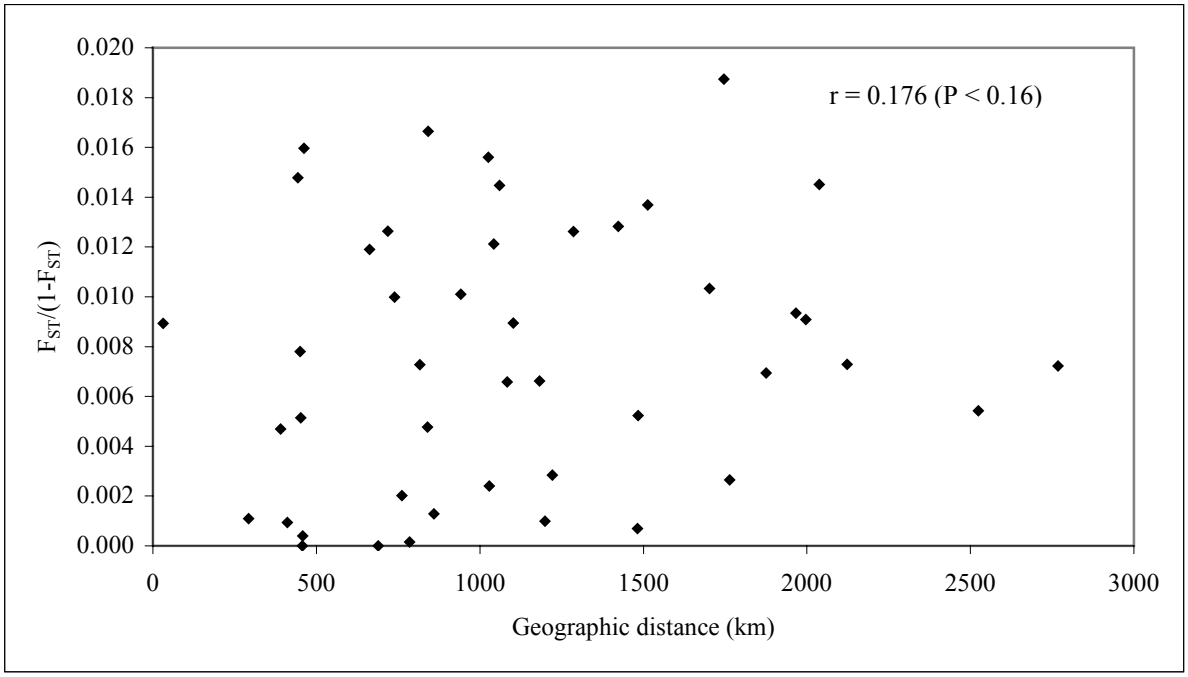


Figure 2.11. Scatterplot of pairwise genetic distances quantified as [$F_{ST}/(1-F_{ST})$] versus pairwise geographic distances (in km) for 10 samples of *S. mentella*. Significance of the correlation was estimated with Mantel's test (r = standardised Mantel test relationship coefficient).

4.6 Multidimensional scaling analysis (MDSA)

Similar to the analyses of pairwise F_{ST} and R_{ST} values, Cavalli-Sforza and Edwards (1967) chord distances (D_{CE}) among all samples of North Atlantic *Sebastes* were first calculated without separating individuals collected and pre-classified as *S. marinus* into the two groups revealed by the results of the FCA (see paragraph 4.2) and the model-based clustering analyses (see paragraph 4.3.1).

The MDSA of Cavalli-Sforza and Edwards (1967) chord distances (D_{CE}) among these samples (one *S. fasciatus* sample, one *S. viviparus* sample, 15 *S. marinus* samples and 10 *S. mentella* samples, see also Table 2.14, 2.16 and 2.19) revealed that 43.6% of the variance could be explained by dimension 1, while 22.1%, 11.7% and 6.3% of the variance were explained by dimension 2, 3 and 4, respectively.

In the MDS plot (Figure 2.12) the samples of the different species were well separated by the first two dimensions. However, even though the 15 samples of *S. marinus* (MA) clustered separately from the samples of *S. viviparus* (VI), *S. fasciatus* (FA) and *S. mentella* (ME), they formed a relatively diffuse cluster and several samples clustered extremely far apart from each other. Furthermore, the grouping was not in congruence with the geographical sampling localities. For example, the sample from Faroe Island West (MAWF02; no. 1 in Figure 2.12) and the sample from the Faroe Islands East (MAEF02; no. 10) clustered far apart from each

other. Also the sample from East Greenland 2001 (MAEGr01; no. 2 in Figure 2.12) and the samples from East Greenland 2000 (MAEGr00 and MAEGr00b; no. 6 and 8) clustered separately. In contrast, the 10 samples of *S. mentella* clustered closely together.

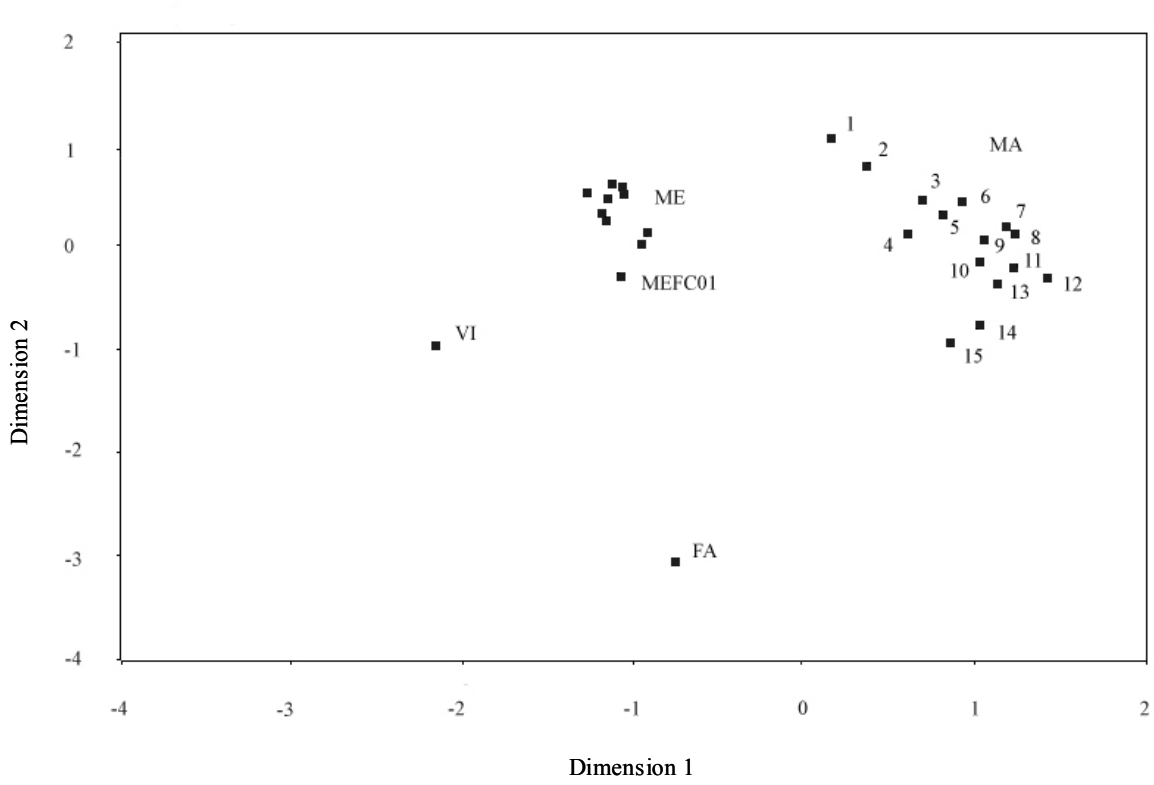


Figure 2.12. Multidimensional scaling plot of Cavalli-Sforza and Edwards (1967) chord distances (D_{CE}) among samples of North Atlantic *Sebastes* species. VI = *S. viviparus,* FA = *S. fasciatus*, ME = *S. mentella* (MEFC01 = Flemish Cap). Samples of *S. marinus* (MA) are coded as following: 1 = MAWF02, 2 = MAEGr01, 3 = MASEIc01, 4 = MAEGrDB01, 5 = MASEIc00, 6 = MAEGr00, 7 = MAWGr01, 8 = MAEGr00b, 9 = MANEIc00, 10 =MAEF02, 11 = MANO01, 12 =MASWIc01, 13 = MASWIc00, 14 = MAFC01, 15 = MAFC01b. Sample codes are explained in Table 2.1. The calculations are based on genetic data obtained from six microsatellite loci (without locus SEB37 and SEB46).

When the *S. marinus* samples were split into two groups - according to the results of the model-based clustering analyses (as described in the analysis of pairwise F_{ST} and R_{ST} values; see paragraph 4.4.1)[1] - clustering of the samples was more distinct. In the MDS plot (Figure 2.13), the samples of the different species were well separated by the first two dimensions, except the *S. fasciatus* sample (FA), which grouped closely with the *S. mentella* sample from the Flemish Cap (MEFC01). However, the *S. fasciatus* sample was well separated from the other samples by dimension 1 and 3 (corresponding plot not shown).

[1] The 15 samples of *S. marinus* were split into 13 samples consisting of individuals that were assigned to the "*S. marinus* cluster" (see also Table 2.17) and six samples consisting of individuals assigned to the "*Sebastes sp.* cluster" (see also Table 2.18).

All samples of *S. mentella* were grouped in a distinct cluster, similar to the results of the previous MDSA. The samples of individuals assigned to the "*Sebastes sp*. cluster" in the *STRUCTURE* analysis (see paragraph 4.3.1) were also grouped into a distinct cluster, well separated from the other samples. In contrast to the previous MDS plot, the samples of *S. marinus* - consisting only of individuals that were assigned to the "*S. marinus* cluster" - also formed a distinct cluster.

The MDSA revealed that 42.1% of the total variance could be explained by dimension 1, while 29.8%, 10.6% and 4.6% of the total variance were explained by dimension 2, 3 and 4, respectively.

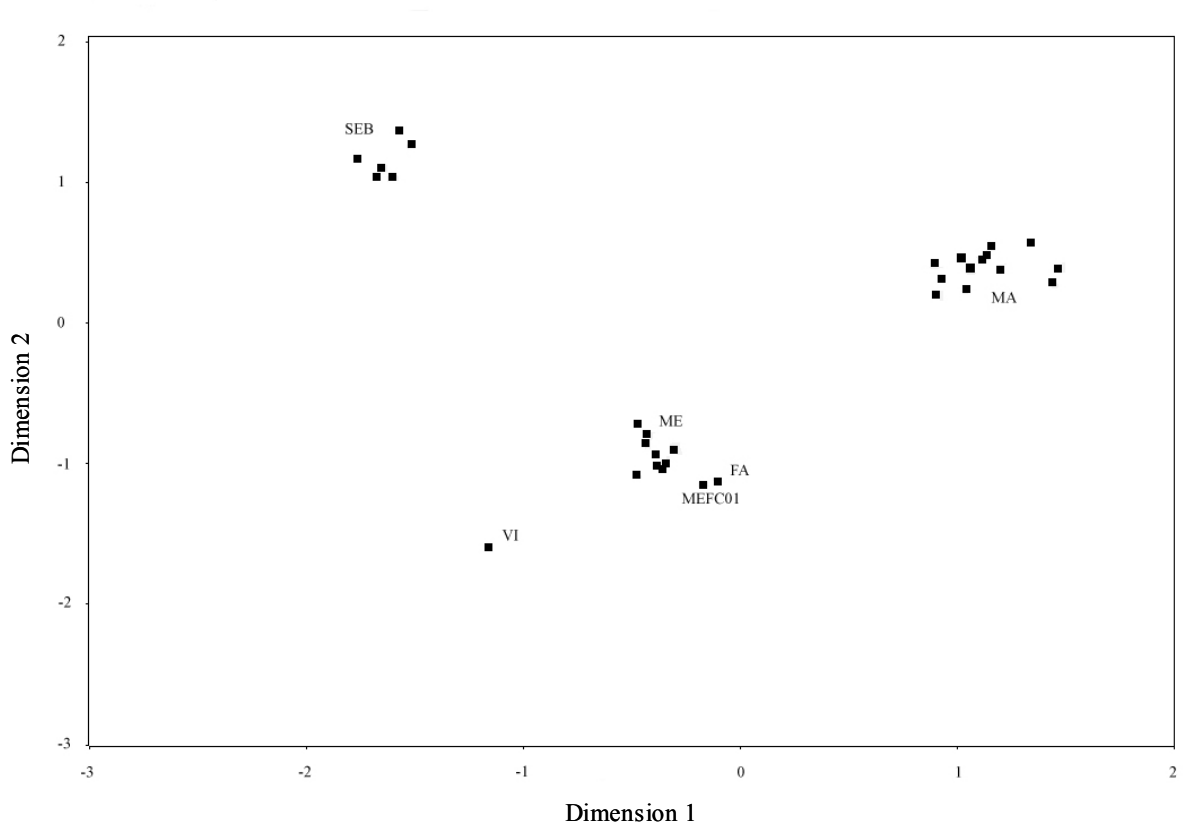


Figure 2.13. Multidimensional scaling plot of Cavalli-Sforza and Edwards (1967) chord distances (D_{CE}) among samples of North Atlantic *Sebastes*. Individuals that were assigned to the "*Sebastes sp.* cluster" (cluster III) and the "*S. marinus* cluster" (cluster V) in the *STRUCTURE* analysis (see paragraph 4.3.1) were separated, as in the analysis of pairwise F_{ST} and R_{ST} values (see paragraph 4.4.1). VI: sample of individuals assigned to the "*S. viviparus* cluster" (one sample); FA: sample of individuals assigned to the "*S. fasciatus* cluster" (one sample); ME: samples of individuals assigned to the "*S. mentella* cluster" (10 samples; see also Table 2.19; MEFC01 = Flemish Cap); MA: samples of individuals assigned to the "*S. marinus* cluster" (13 samples; see also Table 2.17); SEB: samples of individuals assigned to the "*Sebastes sp.* cluster" (six samples; see also Table 2.18). The calculations are based on genetic data obtained from six microsatellite loci (without locus SEB37 and SEB46).

The MDSA of Cavalli-Sforza and Edwards (1967) chord distances (D_{CE}) among the samples of *S. marinus*, when individuals that were assigned to the "*Sebastes sp.* cluster" (cluster III) in the *STRUCTURE* analysis were removed from the data set, revealed that 36.2% of the variance could be explained by dimension 1, while 22.2%, 10.9% and 9.5% of the variance were explained by dimension 2, 3 and 4, respectively.

The corresponding MDS plot (Figure 2.14) shows that the two *S. marinus* samples from the Flemish Cap were well separated from the other samples. The remaining samples were organised along dimension 1, but not in correspondence with their geographical sampling locations. The sample from Norway was placed closer to the samples from the Flemish Cap as for example the samples from Greenland. The two *S. marinus* samples from SW-Iceland (collected in 2000 and 2001) grouped together with the sample from the Faroe Islands East. In contrast, the samples from SE-Iceland collected in two subsequent years did not cluster closely together. The samples from Greenland grouped together, but the samples MEEGr00 and MEEGr00b - collected at relatively closely located sampling sites - were placed apart from each other. The first (sample MEEGr00) grouped with the Dohrn Bank sample (MEEGrDB01) and the sample from West Greenland (MAWGr01).

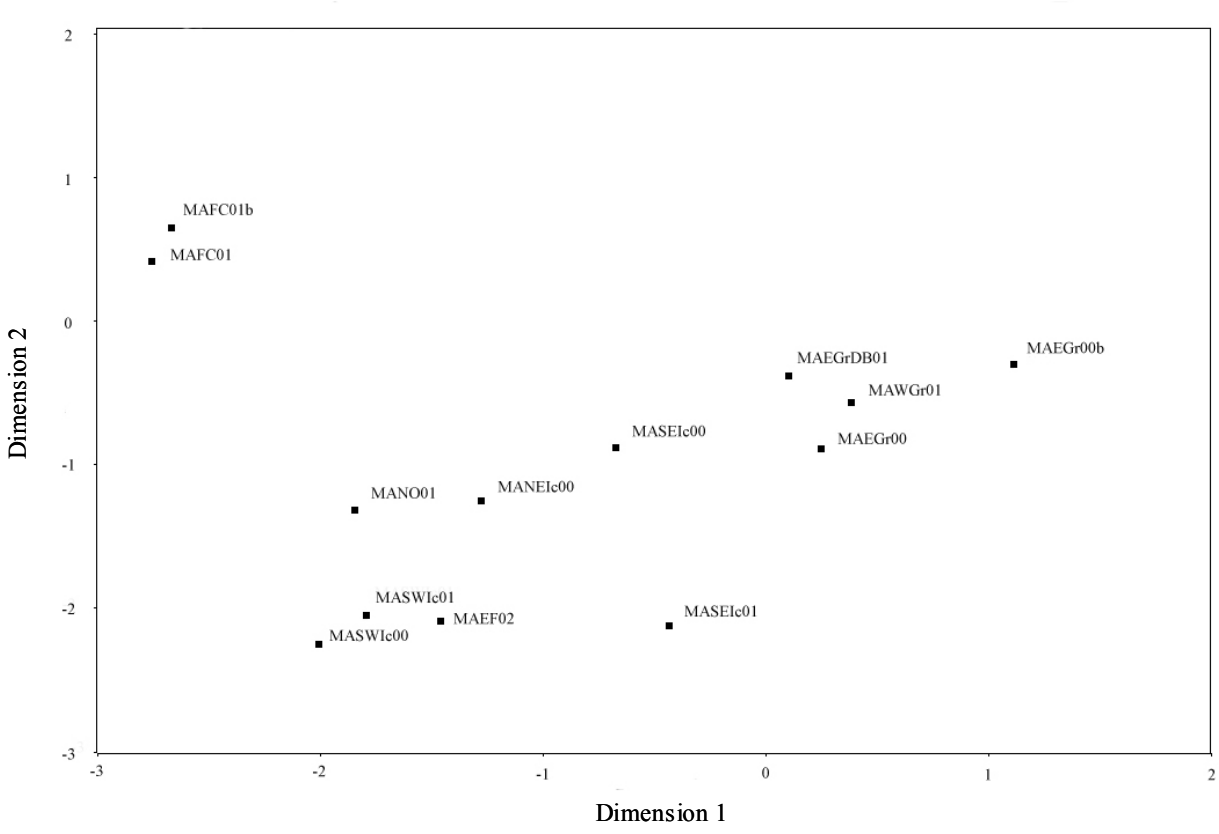


Figure 2.14. Multidimensional scaling plot of Cavalli-Sforza and Edwards (1967) chord distances (D_{CE}) among 13 samples of individuals assigned to the "*S. marinus* cluster" in the *STRUCTURE* analysis (see 4.3.1). The calculations are based on genetic data obtained from seven microsatellite loci (without SEB37). Sample codes are explained in Table 2.1.

The MDSA of Cavalli-Sforza and Edwards (1967) chord distances (D_{CE}) among the 10 samples of *S. mentella* revealed that 33.6% of the variance could be explained by dimension 1, while 20.6%, 12.8% and 9.5% of the variance were explained by dimension 2, 3 and 4, respectively.

The corresponding MDS plot is presented in Figure 2.15. The two *S. mentella* samples from the central Irminger Sea collected in different years (MECIrm01 and MECIrm03) clustered apart from the other samples. Also the *S. mentella* sample from the Flemish Cap was placed separately. The two *S. mentella* samples from the NAFO areas 1F and 2J (MENAFO1F01 and MENAFO2J01), the sample from the southern Irminger Sea (MESIrm01) and the samples from East and West Greenland collected in 2001 (MEEGr01 and MEWGr01) grouped together. The *S. mentella* sample from East Greenland collected one year earlier (MEEGr00) grouped with the sample from SE-Iceland, separately from the sample from East Greenland collected in 2001 (MEEGr01).

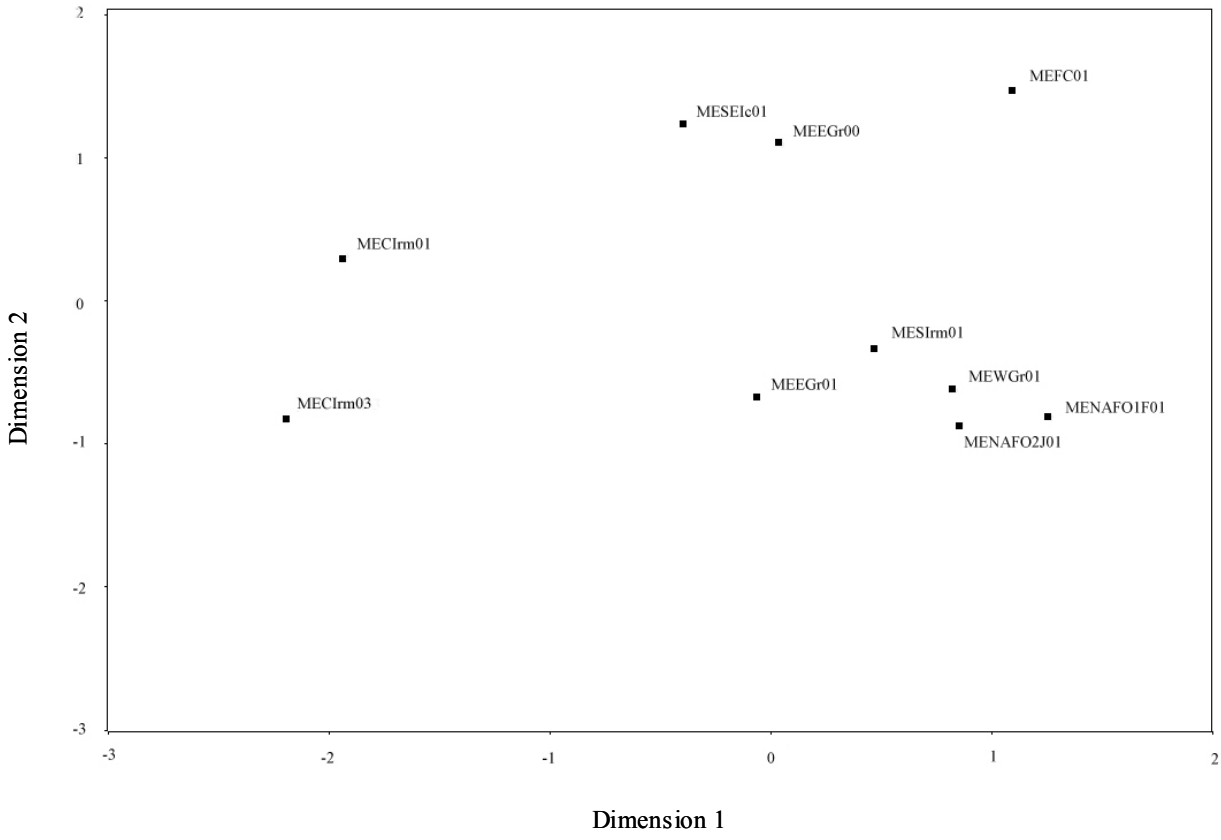


Figure 2.15. Multidimensional scaling plot of Cavalli-Sforza and Edwards (1967) chord distances (D_{CE}) among samples of *S. mentella.* The calculations are based on genetic data obtained from six microsatellite loci (without locus SEB37 and locus SEB46). Sample codes are explained in Table 2.1.

5. Discussion

5.1 Levels of genetic diversity

The high extent of microsatellite allelic diversity and gene diversity estimates observed in this study was comparable to those reported for populations of Northwest Atlantic *Sebastes* (Roques et al. 1999a; 1999b; 2001; 2002), Pacific *Sebastes* species (Miller et al. 2000; Buonaccorsi et al. 2002; Rocha-Olivares et al. 2003) and many other marine fish species (reviewed in DeWoody and Avise 2000). Marine fish species generally possess significant higher levels of genetic diversity (average $H_e = 0.79$) than freshwater (average $H_e = 0.46$) or anadromous (average $H_e = 0.68$) species (DeWoody and Avise 2000). Therefore, the high heterozygosities found in the North Atlantic *Sebastes* samples (average $H_e = 0.81$) are not atypical. The greater genetic variation in marine compared to freshwater or anadromous fishes may reflect the larger evolutionary effective population sizes of marine fish species, which may be related to the larger and more continuous nature of the marine environment, the high potential for dispersal and weak physical barriers to gene flow (Waples 1998; DeWoody and Avise 2000). In redfish, the high genetic variability may be also due to their exceptional longevity, high degree in generation overlap and low temporal variability in effective population size (Rocha-Olivares et al. 1999a).

5.2 Heterozygote deficiencies

Heterozygote deficiencies, as observed in the samples analysed in the present study, have been reported for many fish species (see Hoarau et al. 2002b and references therein). They can be the result of technical artefacts, such as genotyping errors caused by stuttering, null alleles or short allele dominance, as well as linkage of loci. Heterozygote deficiencies can also be caused by biological factors, such as Wahlund effects (see below), or processes such as inbreeding, assortative mating and selection against heterozygotes. As shown in paragraph 4.1, technical artefacts due to genotyping errors caused by stuttering or short allele dominance, as well as linkage of loci, are an unlikely cause of the heterozygote deficiencies in the present data set. Two loci (SEB37 and SEB46), however, revealed null alleles. The other six loci did not reveal null alleles and alternative explanations must be sought for the observed heterozygotes deficiencies at these loci, i.e. Wahlund effects, inbreeding, assortative mating or selection.

Wahlund effects are the most frequent explanation given for observed heterozygote deficiencies (Hoarau et al. 2002b). The Wahlund effect is a common cause of the departure of real data from the Hardy–Weinberg ratio, when two or more subpopulations with different allele frequencies are inadvertently sampled as a single subpopulation.

The Wahlund effect could also explain the heterozygote deficiencies observed in the present study, due to several factors. Firstly, as the general knowledge about the mating areas and the mating period of North Atlantic redfish is very limited, it was not possible to sample close to the actual mating (mating area as well as mating period; see also paragraph 5.6.5). Therefore, the samples behind the presented results could potentially have been collected from mixed feeding aggregations of redfishes from various populations (and even species; see paragraph 5.3). This is supported by the results of the different statistical analyses, which indicated a strong genetic heterogeneity within several samples, implying that samples were often composed of a mixture of redfish from different gene pools. In particular in case of *S. marinus*, a Wahlund effect is a likely explanation for the heterozygotes deficits observed at several loci, as the individual based analyses (FCA and Bayesian model-based clustering using *STRUCTURE*; Pritchard et al. 2000) revealed that the samples of *S. marinus* were genetically extremely heterogeneous (see paragraph 5.3). The same applies to *S. mentella*, even though the heterogeneity was not as prominent as in *S. marinus*. Nevertheless, genetic differences between years (temporal instability) as well as indication of population structure were observed (see paragraph 5.6.4). Roques et al. (2002) observed similar heterozygote deficits in samples of *S. mentella* from Iceland, the Irminger Sea, East Greenland and West Greenland and discussed the possibility of intermingling of populations in these areas.

Furthermore, some samples had to be pooled due to too small sample sizes: The sample of *S. viviparus*, which was used as reference sample in the analysis of species structure, was a pooled sample consisting of individuals from Norway and Iceland. The Icelandic sample was composed of two samples of *S. viviparus* collected in different years (1996 and 2001, respectively). The same applies to the *S. fasciatus* sample from the Flemish Cap, which was also composed of individuals from different stations and years (see Table 2.1 and Appendix I). Also the heterozygote deficits observed at several loci in one of the *S. marinus* samples from the Flemish Cap (MAFC01b) were most likely caused by the heterogeneity of the sample, as it consisted of individuals from thirteen different stations (see Appendix I). Supportingly, the second sample from the Flemish Cap (MAFC01) did not reveal comparable departures from HW expectations.

Non-random mating due to inbreeding or positive assortative mating can also cause heterozygotes deficits. Inbreeding results from mating between relatives, whereas the basis for positive assortative mating is not relatedness but phenotypic similarity.

Inbreeding seems to be a relatively unlikely explanation for the observed heterozygote deficiencies in fish species with presumably large census population sizes (= total number of individuals present) such as redfish. Estimates of population sizes of *S. marinus* and *S. mentella* on the shelf of Greenland range from 10^9 individuals in *S. marinus* to 3 x 10^9 individuals in *S. mentella*, respectively (including only individuals >=17 cm; see ICES 2004c). Until recently, it was generally believed that the majority of wild marine fish populations were sufficiently large to be unaffected by random change between generations

during population declines (reviewed in Bekkevold et al. 2004) - and therefore also inbreeding. This was supported by numerous previous genetic studies, which revealed that marine fishes show usually higher levels of genetic variability and lower levels of genetic differentiation than freshwater or anadromous species (reviewed by Ward et al. 1994; Carvalho and Hauser 1998; Ward 2002). However, earlier studies already had suggested that some marine species might have considerably lower effective population sizes (N_e; the number of individuals in a population who contribute offspring to the next generation) than census sizes (e.g. Hedgecock 1994; Hedgecock et al. 1994) due to marked variance in reproductive success. Recent data (Hauser et al. 2002; Turner et al. 2002; Hutchinson et al. 2003) now indicate that the effective population sizes of populations of marine fish species - especially those characterised by high fecundity and high larval mortality - are typically 10^2 to 10^6 orders of magnitude smaller than census population sizes (reviewed in Bekkevold et al. 2004).

However, reduction in N_e to the point where inbreeding would become significant is not indicated by the levels of variability observed at the microsatellite loci analysed, with between 23 and 107 alleles at individual loci. The same applies to assortative mating, even though assortative mating probably plays an important role in redfish evolution, at least on species level (Love et al. 2002).

A deficiency of heterozygotes will also occur, if two species or populations interbreed to some extent (thus producing hybrids) but mate assortatively within species or populations (pre-mating isolation) - or if selection acts against hybrids/heterozygotes (post-mating isolation: Dowling and Moore 1985; Campton 1987). Selection against heterozygotes can result from outbreeding depression: Mating between species or between populations that have been separated (e.g. geographically) in the past can potentially produce outbred progeny with lower fitness (Hoarau et al. 2002b). Species-level hybridisation and selection against heterozygotes has been shown in two species of *Notropis* (Cyprinidae; Dowling and Moore 1985). Population-level hybridisation and selection against heterozygotes has been discussed in plaice (Hoarau et al. 2002b).

However, only few individuals with admixed ancestry have been observed in the present data set and there is no indication of broad-scale interspecific hybridisation (see paragraph 5.5). Nevertheless, intraspecific hybridisation between populations cannot be ruled out. Therefore, selection and assortative mating could also affect the Hardy-Weinberg equilibrium in North Atlantic *Sebastes* populations.

However, Wahlund effects (and, in case of SEB37 and SEB46, null alleles) are more likely to account for the heterozygotes deficits observed in this study, due to the strong genetic heterogeneity observed in the samples. As a Wahlund effect generally inflates heterozygotes deficits, it was not possible to distinguish the effects of intra- and interspecific hybridisation and assortative mating from the Wahlund effect. Nevertheless, it cannot be ruled out that a combination of selective factors, assortative mating and Wahlund effects has caused the

observed deficits. The importance of the different factors remains hypothetical, given the limited knowledge on the life history and ecology of the taxa, and the difficulty of rigorously quantifying selective effects in these species, as in any marine pelagic organism (Roques et al. 2001).

The heterozygote excess observed at locus SEB30 in the *S. mentella* samples MEEGr01, MEWGr01 and MESIrm01 and at locus SEB45 in sample MESEIc01 could be explained by the occurrence of positive selection at these loci (heterozygote advantage). However, as heterozygote excess was observed in only few samples and did not remain significant after Bonferroni correction, selection is relatively unlikely. Also outbreeding/hybridisation is an unlikely explanation, as only very few loci and samples were involved. These few cases of heterozygotes excess are no more than expected by chance and the significance disappeared when the Bonferroni correction was applied.

5.3 Species structure of North Atlantic redfish – a case of cryptic speciation in the genus Sebastes?

The most salient finding of this study was that redfish individuals, pre-classified morphologically into the four currently known North Atlantic *Sebastes* species, clustered into five genetically distinct groups. This was revealed by both methods of individual-based analysis applied in the present study, the factorial correspondence analysis (FCA) and the Bayesian cluster analysis using *STRUCTURE* (Pritchard et al. 2000), respectively. These clusters were genetically so distinct from each other that it was possible to assign about 94% (average over all samples) of the individuals to one of the five clusters using Bayesian cluster analysis, without using prior information about the sample of origin (see also paragraph 5.4).
Except a few individuals that were most likely pre-classified to the wrong species on board and a few individuals that revealed admixed genotypes, most of the individuals sampled and pre-classified as *S. viviparus*, *S. fasciatus* and *S. mentella* grouped into separate genetic clusters, in agreement with the morphological species identification (see paragraph 5.4). Furthermore, individuals collected as *S. marinus* clustered into two genetically distinct clusters. Samples from Norway, the Faroe Islands East, SW-Iceland and the Flemish Cap as well as the sample of so-called "giant *S. marinus*" consisted almost exclusively of individuals assigned to the first cluster. As the individuals assigned to this cluster were found in samples from the whole sampling area covered in this study (including samples from the Flemish Cap and Norway) and species discrimination on the Flemish Cap and in Norwegian waters using morphological and morphometrical data yielded relatively good results in previous studies (Saborido-Rey 1994; Garabana 2005) - in contrast to areas such as Greenland, the Faroe Islands and Iceland - it can be assumed that these individuals are representatives of *S. marinus*.

Samples from the Faroe Islands West, SE-Iceland and Greenland revealed two groups of individuals: One group clustered with the individuals from Norway, the Faroe Islands East, SW-Iceland and the Flemish Cap and "giant *S. marinus*" (= the "*S. marinus* cluster"), whereas the remaining individuals clustered separately from the "*S. marinus* cluster" - as well as from all other clusters. Due to its genetic distinctness, this cluster was designated "*Sebastes sp.* cluster" (see paragraph 4.3.1). Even though many samples of *S. marinus* collected in the area Faroe Islands/Iceland/Greenland consisted of individuals of both clusters (the "*S. marinus* cluster" and the "*Sebastes sp.* cluster"), the level of genetic differentiation was sufficiently high, so that most of the individuals collected in these areas (about 95%) could be assigned to one of the two genetic clusters with individual admixture proportions equal to or larger than 0.80. This was even possible, when only six microsatellite loci were used in the Bayesian cluster analysis.

Furthermore, the observed genetic pattern was stable over time, as the two clusters were observed in samples of *S. marinus* collected in subsequent years.

The high degree of genetic differentiation among the five clusters was reflected in the results of the corresponding AMOVA (Table 2.21), which revealed that a very high amount of the total genetic variance was imputable to differences between the genetic clusters. This variance component was much higher than the variance among samples within clusters (in the order of 6 to 10 times, depending on the parameter used: variance in allele sizes or allele frequencies, respectively). The results of the locus-by-locus AMOVA (Table 2.22) showed that almost all loci were responsible for the observed genetic variation between clusters. The strong concordance among loci and the strong partitioning of genetic variance among the clusters indicate a robust signal of genetic differentiation between the genetic clusters.

Furthermore, the comparison of the results of the two AMOVAs analysing the (i) variance in allele frequencies among species and within species before the two groups of individuals found in the *S. marinus* samples were separated (see Table 2.20) and (ii) among and within the genetic clusters after the two groups of individuals found in the *S. marinus* samples were sorted into the "*S. marinus* cluster" and the "*Sebastes sp.* cluster" (see Table 2.21) further supported the heterogeneity of the original species samples. The results revealed that when these individuals were not separated, a high part of the total variance in allele frequencies and an even higher part of the total variance in allele sizes was found between samples within species. The variance in allele sizes was even higher between samples within species than between species. This indicates that the species samples (pre-classified into species on board using morphological traits and in a few cases hemoglobin analysis[1]; see also paragraph 3.1) were extremely heterogeneous.

Also the temporal instability observed in the AMOVA of the *S. marinus* data set (Table 2.23) - which revealed that a significant part of the total variance in allele frequencies was due

[1] Hemoglobin analyses were performed by Torild Johansen, University of Bergen (UIB), Norway.

to differences between temporal samples of *S. marinus* from the same location - was very likely caused by the heterogeneity of the *S. marinus* samples, as the observed temporal instability disappeared, when the individuals assigned to the "*Sebastes sp.* cluster" were excluded from the AMOVA (see Table 2.24).

The existence of two genetic clusters explains also the high F_{ST} values between several samples of *S. marinus*, when individuals assigned to the "*Sebastes sp.* cluster" were not excluded from the analyses (maximum F_{ST} value: 0.074, see Table 2.16). Some of these F_{ST} values were comparable to some of the interspecific F_{ST} values observed in this study (see Table 2.14 and 2.15 and discussion below). Furthermore, tests for isolation by distance were not significant when all individuals pre-classified as *S. marinus* were included in the analysis, but significant, when individuals assigned to the "*Sebastes sp.* cluster" were excluded (see paragraph 4.5), further supporting the existence of two genetically distinct groups in the *S. marinus* samples.

The heterogeneity of the unsorted *S. marinus* samples was also illustrated by the MDS plots, which revealed a rather diffuse clustering of samples of *S. marinus*, when all individuals were included in the analysis - in contrast to the distinct clustering of samples after individuals were sorted according to the results of the assignment tests (Figure 2.12 and 2.13).

The distinctness of the five genetic clusters was reflected in highly significant genetic differences with moderate to high pairwise F_{ST} and R_{ST} values between the clusters (Table 2.15). The highly significant F_{ST} values (between 0.079 and 0.169) observed between the "*S. viviparus* cluster", the "*S. fasciatus* cluster", the "*S. marinus* cluster" and the "*S. mentella* cluster" were very similar to F_{ST} values between *S. viviparus*, *S. fasciatus*, *S. marinus* and *S. mentella* reported in other studies (Roques et al. 1999a: F_{ST} values between species: 0.088 and 0.199). The same applies, at least partly, to the pairwise R_{ST} values (ranging from 0.079 to 0.391), which were in some cases approximately twice as high as the corresponding F_{ST}, and comparable to the values observed by Roques et al. (1999a; R_{ST} values between 0.120 and 0.486). Slatkin (1995) suggested that increases in R_{ST} above F_{ST} might be due to extra mutations over and above the variation generated by genetic drift (Adams and Hutchings 1999). This is the case when the levels of genetic exchange are low and/or when coalescence time is long and the accumulation of mutations contributed substantially to differentiation among taxa. In some pairwise comparisons, however, the R_{ST} values were only slightly higher or equal to the corresponding F_{ST} values. F_{ST} is similar to R_{ST}, when the differentiation is roughly independent of the mutation process and genetic drift is the driving force, that is, with large migration rates and/or recent times of divergence among populations or species (Slatkin 1995; reviewed in Estoup and Angers 1998). Furthermore, in three cases, F_{ST} was larger than R_{ST}. The contradicting pattern of the R_{ST} estimators might be explained by the higher variance of R_{ST} (Balloux and Goudet 2002). The different performances of R_{ST} and F_{ST} values in detecting genetic differentiation in North Atlantic redfish on species as well as on population level will be discussed in detail in paragraph 5.7.

Nevertheless, both F- and R-statistics indicated that the five genetic clusters detected in the Bayesian cluster analysis were genetically distinct, with up to ten times higher pairwise F_{ST} and R_{ST} values than those observed on the intraspecific level. In particular, the genetic differences (F_{ST} as well as R_{ST}) found between the "*Sebastes sp.* cluster" and the "*S. viviparus* cluster", the "*S. fasciatus* cluster", the "*S. marinus* cluster" and the "*S. mentella* cluster", respectively, were of the same order of magnitude as the genetic differences observed between the latter four clusters.

The observation that the samples of *S. marinus* collected on the shelves of Greenland, Iceland and the Faroe Islands consist of two genetically distinct groups of individuals - separated by similar levels of genetic differentiation as the four known North Atlantic *Sebastes* species - points to the existence of a fifth cryptic *Sebastes* species in these areas. The high degree of genetic differentiation provides strong evidence of species separation, in particular as the observed genetic clusters occurred sympatrically (Ward and Grewe 1995). Cryptic species (also called sibling or hidden species; Knowlton 1993; 2000) are species or groups, which appear morphologically identical or very similar, but show genetic and reproductive differentiation and are reproductively isolated. They are well known to occur in many marine fish species (Avise 1994; Knowlton 1993 and 2000; see also Ward and Grewe 1995 and references therein).

The existence of cryptic speciation in North Atlantic *Sebastes* is strongly supported by the results of the mitochondrial ND3 gene sequence analysis (presented in Chapter 1 of this thesis). Several individuals, collected off Iceland and Greenland, including 15 individuals assigned to the "*Sebastes sp.* cluster" in the microsatellite analysis, were analysed with both methods (mitochondrial ND3 gene sequence analysis and microsatellite analysis). All of these 15 individuals (sample MASEIc01, MADB01 and MAEGr01; see Table 1.7 in Chapter 1) revealed a genetically distinct mitochondrial haplotype, MA1, which differed from the other haplotypes observed in *S. marinus*. Therefore, the two genetically distinct groups of individuals in *S. marinus* were not only observed using microsatellite analysis, genetic differences were also observed on the level of mitochondrial DNA. Similar results were obtained from the AFLP analysis (amplified fragment length polymorphisms), which also supported the existence of cryptic speciation in North Atlantic *Sebastes* (see Chapter 3).

Alternate hypotheses, which could explain the observed high level of genetic divergence, are selection or hybridisation. However, the fact that mitochondrial DNA as well as several nuclear loci (microsatellite and AFLP loci) revealed high levels of genetic divergence indicates that the observed pattern has genome-wide effects. Therefore, selection is an unlikely explanation for the observed genetic differences, as selection is likely to differentiate a minority of functional coding loci (reviewed in Carvalho and Hauser 1995).

Also introgressive hybridisation was not likely a substantial cause of the observed genetic divergence. Even though there is some indication of ancient introgression of mitochondrial DNA in North Atlantic *Sebastes*, e.g. between *S. fasciatus* and *S. marinus* on the Flemish Cap

(see Chapter 1), recent broad-scale introgressive hybridisation was not observed in the current data set, as most of the individuals did not reveal admixed genotypes (see paragraph 5.5). Therefore, cryptic speciation remains the most likely explanation for the observed genetic structure.

The occurrence of cryptic species and species complexes is not at all uncommon for the genus *Sebastes*: In Pacific and South Atlantic *Sebastes* species, cryptic species have often been the centre of taxonomic and nomenclature revisions (Rocha-Olivares et al. 1999c). MtDNA analyses have detected at least one cryptic rockfish species in the southern hemisphere (Rocha-Olivares et al. 1999b; 1999c). Genetic differences between three morphotypes of *S. inermis* from the Northwest Pacific (based on AFLP and mitochondrial DNA analyses) indicated that the morphotypes represent different species, which have evolved recently (Kai et al. 2002a). Furthermore, Kai and Nakabo (2004) described a new *Sebastes* species (*S. kiyomatsui*) from the Pacific coast of southern Japan, which had previously been treated as one of two colour morphotypes of *S. scythropus*. Genetic analyses (based on mitochondrial DNA), however, revealed that the two colour morphotypes are indeed different species (Kai et al. 2002b). Recently, cryptic speciation has been described in *Sebastes aleutianus* (Gharrett et al. 2005).

The fact that no indication of cryptic speciation in North Atlantic *Sebastes* was found in previous genetic studies can be explained by several factors. Most of the previous genetic studies have been dominated by studies on the electrophoretic mobility patterns of proteins, such as hemoglobin and allozyme analyses (Payne and Ni 1982; McGlade et al. 1983; Trottier et al. 1989; Nedreaas and Nævdal 1989; 1991a; 1991b; Rubec et al. 1991; Nedreaas et al. 1994). These methods did not detect the same degree of genetic differentiation as observed in the present study, most likely because of the limited resolution of the markers used. No single protein marker system was sensitive enough to discriminate between all four species in all life stages. Except the microsatellite studies by Roques et al. (1999a; 1999b; 2001), also all previous DNA-based genetic studies (Bentzen et al. 1998; Desrosiers et al. 1999; Sundt and Johansen 1998) revealed very low levels of genetic differentiation between the species. Furthermore, all previous studies either did not include all North Atlantic *Sebastes* species or samples from a very limited geographic area were analysed - making it impossible to detect the genetic variation of North Atlantic *Sebastes* over its range of distribution (e.g. McGlade et al. 1983; Desrosiers et al. 1999; Roques et al. 1999a). Roques et al. (1999a) for example, who found that microsatellite data from eight loci could be used to assign individual fish to the correct species more precisely than other genetic techniques, analysed samples from a very limited geographic area (e.g. only *S. marinus* samples from Norway were analysed). In contrast, the present study covered a larger geographic area and included specifically samples from Greenland, Iceland and the Faroe Islands.

However, the results of the hemoglobin analyses by Nedreaas and Nævdal (1989; 1991a) and Nedreaas et al. (1994) - performed on a larger geographical scale - support the results of the

present study with respect to the genetic heterogeneity of *S. marinus* in the areas off Iceland and Greenland. The results of the analyses by Nedreaas and Nævdal (1989; 1991a) and Nedreaas et al. (1994) revealed that the hemoglobin pattern of *S. marinus* is more complicated in *S. marinus* occurring on the shelves of Iceland and Greenland, in contrast to *S. marinus* occurring in Norwegian waters. According to Nedreaas et al. (1994), this pattern could be explained by hybridisation between *S. marinus* and *S. mentella* or by variation within *S. marinus*. As the results of the present study did not indicate extensive hybridisation between *S. marinus* and *S. mentella* (see paragraph 5.5), the more complicated hemoglobin pattern found by Nedreaas and Nævdal (1989; 1991a) and Nedreaas et al. (1994) could indeed support the hypothesis of cryptic speciation in *S. marinus*.

5.4 Precision of species assignment and admixed genotypes

In congruence with the genetic distinctness of the five clusters revealed by the FCA and the Bayesian cluster analysis, as well as with the high levels of genetic differentiation revealed by the F-statistics, a high percentage of individuals could be assigned to these five clusters. On average, it was possible to assign about 94% of the individuals (between 90% and 100%, depending on the sample) to one of the five clusters inferred by the Bayesian cluster analysis, with individual proportions of membership $Q \geq 0.80$ in one of the five cluster. Assignment success was highest in *S. viviparus* (98% assigned individuals), *S. fasciatus* (99% assigned individuals) and *S. marinus* (95% assigned individuals). The individuals of *S. mentella* were assigned a bit less accurately on average, as only 90% of the individuals could be assigned to the "*S. mentella* cluster" with individual proportions of membership $Q \geq 0.80$. Nevertheless, the high percentage of assigned individuals shows that that assignment power was high.

A comparison of the assignment results using microsatellite data and the pre-classification on board - using external morphological and morphometrical traits and in some cases hemoglobin analysis (in case of the samples from Greenland 2001; see paragraph 3.1) - revealed that in case of the samples of *S. viviparus*, *S. fasciatus* and *S. mentella*, the results of the assignment tests were in congruence with the pre-classification on board. Only very few individuals of the samples of *S. viviparus*, *S. fasciatus* and *S. mentella*, respectively, were assigned to another genetic cluster as the major part of the individuals of these samples (Table 2.8). The same applies - at least partly - to the samples of *S. marinus* (with respect to the fact that only very few individuals sampled as *S. marinus* were assigned to the "*S. viviparus* cluster", the "*S. fasciatus* cluster" or the "*S. mentella* cluster" - morphologically the two genetically distinct clusters in the samples of *S. marinus* were not detectable).

The few individuals, assigned to another species using Bayesian cluster analysis, were most likely misclassified or mislabelled on board the research vessels. This is supported by the results of independent analyses performed by specialists, who applied morphological and

morphometrical analyses[1], hemoglobin[2], MDH[2] (malate dehydrogenase; only performed on the Flemish Cap samples) or AFLP analysis (see Chapter 3). Even though not all of the presumably misclassified individuals could be analysed with all these methods, in most cases at least one additional method was applied to classify these individuals.

The four individuals of *S. marinus* collected on the Flemish Cap, which were assigned to the "*S. fasciatus* cluster", were most likely misclassified on board, supported by the results of AFLP (see Chapter 3) and MDH analyses (Johansen pers. comm. 2003). The same applies to the four individuals of the *S. fasciatus* sample, which were assigned to the "*S. marinus* cluster" (two individuals) and the "*S. mentella* cluster" (two individuals). Also in this case, the assignment results were supported by MDH analysis, indicating that these four individuals did not show the *S. fasciatus* specific MDH pattern (Johansen pers. comm. 2003). Furthermore, AFLP analyses, performed on three of the four individuals, revealed the same results as the microsatellite analysis (see Chapter 3).

Of the 13 individuals sampled as *S. marinus* but assigned to the "*S. mentella* cluster", nine were subsequently classified as *S. mentella*, based on external morphology and hemoglobin analysis[1,2] (Garabana pers. comm. 2003; Johansen pers. comm. 2003). For the remaining four individuals, no other data were available, but it is likely that also these individuals were misclassified or mislabelled on board the research vessels.

The single *S. marinus* individual assigned to the "*S. viviparus* cluster" represented a mislabelled sample of *S. viviparus* (Bárðarsson[3] pers. comm. 2002).

Also the individuals collected as *S. mentella*, but assigned to the "*S. marinus* cluster" (two individuals) and to the "*Sebastes sp.* cluster" (nine individuals) were most likely misclassified, as the results of the assignment tests based on AFLP data (see Chapter 3) supported the result of the assignment tests based on microsatellite data in almost all cases. Only a single individual from the sample MESIrm01 - assigned to the "*Sebaste*s *sp.* cluster" using microsatellite analysis - was assigned to *S. mentella* in the AFLP analysis. This could be due to the fact that in the AFLP analysis too few individuals of the "*Sebaste*s *sp.* cluster" were analysed (see Chapter 3 for details). However, except for this single individual, the result of the AFLP analysis supported the results of the microsatellite analysis, indicating that the eight individuals collected as *S. mentella* off East Greenland 2001, which were assigned to the "*Sebastes sp.* cluster" in the microsatellite analysis, belonged to a genetically distinct group. Furthermore, the results of the AFLP analysis support the results of the microsatellite analysis, indicating that the two individuals of the *S. mentella* sample from Iceland (MESEIc01) belonged to the species *S. marinus*.

In summary, assignment success over all samples and species was high and allowed to detect previously misclassified individuals. The microsatellite-based assignment was supported by

[1] Morphometrical analyses were performed by Dolores Garabana, Institute of Marine Research, Vigo, Spain.

[2] Hemoglobin and MDH analyses were performed by Torild Johansen, University of Bergen (UIB), Norway.

[3] Birkir Bárðarsson, Marine Research Institute, Division of Marine Stocks and Fisheries, Reykjavík, Iceland.

other genetic and non-genetic methods such as morphological and morphometrical analyses[1], hemoglobin[2], MDH[2] and AFLP analyses. Except the AFLP analysis (see Chapter 3), however, these methods could not detect the potential cryptic species in the samples of *S. marinus* (Garabana pers. comm. 2003; Johansen pers. comm. 2003). Furthermore, neither morphological and morphometrical analyses, nor hemoglobin or MDH analyses are applicable to younger life stages (Ni 1981; Power and Ni 1982; Nedreaas and Nævdal. 1989; 1991a; Nedreaas 1994).

The results of the present study therefore demonstrate the potential of microsatellite analysis for fast and reliable species identification in the North Atlantic *Sebastes* species complex. This supports the findings by Roques et al. (1999a), suggesting that microsatellites can be used to assign individual redfish to the correct species more precisely than other previously applied techniques.

On average, about 6% of the individuals could not be assigned to one of the genetic clusters/species, as they revealed admixed genotypes.

There are two possible explanations for the occurrence of admixed genotypes: Interspecific hybridisation or statistical uncertainty.

There are some factors that support the latter explanation. First, the individual 90% confidence intervals of *Q* were often large for individuals with admixed genotypes (see Table 2.7; other data not shown). Furthermore, several (even though not all) of the admixed genotypes were composed of rather unlikely genotype combinations. For example, a high percentage of admixed genotypes/unassigned individuals were observed in the two samples from the central Irminger Sea (MECIrm01 and MECIrm03). Several of these admixed genotypes revealed a proportion of membership in the "*S. viviparus* cluster" or in the "*S. fasciatus* cluster". Hybridisation between *S. mentella* and *S. viviparus* as well as between *S. mentella* and *S. fasciatus* in the central Irminger Sea is rather unlikely, as neither *S. viviparus* nor *S. fasciatus* are known to occur in this area. Furthermore, *S. viviparus*, *S. fasciatus* and Irminger Sea *S. mentella* do not share similar depth preferences, habitats, mating areas and mating periods (Barsukov et al. 1985; Whitehead et al. 1986; Magnússon and Magnússon 1995). Similar unlikely genotype compositions were found in other samples of *S. marinus* and *S. mentella* (e.g. admixed genotypes found in samples of *S. marinus* revealed in some cases a proportion of membership in the "*S. viviparus* cluster" and the "*S. fasciatus* cluster", or compositions of three and more clusters in different proportions; see Appendix IV).

Therefore, statistical uncertainty due to low sample sizes or - more likely - too few loci (as SEB37 and SEB46 were removed from the data set due to null alleles) represents the most likely explanation for the occurrence of admixed genotypes. A higher number of microsatellite loci may increase statistical power for assigning individuals to species and

[1] Morphometrical analyses were performed by Dolores Garabana, Institute of Marine Research, Vigo, Spain.

[2] Hemoglobin and MDH analyses were performed by Torild Johansen, University of Bergen (UIB), Norway.

reduce the number of admixed genotypes. Simulation studies have shown that with increasing genetic differentiation, there was a larger gain in assignment power by increasing the number of loci rather than sample sizes (Cornuet et al. 1999; Pritchard et al. 2000; reviewed in Hansen et al. 2001).

Therefore, the high assignment success will probably further improve by increasing the number of loci rather than the sample sizes in routine genetic assignment analyses of North Atlantic *Sebastes* samples.

Even though it is likely that a part of the unassigned individuals/admixed genotypes resulted from statistical uncertainty, it cannot be ruled out that a few individuals with admixed genotypes were interspecific hybrids.

The results of the model-based cluster analyses investigating specifically hybridisation in areas where several *Sebastes* species occur sympatrically, such as the Flemish Cap and the shelves of Iceland, Greenland and the Faroe Islands, are discussed in paragraph 5.5.

5.5 Interspecific hybridisation

In general, a lower number of unassigned individuals was observed when only the samples relevant to the evaluation of (i) hybridisation between *S. mentella*, *S. marinus* and the newly identified cryptic *Sebastes* species in the areas off the Faroe Islands, Iceland, East and West Greenland and (ii) hybridisation between *S. mentella*, *S. marinus* and *S. fasciatus* on the Flemish Cap were included in the model-based cluster analyses. This is in concordance with the recommendations of Hansen et al. (2001), who proposed that to minimise statistical uncertainty and maximise assignment power when applying assignment methods, it is advisable to include only samples relevant to a specific problem.

The results of the model-based cluster analysis investigating hybridisation between *S. marinus* and *S. mentella* and the newly identified cryptic species in the areas off the Faroe Islands, Iceland, East and West Greenland revealed that only 4.12% of the individuals displayed admixed genotypes, whereby most of these individuals revealed a higher proportion of membership in one of the three clusters inferred by the *STRUCTURE* analysis. Only very few individuals with intermediate genotype compositions were observed (see Figure 2.6). Also in this analysis, the 90% confidence intervals of the individual admixture proportions of the admixed genotypes were often high, indicating statistical uncertainty regarding the ancestry of these individuals (data not shown).

However, even if we assume that there was no statistical uncertainty and that all of the admixed/unassigned individuals were hybrids, the small number of admixed genotypes observed indicates that there is no extensive hybridisation between the three genetic groups occurring sympatrically in the areas off the Faroe Islands, Iceland and East and West Greenland.

Furthermore, the results of the microsatellite analysis did not indicate that "giant" *S. marinus* are the product of hybridisation between *S. mentella* and *S. marinus* - as proposed by Altukhov and Nefyodov (1968) - as none of the 12 individuals of "giant" *S. marinus* revealed admixed genotypes. However, even though the results of the present microsatellite analysis did not indicate that "giant" *S. marinus* are hybrids, the results of the analysis of the mitochondrial ND 3 gene of "giant" *S. marinus* indicated that some of these individuals could be the product of an ancient introgressive hybridisation event between *S. mentella* and *S. marinus* (see Chapter 1). Nevertheless, the results of the microsatellite analysis indicate that introgression did not occur on the level of nuclear DNA - or has long been diluted by repeated backcrossing of female hybrids with male *S. marinus* (see Chapter 1 for a detailed discussion)[1].

Previously, it has often not been possible to rule out hybridisation as a reason for the overlap observed in morphological and morphometrical characters and in electrophoretic mobility patterns of proteins used to discriminate *S. marinus* and *S. mentella* (Altukhov and Nefyodov 1968; Nævdal 1978; Nedreaas and Nævdal 1991a; Nedreaas et al. 1994). In Icelandic waters and off East and West Greenland, Nævdal (1978), Nedreaas and Nævdal (1991a) and Nedreaas et al. (1994) found a variation in the hemoglobin pattern of *S. marinus*. According to Nedreaas et al. (1994), this pattern could be explained by hybridisation between *S. marinus* and *S. mentella*, or variation within *S. marinus* (see also paragraph 5.3). However, the genetic basis of these variations has never been revealed and methodological difficulties hampered the distinction between the two hypotheses (reviewed in Johansen 2003). The results of the present study favour the latter hypothesis, indicating that the variation in the hemoglobin pattern found in samples of *S. marinus* in these areas is not the result of hybridisation between *S. mentella* and *S. marinus*, but represents rather intraspecific variation or, more likely, interspecific variation - as the results of the microsatellite analysis point to the existence of a fifth cryptic *Sebastes* species off the coast of Greenland, Iceland and the Faroe Islands (see paragraph 5.3).

Also the results of the model-based cluster analysis of samples of *S. mentella*, *S. marinus* and *S. fasciatus* from the Flemish Cap revealed only few unassigned individuals with admixed genotypes (4.58% = 12 individuals). These few individuals with admixed genotypes were genetically mainly composed of the "*S. mentella* cluster" and the "*S. marinus* cluster". Only three individuals with admixed genotypes showed a small proportion of membership in the "*S. fasciatus* cluster" (see Figure 2.7).

Therefore, the results of the microsatellite analysis do not indicate broad-scale hybridisation between *S. mentella*, *S. marinus* and *S. fasciatus* on the Flemish Cap.

[1] Also the hypothesis that "giant" *S. marinus* represent a genetically isolated group or separate species of North Atlantic *Sebastes*, as proposed by Johansen et al. (2000a), could not be supported by the microsatellite analysis, as all individuals analysed clustered with the "*S. marinus* cluster". For further discussion of the taxonomic status of "giant" *S. marinus* see paragraph 5.3.2 in Chapter 1.

These results are somewhat in contrast to the results of the analysis of mitochondrial ND3 gene sequences (see Chapter 1). All twenty individuals of *S. marinus* from the Flemish Cap (sequenced in the study presented in Chapter 1) revealed ND3 haplotypes that were specific for *S. fasciatus*, indicating an incorporation of *S. fasciatus* haplotypes into *S. marinus* through introgressive hybridisation events. However, as the microsatellite analysis (in which these twenty introgressed individuals were included) did not reveal individuals with intermediate allelic compositions composed of the "*S. fasciatus* cluster" and the "*S. marinus* cluster", recent hybridisation events between *S. fasciatus* and *S. marinus* are unlikely. This suggests that the introgressive hybridisation events indicated by the results of the ND3 gene analysis are ancient.

The generally large 90% confidence intervals of the individual admixture proportions of the few admixed individuals observed in this study make it difficult to determine whether these individuals are hybrids. The observation of admixed genotypes could also result from a lack of assignment power. An analysis of a higher number of microsatellite loci is required to exclude the effect of statistical uncertainty. However, if we assume that all individuals with admixed genotypes were real hybrids, hybridisation between the North Atlantic *Sebastes* species is both (i) restricted (as only few genotypes were admixed) and (ii) extremely bimodal (as most of the individuals were genetically similar to one or the other parental genotype - with extremely few intermediates).

A similar pattern of bimodal introgressive hybridisation on nuclear level was observed in *S. mentella* and *S. fasciatus* in the Gulf of St Lawrence (Roques et al. 2001). Roques et al. (2001) used microsatellite analysis to investigate hybridisation between *S. mentella* and *S. fasciatus* in the Northwest Atlantic and found evidence of broad-scale introgressive hybridisation between the two species in the Gulf of St. Lawrence and adjacent areas. This hybridisation resulted in the co-existence of two introgressed populations of *S. mentella* and *S. fasciatus*, which remained reproductively isolated in the main zone of sympatry, despite some gene flow between them. Most fish were relatively similar to one or the other parental species, rather than being of intermediate genetic composition. The lack of intermediate genotypes is in congruence with the findings of the present study. This is suggestive of rare hybridisation and/or of numerous backcrosses between F1 hybrids and one or the other parental species (Roques et al. 2001).

Therefore, the results of the present microsatellite analysis indicate that there is only restricted, if any, hybridisation between *S. marinus*, *S. mentella* and the newly identified cryptic species of North Atlantic *Sebastes* in the areas off the Faroe Islands, Iceland, East and West Greenland, as well as between *S. fasciatus*, *S. marinus* and *S. mentella* on the Flemish Cap. Nevertheless, ancient introgressive hybridisation events seem to have occurred, e.g. on the Flemish Cap (see Chapter 1).

In this context it has to be noted that the broad-scale introgressive hybridisation between *S. mentella* and *S. fasciatus* in the Gulf of St. Lawrence and adjacent waters found by Roques et al. (2001) is one of the few rare examples of hybridisation in the genus *Sebastes*. In the Pacific, where a high number of *Sebastes* species exists - which are often closely related and live in close proximity - one might expect that hybridisation would be fairly common (Love et al. 2002). However, this does not appear to be the case. With only a few exceptions (e.g. Seeb 1998), hybrids have rarely been seen or at least recognised (reviewed in Love et al. 2002).
The extent of hybridisation in *Sebastes* is probably limited by factors such as the mode of reproduction, differing mating periods and mating areas and the relative abundance of the parental species (Desrosiers et al. 1999).
In particular, all *Sebastes* species are viviparous with internal fertilisation (reviewed in Wourms 1991) and internal fertilisation may permit more restrictive mate selection than occurs in fishes with external fertilisation.
However, identifying the factors that restrict hybridisation between the North Atlantic *Sebastes* species is difficult, as our knowledge about biology and behaviour of these species is still limited.

5.6 Population structure

5.6.1 General patterns of population structure in North Atlantic redfish

The general patterns of population structure observed in *S. mentella*, *S. marinus* and the newly identified cryptic species of North Atlantic *Sebastes* were congruent with the weak genetic structuring usually reported for marine organisms (Palumbi 1994). The apparent genetic homogeneity of many marine species is thought commonly to be due to two factors that minimise accumulation of genetic differentiation among populations: Large effective population sizes that limit genetic drift (DeWoody and Avise 2000) and life history characteristics that favour dispersal (e.g. planktonic eggs and larvae, juvenile and adult vagility) in continuous dynamic oceanic environments (Ward et al. 1994; Waples 1998).
However, even though the levels of genetic differentiation observed in *S. mentella* and *S. marinus* were subtle, they were nevertheless significant and comparable to those found in other marine fishes (Ward et al. 1994; Waples 1998). The significant F_{ST} values observed between samples of *S. marinus* (significant F_{ST} values between 0.009 and 0.045) and *S. mentella* (significant F_{ST} values between 0.007 and 0.018) are comparable to those found between populations of cod (*Gadus morhua*; Bentzen et al. 1996; Nielsen et al. 2001; 2003), European hake (*Merluccius merluccius*; Lundy et al. 1999) and sea bass (*Dicentrarchus labrax*; Naciri et al. 1999). Also Pacific *Sebastes* species revealed similar levels of genetic differentiation between populations: A significant F_{ST} value of 0.04 was observed between two populations of *S. alutus* (Withler et al. 2001), significant F_{ST} values between populations

of *S. caurinus* ranged from 0.007 to 0.087 (Buonaccorsi et al. 2002) and significant F_{ST} values between populations of *S. macdonaldi* ranged from 0.007 to 0.010 (Rocha-Olivares et al. 2003).
Differential patterns of population structure, however, were observed between *S. mentella* and *S. marinus*, which may be partly explained by their respective life histories. The fact that *S. mentella* revealed lower population structure than *S. marinus* might be explained by the ecology of the species. In general, lower population structure is expected for pelagic and widely distributed marine species - compared to benthic ones or those more geographically restricted (Avise et al. 1987; reviewed in Roques et al. 2001). *S. mentella* is more widely distributed than any other North Atlantic *Sebastes* species and is more frequently encountered in the pelagic zone, while *S. marinus* occurs mainly on the continental slopes, closer to the coast (Magnússon and Magnússon 1995; Roques et al. 2001). Roques et al. (2001) observed a similar pattern in the West Atlantic, where *S. mentella* revealed lower levels of genetic differentiation than *S. fasciatus*, which is more associated with the epibenthic zone of shallower waters. Also studies investigating the population structure of *Sebastes* species along the Pacific coast showed this general trend. Nearshore rockfishes demonstrated more population substructuring than offshore species (Cope 2004). Though mechanisms behind the pattern are unknown, oceanographic conditions and habitat discontinuities may be of particular relevance (Rigions and Nachman 2001; see also Cope 2004).

Most of the pairwise R_{ST} values between samples of *S. mentella* and *S. marinus*, respectively, were not significant, in contrast to their corresponding F_{ST} estimates. R_{ST} can be less accurate at reflecting population differentiation than F_{ST} due to different factors (Gaggiotti et al. 1999). As all other statistical analyses (except the pairwise R_{ST} values) revealed similar results, the pairwise R_{ST} values will not be further discussed in the next three paragraphs. A detailed discussion of the different performances of pairwise F_{ST} and R_{ST} values on species as well as on population level is given in paragraph 5.7.

5.6.2 Population structure of S. marinus

Overall, the results of the analysis of microsatellite variation in the "*S. marinus* cluster" revealed indication of population structure. According to the results of the model-based cluster analysis (*STRUCTURE*; Pritchard et al. 2000), the "*S. marinus* cluster" consisted of two to three populations, respectively, depending on the model of allele frequencies applied (independent or correlated). However, the results of the *STRUCTURE* analysis applying the "independent allele model" were inconclusive, as all samples were rather admixed and none of the samples showed a high proportion of membership in one of the two inferred clusters (see Table 2.11). The results of the *STRUCTURE* analysis applying the more sensitive "correlated allele frequency model", which often achieves better performance by improving

clustering for closely related populations (Falush et al. 2003), revealed a more detailed picture. The *S. marinus* individuals from the Flemish Cap were mainly grouped in a separate cluster (cluster I; red segments in Figure 2.9), indicating that they belong to the same population. The same applies to the individuals from Greenland (cluster III; blue segments in Figure 2.9) and the individuals from Iceland, the Faroe Islands and Norway (cluster II; green segments in Figure 2.9). Furthermore, the plot of the individual admixture proportions (see Figure 2.9) indicates that some samples represent an admixture of individuals from different populations. Several individuals of the samples from SE-Iceland showed a relatively high membership in the cluster to which many individuals from Greenland were assigned (cluster III: blue segments in Figure 2.9), and a few individuals of the sample from the Dohrn Bank showed a relatively high membership in the cluster to which many individuals from Iceland, the Faroe Islands and Norway were assigned (cluster II: green segments in Figure 2.9).

However, the levels of individual admixture were high and only about half of the individuals could be assigned to one of the three clusters/populations inferred by *STRUCTURE* (Table 2.12), most likely because the genetic signal in the data set was too weak. As the levels of genetic differentiation (F_{ST}) observed between *S. marinus* samples were relatively small - even though highly significant in many cases (see discussion below) - the statistical power was presumably not sufficient to assign a high percentage of individuals to populations.

Nevertheless, the observation of genetic structure in the "*S. marinus* cluster" was supported by the results of the AMOVA (Table 2.24), indicating that a small but significant part of the total genetic variance in allele frequencies as well as in allele sizes was due to genetic differentiation between samples from different locations. The results of the locus-by-locus AMOVA (Table 2.25) showed that almost all loci were responsible for the observed genetic variation between samples from different locations. The strong concordance among loci indicates a robust signal of genetic differentiation. Furthermore, the results of the AMOVA indicate temporal stability, as the part of the total genetic variance in allele frequencies as well as in allele sizes that was due to genetic differentiation between temporal samples from the same location was not significant and only one locus revealed significant differences in allele frequencies between temporal samples.

Also the finding that several pairwise comparisons (F_{ST} values) were significantly different from zero allows the rejection of the null hypothesis of genetic homogeneity within the "*S. marinus* cluster". The levels of genetic differentiation were small, with significant F_{ST} values ranging from 0.009 to 0.045, but comparable to those reported for populations of many other marine fish species (Ward et al. 1994; see paragraph 5.6.1). In general, the samples from Greenland and from the Flemish Cap, respectively, were genetically different from the samples from Iceland, the Faroe Islands and Norway, whereas the samples from the latter three areas were genetically very similar. In particular, the samples from Norway, the Faroe Islands (East) and SW-Iceland revealed no indication of genetic differentiation, but differed

significantly from the samples from the Flemish Cap and Greenland. In contrast, the samples from NE-Iceland and SE-Iceland revealed some heterogeneity, as they did not show significant differences to the samples from Greenland in several pairwise comparisons. Also the sample from the Dohrn Bank off East Greenland showed some heterogeneity, as it revealed significant genetic differences to the two samples from East Greenland.

This heterogeneous pattern is reflected in the MDS plot (Figure 2.14). The two samples from the Flemish Cap were placed in a distinct cluster separately from all other samples, which supported the genetic distinctness of these samples. The samples from SW-Iceland and the Faroe Islands clustered together, and the sample from Norway clustered closely with these three samples, supporting that these samples were genetically similar. The same applies to the samples from Greenland, which grouped together. The samples from NE-Iceland, SE-Iceland and East Greenland/Dohrn Bank were placed in-between these major groups. The intermediate position of these samples, the fact that samples that were geographically more closely located (e.g. the samples from SE-Iceland and the Faroe Islands), but clustered separately in the MDSA, and the heterogeneity in pairwise genetic differences (F_{ST}) indicate that these "intermediate" samples represent a mixture of individuals from different populations. As discussed above, also the results of the model-based cluster analysis strongly support the hypothesis that these samples are a mixture of individuals from different populations.

Therefore, the results of the present microsatellite analysis indicate that there are at least three genetically distinct populations within *S. marinus* in the geographic area covered in this study, represented by (i) the samples from the Flemish Cap, (ii) the samples from Greenland and (iii) the samples from Norway, the Faroe Islands and Iceland, whereby the latter two populations apparently overlap in some areas.

The genetic distinctness of Flemish Cap *S. marinus* probably results from the special characteristics of the Flemish Cap, which is a relatively isolated bank separated from the continental shelf by the about 1,100 m deep Flemish Pass. It is generally assumed that this physical barrier probably contributed to the isolation of some of the Flemish Cap fish populations, including populations of *S. marinus*, *S. fasciatus* and *S. mentella* (Templeman 1976; Ávila de Melo et al. 1998). Furthermore, the large geographic distances separating the Flemish Cap from other sampling locations covered in the present study probably act as barriers to gene flow. This is supported by the significant, though perhaps low, correlation between geographic and genetic distances observed at the spatial scale of this study.

The finding that the *S. marinus* samples from Norway, the Faroe Islands and Iceland belong to the same population, which was indicated by no or extremely weak genetic differences, is inconsistent with the current management strategy. *S. marinus* in the Northeast Arctic (ICES subareas I and II) and *S. marinus* on the continental slopes of Greenland, Iceland and the Faroe Islands (ICES Divisions V, VI, XII, XIV and NAFO area 1) are treated as two different

stocks (ICES 1998a; Saborido-Rey et al. 2005). The concept that *S. marinus* on the continental slopes of Greenland, Iceland and the Faroe Islands belong to the same population, separated from *S. marinus* from Norway and the Barents Sea, is mainly based on the location of the currently known areas of larval extrusion and the nursery areas. The main area of larval extrusion of the *S. marinus* stock on the continental slopes of Greenland, Iceland and the Faroe Islands is believed to be southwest of Iceland, although larval extrusion has also been observed south of the Faroe Islands in some years (ICES 1983; ICES 1998a). The nursery grounds of this stock are assumed to be located on the continental slopes of Iceland and Greenland (ICES 1998a). As no extrusion of larvae has been detected on the shelf of Greenland, it has been assumed that the juveniles present around Greenland and on the Icelandic shelf belong to the same stock (ICES 1998a). In contrast, areas of larval extrusion of the Northeast Arctic stock are found on the Norwegian shelf, outside the Lofoten/Vesterålen, the Halten Bank and Storegga outside Møre. The Barents Sea is believed to be the nursery area of this stock (ICES 1998a).

Even though the location of the areas of larval extrusion could indicate that *S. marinus* occurring on the continental slopes of Greenland, Iceland and Faroe Islands belong to the same population - separated from the Northeast Arctic population - already previous analyses of allozymes raised some doubts about the current stock division and pointed to a closer relationship between Norwegian and Faroe waters (Nedreaas and Nævdal 1991a). Furthermore, gene flow caused by migration between Norway, the Faroe Islands and Iceland cannot be ruled out. According to Reinert and Lastein (1992), *S. marinus* occurring in Faroe waters could derive from the Northeast Arctic stock, as the Faroe Islands are situated on the submerged ridge between Scotland and Greenland via Iceland. This is further supported by the fact that adult *S. marinus* are found all over the Norwegian shelf and in the North Sea and have also been observed off the coast of Brittany (ICES 1998a), revealing the high dispersal potential of the species.

However, there are alternative hypotheses that could explain the observed genetic homogeneity. *S. marinus* from Norway and the area Iceland/Faroe Islands, respectively, could represent two reproductively isolated populations, but population separation could not be detected due to a recent common ancestry of the two populations (= genetic drift has not yet produced large amounts of differentiation) or due to high levels of gene flow. One of the major limitations of molecular genetics in detecting population separation in marine fish populations is its sensitivity to rates of gene flow. A small number of migrants per generation is sufficient to homogenise the gene pool, at least empirically given the relatively small sample sizes generally employed in population genetic studies (Carvalho and Hauser 1995). Therefore, it is possible that there are two effectively self-recruiting populations of *S. marinus* in Norwegian waters and the waters off Iceland and the Faroe Islands, respectively, which are not genetically distinct yet. This is supported by the existence of distinct areas of larval extrusion and nursery areas of the Northeast Arctic stock (ICES 1998a).

Furthermore, even though the analysis of temporal samples from Greenland and Iceland revealed temporal stability, it cannot be ruled out that the lack of genetic differentiation among the samples from Norway, Iceland and the Faroe Islands is an ephemeral phenomenon, as only one sample from Norway was analysed. The sample from Norway could have represented a group of migrants from the Faroe/Icelandic *S. marinus* population. More temporal replicates are needed to verify whether the observed genetic pattern is stable over time.
The same applies to the finding that the samples from Greenland and Iceland represent two genetically distinct populations of *S. marinus.* Even though the results are strongly supported by the fact that the structure was observed in two subsequent years - indicating temporal stability - it is crucial to investigate the population structure over a longer time period. Two subsequent years represent a small time window, in particular considering the longevity of the species and its high potential for dispersal.

5.6.3 Population structure of the newly identified potential cryptic species of North Atlantic Sebastes

The model-based clustering analysis (*STRUCTURE*; Pritchard et al. 2000) revealed that the "*Sebastes sp.* cluster" (cluster III), which most likely represents a cryptic species of North Atlantic *Sebastes* (see 5.3), did not show sub-structuring. The most likely number of populations present in the data set consisting solely of individuals assigned to the "*Sebastes sp.* cluster" was one, when the model assuming independent allele frequencies was applied. When the model assuming correlated allele frequencies was applied, the most likely number of populations was ten. However, the visual inspection of estimated values of Q revealed a uniform distribution with $Q \sim 1/K$ (= 1/10) such that all individuals were deemed equally derived from each of the ten populations inferred by *STRUCTURE*. Therefore, according to the guidelines by Pritchard et al. (2000) and Pritchard and Wen (2003), it can be concluded that a value of $K = 1$ (= one population) was the most appropriate choice of model for the data (see also paragraph 3.8.3). The lack of genetic structure is also reflected in the pairwise F_{ST} and R_{ST} values, which were all not significantly different from zero. Therefore, the null hypothesis of no population structure in the newly identified cryptic species of North Atlantic *Sebastes* cannot be rejected with the current data set.

5.6.4 Population structure of S. mentella

Overall, the microsatellite analysis revealed also weak but significant genetic structure in *S. mentella*. However, the levels of genetic differentiation were smaller than those observed in *S. marinus*. In fact, the levels of genetic differentiation were so small that the model-based cluster analysis (*STRUCTURE*; Pritchard et al. 2000) indicated that the most likely number of

populations (*K*) in the *S. mentella* data set was one, regardless of the model of allele frequencies applied (see Table 2.13).

The weak genetic structure is also reflected in the results of the AMOVA, which revealed that the part of the total variation (in allele frequencies as well as in allele sizes; Table 2.26) that was due to differences between samples from different locations was not significant. Overall, only two out of six loci contributed significantly to the variation in allele frequencies among samples from different locations (Table 2.27). This indicates that the overall signal of genetic differentiation is weak and not as robust as in *S. marinus*. Furthermore, a small but still significant part of the total genetic variation in allele frequencies was distributed among temporal samples, indicating temporal instability of the observed genetic structure - although the temporal instability was caused by only one locus.

Even though the overall pattern of genetic variation indicates weak, if any, genetic structure in *S. mentella*, the pairwise F_{ST} values revealed significant genetic differentiation between several samples. Small but significant F_{ST} values (between 0.007 and 0.018) indicate genetic distinctness of the samples from the Flemish Cap, the central Irminger Sea and the samples from Greenland, Iceland, the southern Irminger Sea, NAFO area 1F and NAFO area 2J, whereby the latter five areas were genetically relatively homogeneous. The lack of genetic differentiation between the samples from East Greenland (collected in the year 2001), West Greenland, the southern Irminger Sea, NAFO area 1F and NAFO area 2J suggests that these samples belong to the same *S. mentella* population. However, the sample from East Greenland collected in the year 2000 was genetically different from the sample collected in the same area one year later, indicating temporal instability of the genetic structure. In contrast, the samples collected in the central Irminger Sea in two different years were genetically similar, indicating temporal stability. The genetic pattern revealed by the pairwise F_{ST} values was reflected in the results of the MDSA (Figure 2.15), in which the samples from the central Irminger Sea clustered separately from the other samples. The same applies to the Flemish Cap sample, as well as to the samples from East Greenland (collected in 2001), West Greenland, the southern Irminger Sea and the two samples collected in the NAFO areas 1F and 2J.

However, in some cases the F_{ST} values were small and not significant after sequential Bonferroni correction (e.g. between the sample from the central Irminger Sea and East Greenland collected in the year 2001, or between the sample from the Flemish Cap and the samples from NAFO area 1F and West Greenland). This may reflect high levels of gene flow (migration, larval drift) or recent population separation. High levels of gene flow or recent population separation are also indicated by the extremely small genetic differences between the sample from SE-Iceland and almost all other samples, which were in most cases not significant. The fact that the sample from SE-Iceland differed little from the other samples, but clustered with the sample from East Greenland collected in the year 2000 (see Figure 2.15) - which was significantly different from all other samples - suggests a rather heterogeneous genetic pattern. This heterogeneity can be explained by an admixture/overlap

of population components: The samples from SE-Iceland and East Greenland (collected in the year 2000) could represent mixed samples of individuals from different populations. The distinct clustering of these two samples in the MDSA could furthermore indicate that some of the individuals of these samples are migrants from other areas/populations not covered in the present study. This is supported by the highly significant genetic differences observed between the two samples from East Greenland collected in subsequent years.

In summary, the results of the present study support the existence of at least three genetically distinct populations of *S. mentella* in the geographical area covered in this study, represented by (i) the sample from the Flemish Cap, (ii) the samples from the central Irminger Sea and (iii) the samples from East Greenland (collected in the year 2001), West Greenland, the southern Irminger Sea and the two samples collected in the NAFO areas 1F and 2J. However, it cannot be ruled out that also migrants from other areas/populations were present in the samples, as the geographical area covered in this study was limited.

Even though the observed genetic distinctness of Flemish Cap *S. mentella* is not as prominent as in *S. marinus* and there is indication of gene flow or a relatively recent isolation of the population, the results indicate that *S. mentella* on the Flemish Cap represents a distinct population. This is in congruence with the general assumption that the redfish populations of the Flemish Cap are relatively isolated, e.g. from those of the Grand Banks (reviewed in Templeman 1976; see also paragraph 5.6.2). The genetic distinctness of Flemish Cap fish populations has also been described for cod (*Gadus morhua*). Bentzen et al. (1996) found that Flemish Cap cod were genetically more divergent from geographically proximate Grand Bank samples, suggesting that the deep ocean trenches act as barriers to gene flow.

In contrast, the observed genetic structure of *S. mentella* in the Irminger Sea and on the continental slopes of Greenland and Iceland is more difficult to explain. The population structure of *S. mentella* in these areas has been discussed extensively in many studies and there is still a strong controversy about the three alternate hypotheses that have been put forward to describe this structure (ICES 1998a; ICES 2002; ICES 2004b; Saborido-Rey et al. 2005; see also Figure 1 in the GENERAL INTRODUCTION chapter). The single stock hypothesis suggests that the mature individuals of a single stock of *S. mentella* - distributed from the Faroe Islands to Greenland - segregate according to age/size, whereas the two stock hypotheses suggests that the demersal *S. mentella* living on the shelves (so-called deep-sea *S. mentella*) and the *S. mentella* occurring in deeper pelagic waters in the Irminger Sea (so-called pelagic deep-sea *S. mentella*) constitute one stock unit which is separated from the oceanic *S. mentella* occurring pelagic in the Irminger Sea. According to the three stock hypotheses each of these three types constitutes a distinct stock.

Earlier genetic studies based on allozyme or hemoglobin variation revealed low levels of genetic polymorphism within *S. mentella* across the North Atlantic. No differences were observed from six locations in the Irminger Sea, based on six allozyme loci (Dushchenko 1987), nor among locations off the Faroe Islands, the eastern Davis Street and West

Greenland (Nedreaas and Nævdal 1991a; Nedreaas et al. 1994). However, Johansen et al. (2000b) studied allozyme and hemoglobin variation in *S. mentella* and found indication that *S. mentella* occurring in deeper pelagic waters in the Irminger Sea (so-called pelagic deep-sea *S. mentella*) and the so-called oceanic *S. mentella* occurring pelagic in shallower waters in the Irminger Sea do not share a common gene pool. In contrast, a recently published work by Roques et al. (2002), based on microsatellite analysis, revealed genetic homogeneity and corresponding low genetic differences within a large "panoceanic" population of *S. mentella* over a distance of 6,000 km, from the Grand Banks and Labrador to the Faroe Islands.

The pattern of genetic structure in *S. mentella* observed in the present study is at least partly congruent with the results of the microsatellite-based population genetic study by Roques et al. (2002). Similar to the present study, Roques et al. (2002) found no or only weak genetic differences between samples from Iceland, the Irminger Sea, East and West Greenland. In contrast to the findings by Roques et al. (2002), however, the *S. mentella* samples from the central Irminger Sea analysed in the present study revealed significant levels of genetic differentiation from other samples collected in these areas. The different outcomes of the two studies can be explained by the higher number of samples from the Irminger Sea analysed in the present study. Furthermore, the results of the present study indicated temporal instability, most likely due to a different distribution of population units in different years, which could also explain the differences to the study by Roques et al. (2002), in which *S. mentella* samples from 1995 to 1998 were analysed.

Therefore, some evidence of population structuring at closer geographical scales as reported for allozymes (Johansen et al. 2000b) is suggested by the present microsatellite data, although the results were not always consistent across sampling years and were not a significant source of variation in the AMOVA.

However, even though the results of the present study do not support the one stock hypothesis, they also do not support a separation of demersal *S. mentella* living on the shelves and so-called oceanic *S. mentella* occurring pelagic in shallower waters in the Irminger Sea, as proposed by the two and the three stock hypothesis. The samples from the shelf of Greenland[1] and the samples of pelagic *S. mentella* collected in the southern Irminger Sea and the NAFO areas 1F and 2J in shallower depths (between 159 and 425 m depth; see Appendix I) did not show significant genetic differences. Only the samples of pelagic *S. mentella* from the central Irminger Sea collected in greater depths (between 640 and 700 m depth; see Appendix I) revealed significant genetic differences to the other samples. Therefore, a strict separation of pelagic and demersal population components of *S. mentella* is not indicated by the results of the present study.

In contrast to the results presented here, the authors of a recent publication (Saborido-Rey et al. 2005) come to the conclusion that the general ecology (i.e. life cycle, spawning behaviour,

[1] except the sample from East Greenland collected in the year 2000 (MEEGr00)

distribution of different life stages, drift and migration patterns) of *S. mentella* inhabiting the Irminger Sea and adjacent waters points to the existence of one single stock. According to Saborido-Rey et al. (2005), in particular the existence of a single and wide area of larval release southwest of Iceland and above the Reykjanes Ridge, with a wide depth distribution and with a single spawning peak from April to June, supports the existence of a single stock. Also the distribution of juveniles and adults, indicating that the major nursery grounds are located on the shelf of Greenland and that individuals migrate to the feeding grounds (located around the Faroe Islands, Iceland and in the Irminger Sea) as they grow and mature, support the single stock hypothesis (Saborido-Rey et al. 2005). Saborido-Rey et al. (2005) conclude that the genetic differences observed for example by Johansen et al (2000b) could be explained by selection or differences between different year-classes (age dependency).

Selection could indeed have caused the genetic differences found by Johansen et al. (2000b), as mainly two allozyme loci (isocitrate dehydrogenase IDHP and malic enzyme MEP) caused the observed genetic differences and the levels of divergence were many times higher than for supposed neutral microsatellites. This could indicate the genetic differences observed by Johansen et al. (2000b) are determined by locus-specific selection and adaptation to certain environmental conditions - such as depth.

As microsatellites are non-coding and thus likely to be selectively neutral (Jarne and Lagoda 1996), selection most likely did not cause the genetic structure observed in the present data set.

However, it cannot be entirely ruled out that the genetic differences between certain samples observed in the present study may reflect genetic differentiation between year-classes rather than geographic isolation. A comparison of the total lengths of all *S. mentella* specimens used in the microsatellite analysis revealed that the two genetically distinct samples from the central Irminger Sea collected in greater depths consisted of much larger individuals (median of the total lengths: approximately 40 cm) as for example the samples from Greenland (median approximately 25 cm; see Appendix IIb) and most likely represent different age groups. Even though aging of *S. mentella* in the Irminger Sea has been conducted only sparsely, studies have shown that the spatial and vertical distribution of the age composition fully reflects the dynamics of the length composition in redfish (Saborido-Rey et al. 2005). Therefore, the observed genetic structure may reflect genetic differences between year-classes (no geographic isolation) derived from a sweepstakes chance effect in which a few adult spawners successfully contribute offspring each season (Hedgecock 1994; Geiger et al. 1997). Evidence for sweepstakes recruitment in marine fishes comes from the observation of genetic differences among individual schools of anchovies (e.g. Hedgecock et al. 1994) but also *S. mentella* (Altukhov et al. 1990). Genetic differentiation without a clear geographic pattern is expected under this effect (Matala et al. 2004a and 2004b). In fact, no significant correlation between geographic and genetic distances was observed in *S. mentella* at the spatial scale of this study (from SE-Iceland to the Flemish Cap) - in contrast to *S. marinus*, which revealed a

significant correlation between genetic and geographic distance. Similar observations of genetic structure without clear geographic pattern are reported for shortraker redfish, *Sebastes borealis*, and bocaccio, *Sebastes paucispinis* (Matala et al. 2004a and 2004b). Also Matala et al. (2004a and 2004b) discuss the possibility of genetic structure reflecting differences between year-classes.

Cohort effects would require that there is a non-trivial level of genetic differentiation among cohorts, requiring relatively little overlap (gene flow) among generations and small effective population sizes N_e (Nielsen 2004). In long-lived species such as *S. mentella*, with a relatively high reproductive potential (due to the viviparous mode of reproduction, offspring are more likely to survive) a population can be sustained by few, abundant and temporally closed, year-classes, which reproduce during many years (Saborido-Rey et al. 2005). According to Saborido-Rey et al. (2005) the single population of *S. mentella* occurring in the Irminger Sea and adjacent areas could be sustained if these few year-classes produce a series of successful new year-classes, even if there is recruitment failure during a long period of time. This is supported by the fact that after a severe depletion of demersal *S. mentella* off East Greenland during the 1980's and early 1990's, a sharp increase in abundance was observed in 1995-1997, mainly due to an individual strong year-class of young fish, which later migrated into the central Irminger Sea recruiting to the adult stock there (Stransky 2000).

These observations could therefore support a cohort effect. Furthermore, even though the consensus population size of *S. mentella* is high (about 3×10^9 individuals on the shelves of Greenland; ICES 2004c), recent data indicate that N_e in marine fishes can be 10^2 to 10^6 orders of magnitude smaller than census population sizes (Hauser et al. 2002; Turner et al. 2002; Hutchinson et al. 2003; reviewed in Bekkevold et al. 2004).

However, small effective population sizes are not indicated by the levels of variability observed at the microsatellite loci analysed, with between 23 and 107 alleles at individual loci. Therefore, to validate cohort differentiation as a possible explanation for genetic differentiation, more investigations, e.g. of the qualitative estimation of the required effective population sizes and the overlap among cohorts, are needed (Nielsen 2004).

Distinguishing between the two hypothesis (i) population structure and (ii) cohort effect is beyond the scope of this study. Both hypotheses may be consistent with the finding of this study, namely significant genetic differentiation between *S. mentella* collected in greater depths in the central Irminger Sea and *S. mentella* from the shelf of Greenland and in shallower waters in the southern Irminger Sea. However, even though a sweepstakes chance effect cannot be entirely discounted, it is relatively unlikely that the observed genetic differentiation is caused by differences between cohorts alone, due to the high levels of genetic diversity and the high number of alleles observed in the microsatellite analysis.

Furthermore, there is strong indication of temporal instability of the observed genetic structure. This suggests that geographic patterns could result from movement of population components across putative stock boundaries. The number of temporal genetic samples in this

study is limited - in particular considering the longevity of the species and its high potential for dispersal. Therefore, more temporal replicates are needed before definite conclusions about the population structure of *S. mentella* in the Irminger Sea and adjacent waters can be drawn.

5.6.5 Factors complicating the analysis of the population structure of North Atlantic redfish species

One of the main drawbacks that hamper investigations of the population structure (and also of the species structure) of North Atlantic redfish is the general lack of information about the biology of the species. In particular, our knowledge about the mating process is limited. Even though it is generally believed that *S. marinus* as well as *S. mentella* mate in autumn (Saborido-Rey et al. 2005) and that the copulation areas of *S.marinus* and *S. mentella* are located along the coast of East Greenland, around Iceland and the Faroe Islands (Reinert and Lastein 1992), we do not know exactly where and when mating takes place.
Therefore, samples of North Atlantic *Sebastes* could potentially be mixed feeding aggregations of fish from various populations. This was also observed in the present analysis of the population structure of *S. marinus*. However, only *S. marinus* revealed levels of genetic differentiation that were - at least to some extent - sufficient for applying individual-based assignment methods on the population level.
S. mentella revealed lower levels of genetic differentiation, which is not atypical for marine fish species. However, due to the low levels of genetic differentiation, it was not possible to assign individuals of *S. mentella* to populations - and to determine unequivocal whether samples of *S. mentella* consisted of individuals from different populations. This makes it extremely difficult to interpret the results of the genetic analyses, in particular as the analysis of samples collected in different years indicated temporal instability.
This problem can be overcome if samples could be collected as close (temporally and/or spatially) to the actual mating as possible - as it can be assumed that populations that are reproductively isolated are separated (temporally and/or spatially) during mating. However, this would require a more detailed knowledge about the mating time and the mating area of the North Atlantic redfish species.

5.7 A comparison of F- and R-statistics and their efficiency in detecting inter- and intraspecific genetic differentiation in North Atlantic redfish

In the present study, a comparison of F- and R-statistics revealed that pairwise F_{ST} and R_{ST} values often differed in a pronounced manner. This observation is not atypical and has been reported in several studies (reviewed in Lugon-Moulin et al. 1999).

As already mentioned in paragraph 3.8.4, F- and R-statistics are based on different assumptions about the mutation model underlying microsatellite variation. Whereas Wright's (1951) pairwise F_{ST} (Weir and Cockerham 1984) is based on the infinite alleles model (IAM; Kimura and Crow 1964) of microsatellite mutation, Slatkin's (1995) R_{ST}, assumes a stepwise mutation model (SMM; Kimura and Ohta 1978). Under the IAM, a mutation involves any number of tandem repeats and always results in an allelic state not previously encountered in the population (Estoup and Angers 1998). Therefore, the comparison between any two different alleles bears the same information. Slatkin (1995) showed that if the mutational process follows a SMM, the expected squared difference between allele sizes is a linear function of the expected coalescence time of the alleles compared. The larger the difference, the higher the number of mutation events is expected to have occurred since common ancestry. Consequently, alleles of very different sizes will be more distantly related than alleles of similar sizes. There is thus a "memory" of past mutation events. Slatkin (1995) showed in a simulation study that R_{ST} performs better than F_{ST}, when microsatellites mutate in a stepwise fashion.

Comparing pairwise F_{ST} and the corresponding R_{ST} values can provide valuable insights into the main causes of population differentiation, i.e. drift and mutation (reviewed by Hardy et al. 2003). According to Slatkin (1995), the IAM based F_{ST} is expected to be similar to the SMM based R_{ST}, when genetic drift is the dominant process in creating differentiation and mutation plays little role, that is, with large migration rates and/or recent times of divergence among populations. In case of low migration rates and/or longer time since divergence among populations, R_{ST} is expected to be larger than F_{ST}, and F_{ST} will tend to underestimate the true level of genetic differentiation (Slatkin 1995; reviewed in Estoup and Angers 1998).

As mentioned in paragraph 5.3, genetic differentiation among species/genetic clusters measured with R_{ST} was in most cases much higher than that measured with F_{ST}[1], indicating that the species are genetically well differentiated and there was ample time for mutations to accumulate since common ancestry (Slatkin 1995).

These results are concordant with the empirical findings reviewed by Lugon-Moulin et al. (1999). These authors reviewed empirical studies based on microsatellite data and found that for moderate to strong genetic differentiation, R_{ST} seemed to be a better estimator. This was further supported by a simulation study by Balloux and Goudet (2002).

In a few cases, however, the R_{ST} values observed in the present study were similar or only slightly higher than F_{ST}[2]. Furthermore, even opposite results were obtained: A few R_{ST} values were smaller than the corresponding F_{ST} values[3]. Similar values for R_{ST} and F_{ST} could indicate a more recent ancestry of the species. R_{ST} values that are smaller than the

[1] e.g. between the "*S. viviparus* cluster" and the "*S. fasciatus* cluster"; between the "*S. marinus* cluster" and the "*Sebastes sp.* cluster" or between the "*S. mentella* cluster" and the "*S. marinus* cluster"

[2] between the "*S. viviparus* cluster" and the "*S. mentella* cluster", between the "*S. fasciatus* cluster" and the "*Sebastes sp.* cluster" and between the "*S. mentella* cluster" and the "*Sebastes sp.* cluster"

[3] between the "*S. viviparus* cluster" and the "*Sebastes sp.* cluster" and between the "*S. marinus* cluster" and the "*S. fasciatus* cluster"

corresponding F_{ST} values are, however, not expected: As differentiation decreases, the expectations of F_{ST} and R_{ST} converge, but F_{ST} never becomes larger than R_{ST} - even under mutation models others than a stepwise mutation model - unless allele size differences are inversely related to evolutionary distances between alleles (Balloux and Goudet 2002). The smaller R_{ST} values might be explained by the higher variance of R_{ST}. F_{ST} estimates may outperform their R_{ST} counterparts, as variances of estimates of R_{ST} can be strongly affected by unequal sample sizes - even under the SMM and even though the inclusion of allele size is expected to make R_{ST} more informative (Gaggiotti et al. 1999). F_{ST} is generally characterised by a lower sampling variance (Slatkin 1995; Gaggiotti et al. 1999; Balloux and Goudet 2002). Indeed, in the present study the range of values covered by R_{ST} (0.079-0.391) was larger than that covered by F_{ST} (0.079-0.169). This and the unequal sample sizes (e.g. 42 individuals of *S. viviparus* and 643 individuals of *S. marinus*, see Table 2.15) most likely explain why the interspecific R_{ST} values were in some cases similar or even smaller than F_{ST}. This is further supported by the fact that the interspecific F_{ST} values observed in the present study were similar to the F_{ST} values reported by Roques et al. (1999a), in contrast to the R_{ST} values, which were always higher than F_{ST} in their study.

Also on the intraspecific level - in case of *S. mentella* as well as *S. marinus* - estimates of F_{ST} and R_{ST} differed. Even though R_{ST} values were in some cases higher than F_{ST}, they were not significantly different from zero in most pairwise comparisons, whereas the corresponding F_{ST} indicated significant genetic differentiation. Pairwise R_{ST} values were therefore more conservative than the corresponding F_{ST}, despite the assertion that F_{ST} underestimates the degree of genetic differentiation among populations (Slatkin 1995). This can be explained by two factors: The large variance of R_{ST} discussed above and/or small levels of genetic differentiation. According to Lugon-Moulin et al. (1999) and Balloux and Goudet (2002), R_{ST} better reflects population differentiation in populations characterised by very low genetic exchange, whereas F_{ST} gives better estimates in cases of low genetic differentiation (due to high levels of gene flow or a recent separation of the populations). In particular, when the numbers of loci scored or the sample sizes are small, F_{ST} may outperform R_{ST} (Gaggiotti et al. 1999). These factors, high variance, relatively few loci (only six and seven, respectively) and unequal sample sizes - in combination with low levels of genetic differentiation - may explain why R_{ST} did not detect the same degree of differentiation as F_{ST}.

Another explanation could be that the mutation pattern of the microsatellites used in the present study does not follow the SMM. It is still unclear how departures from the stepwise mutation model (SMM) affect measures of genetic differentiation based on allele size differences such as R_{ST} (Estoup and Angers 1998).

Because of these issues, some authors (O'Connell and Wright 1997) recommended that analyses of microsatellite variability in fish should be based on conventional F_{ST} statistics (Ward 2002).

6. General conclusions and perspectives

The present study - based on the analysis of highly polymorphic microsatellite markers - has provided new insights into the inter- and intraspecific genetic structure of North Atlantic *Sebastes*.

The most relevant finding of this study was the observation of a group of individuals that was genetically distinct from the four currently known species of North Atlantic *Sebastes*, pointing to the existence of a fifth cryptic *Sebastes* species on the continental slopes of Greenland, Iceland and the Faroe Islands. The fact that these individuals occurred mainly in samples of *S. marinus* indicates that they are - even though genetically very distinct - morphologically similar or indistinguishable from *S. marinus*, at least with the morphological and morphometrical characters currently used to discriminate the North Atlantic *Sebastes* species.

The addition of a new cryptic species to the North Atlantic *Sebastes* species complex underscores the fact that we still know very little about these species - and emphasises the need for further investigations to understand the distribution and the biology of this species complex.

The present work also illustrates the relevance of individual-based analyses such as FCA and in particular Bayesian cluster analysis in studies of species complexes consisting of morphologically similar species. By using these statistical methods, it was possible to identify the major patterns of genetic structure in the data set without any a priori assumptions. This is of particular interest in cases where species identification using morphological traits is difficult and often ambiguous. Furthermore, it is often unknown whether the classification of individuals into species based on phenotypic criteria represents a natural assignment in genetic terms (Pritchard et al. 2000). By applying Bayesian cluster analysis, it was not only possible to infer the relevant genetic units (species) in the data set, but also to assign single individuals to these genetic units and therefore to identify specimens that were misclassified in the field.

Without the application of FCA and Bayesian cluster analysis, the two genetic groups found in the samples of *S. marinus* would have remained undetected. This would have had serious consequences for the interpretation of patterns of population structure obtained by classical population genetic parameters estimating genetic differentiation between samples (e.g. pairwise F_{ST} and AMOVA) - as demonstrated by the different results obtained from the analyses performed with and without separating the two groups of individuals in the samples of *S. marinus*.

The indication of cryptic speciation has also wide-ranging implications for the management and conservation of North Atlantic *Sebastes*, as it reveals the existence of biological diversity that will go unnoticed without the careful examination of genetic variation. Therefore, a clear understanding of species differentiation and distribution of North Atlantic *Sebastes* - essential

for a proper understanding of their biology and for an adequate management of the exploited stocks - requires a reliable technique for accurate species identification (Gascon 2003).

The potential usefulness and simplicity of redfish species discrimination using microsatellite analysis was already demonstrated by Roques et al (1999a). The results of the present study showed that on average 94% of the individuals could be successfully assigned to species using only six microsatellite loci. Furthermore, by using Bayesian cluster analysis it was possible to rule out extensive broad-scale hybridisation between *S. marinus* and *S. mentella* as a reason for the wide overlap in morphological and morphometrical characters frequently observed in certain areas, such as the continental slopes of Greenland. A higher number of loci will probably further improve assignment success. Microsatellite analysis could therefore be considered as reliable standard method for species identification, e.g. on scientific surveys in areas were several species occur sympatric and species identification is particularly difficult, e.g. on the continental slopes of Greenland, Iceland and the Faroe Islands. Moreover, microsatellite markers can easily be used at any life history stage, which can be of interest in recruitment and larval ecology (Roques et al. 1999a).

The present microsatellite study has also provided new insights into the population structure of *Sebastes* in the central North Atlantic. It represents the first analysis of the population structure of *S. marinus* using polymorphic microsatellite markers. The results indicate that the *S. marinus* samples from Norway, Iceland and the Faroe Islands could belong to the same population, or that high levels of gene flow or recent ancestry connect *S. marinus* populations occurring in these areas. Furthermore, there is strong indication that the samples from Greenland and Iceland represent two genetically distinct populations of *S. marinus*. These findings could have implications on the management of the stocks, as *S. marinus* in the area Greenland, Iceland and the Faroe Islands is currently managed as one stock, whereas *S. marinus* on the Norwegian shelf and in the Barents Sea is managed separately (ICES 1998a). However, if *S. marinus* from Norway, Iceland and the Faroe Islands indeed belong to the same population, then this population contributes to the fishery of both regions. Overharvesting in either Norwegian or Icelandic waters may have reciprocal consequences for recruitment and reproduction. Similarly, the existence of a distinct population of *S. marinus* on the shelf of Greenland would require a separate management, in particular as the abundance of *S. marinus* off East and West Greenland has been extremely low in recent years and the fishable stock of *S. marinus* in these areas is already considered depleted (ICES 2004c).

However, as the observed patterns of population structure could represent an ephemeral phenomenon, a more detailed temporal and geographic sampling is warranted to draw further conclusions, in particular as *S. marinus* is long-lived and the distribution of potential populations in these areas can vary temporally.

Also the results of the population genetic analysis of *S. mentella* do not fully support the current management strategy, which treats demersal *S. mentella* living on the shelves and the

so-called oceanic *S. mentella* occurring pelagic in shallower waters of the Irminger Sea as different stocks (ICES 2004c). In contrast, the results of the present study indicate that the samples from the shelf of Greenland as well as the samples collected in pelagic waters in the southern Irminger Sea and in the NAFO areas 1F and 2J in the year 2001 could belong to the same population, which is genetically different from the samples collected in the central Irminger Sea in greater depths.

However, the observed genetic differences were very small and temporal and spatial sampling was too limited to draw final conclusions about the population structure of *S. mentella*. In particular, there is strong indication of temporal instability of the observed structure, suggesting that geographic patterns could result from movement of population components across putative stock boundaries.

Therefore, a broader survey including more sampling sites, additional loci and in particular temporal replicates over a longer period of time will most likely elucidate the population structure of *S. mentella*.

To validate cohort differentiation as a possible explanation for genetic differentiation, more investigations, e.g. genetic analysis of groups arranged by year classes, investigations of the overlap among cohorts, etc., are needed (Nielsen 2004). Furthermore, in future studies one should try to sample as close to the actual mating area/time as possible, to prevent that samples consist of mixed feeding aggregations of fish from various populations (Nielsen 2004).

Appendix I

Summary statistics for redfish samples collected for microsatellite analysis showing species names, morphological types, sampling locations, position of the sampling locations, sampling years, sampling months, average sampling depths, station numbers, coordinating institutes/nations, cruise identification numbers, sample codes and sample sizes (*n*).

			Position (mean)				Depth
Species	**Type**	**Location**	**Longitude**	**Latitude**	**Year**	**Month**	**(m)**
S. marinus		Norway	682030N	0111550E	2001	October	195
S. marinus		Faroe Islands East	623543N	0047361W	2002	September	408
S. marinus		Faroe Islands West	623543N	0090547W	2002	Oktober	497
S. marinus		NE-Iceland	654996N	0120303W	2000	October	177
S. marinus		SE-Iceland	640146N	0131350W	2000	October	300
S. marinus		SE-Iceland	640143N	0131375W	2001	October	222
S. marinus		SW-Iceland	641640N	0232400W	2000	June	102
S. marinus		SW-Iceland	641305N	0233920W	2001	October	117
S. marinus		SW-Iceland	641640N	0232420W	2001	October	117
S. marinus		Greenland East/Dohrn Bank	653457N	0300652W	2001	October	368
S. marinus		Greenland East/Dohrn Bank	652891N	0302102W	2001	October	378
S. marinus		Greenland East	642913N	0350859W	2000	October	-
S. marinus		Greenland East	635350N	0362876W	2000	October	-
S. marinus		Greenland East	641929N	0360459W	2001	October	256
S. marinus		Greenland East	641664N	0355754W	2001	October	278
S. marinus		Greenland West	600914N	0470932W	2001	November	379
S. marinus		Greenland West	601852N	0480720W	2001	November	261
S. marinus		Flemish Cap	468257N	0459862W	2001	July	304
S. marinus		Flemish Cap	468703N	0464920W	2001	July	450
S. marinus		Flemish Cap	472712N	0457550W	2001	July	293
S. marinus		Flemish Cap	476687N	0448737W	2001	July	240
S. marinus		Flemish Cap	476315N	0449568W	2001	July	247
S. marinus		Flemish Cap	469425N	0447085W	2001	July	148
S. marinus		Flemish Cap	472643N	0447045W	2001	July	172
S. marinus		Flemish Cap	474342N	0445983W	2001	July	208
S. marinus		Flemish Cap	476645N	0441683W	2001	July	431
S. marinus		Flemish Cap	474452N	0443257W	2001	July	269
S. marinus		Flemish Cap	473270N	0443403W	2001	July	251
S. marinus		Flemish Cap	472988N	0442058W	2001	July	291
S. marinus		Flemish Cap	476732N	0441738W	2001	July	433
S. marinus		Flemish Cap	470663N	0439083W	2001	July	478
S. marinus		Flemish Cap	471648N	0439082W	2001	July	477
S. marinus	giant	Irminger Sea	604400N	0282000W	1996	August	732
S. marinus	giant	Irminger Sea	604400N	0282000W	1996	August	786
S. marinus		**Total**					

Appendix I (continued)

Summary statistics for redfish samples collected for microsatellite analysis showing species names, morphological types, geographical regions, position of the sampling locations, sampling years, sampling months, station numbers, average sampling depths, coordinating institutes/nations, cruise identification numbers, sample codes and sample sizes (*n*).

Station	Nation	Cruise id	Sample code	*n*
560	Norway[4]	JH-81549	MANO01	78
2720023	Faroe Islands[5]	MH2002	MAEF02	59
2720003	Faroe Islands[5]	MH2002	MAWF02	39
918	Iceland[3]	B13_2000	MANEIc00	52
930	Iceland[3]	B13_2000	MASEIc00	101
881	Iceland[3]	B13-2001	MASEIc01	55
411	Iceland[3]	B82000	MASWIc00	60
752	Iceland[3]	B13-2001	MASWIc01	21
753	Iceland[3]	B13-2001	MASWIc01	31
974	Germany[1]	WH233	MAEGrDB01	24
975	Germany[1]	WH233	MAEGrDB01	26
695	Germany[1]	WH221	MAEGr00	71
707	Germany[1]	WH221	MAEGr00b	47
995	Germany[1]	WH233	MAEGr01	31
996	Germany[1]	WH233	MAEGr01	29
1050	Germany[1]	WH233	MAWGr01	26
1052	Germany[1]	WH233	MAWGr01	24
36	Spain[2]	CAFC01	MAFC01	26
85	Spain[2]	CAFC01	MAFC01	29
45	Spain[2]	CAFC01	MAFC01b	17
75	Spain[2]	CAFC01	MAFC01b	5
76	Spain[2]	CAFC01	MAFC01b	13
109	Spain[2]	CAFC01	MAFC01b	2
110	Spain[2]	CAFC01	MAFC01b	10
113	Spain[2]	CAFC01	MAFC01b	12
118	Spain[2]	CAFC01	MAFC01b	3
122	Spain[2]	CAFC01	MAFC01b	2
123	Spain[2]	CAFC01	MAFC01b	2
124	Spain[2]	CAFC01	MAFC01b	1
126	Spain[2]	CAFC01	MAFC01b	1
134	Spain[2]	CAFC01	MAFC01b	2
140	Spain[2]	CAFC01	MAFC01b	3
49	Iceland[3]	TJ11996	MAG96	10
50	Iceland[3]	TJ11996	MAG96	2
				914

[1] Federal Research Centre for Fisheries (BFAFi), Hamburg, Germany; [2] Institute of Marine Research, Vigo, Spain; [3] Marine Research Institute (MRI), Reykjavík, Iceland; [4] Institute for Marine Research (IMR) and the University of Bergen (UIB), Department of Fisheries and Marine Biology, Bergen, Norway; [5] Faroes Fisheries Laboratory, Tórshavn, Faroe Islands

Appendix I (continued)

Species	Type	Location	Position (mean) Longitude	Position (mean) Latitude	Year	Month	Depth (m)
S. mentella		SE-Iceland	632940N	0114120W	2001	October	385
S. mentella		SE-Iceland	631270N	0121140W	2001	October	412
S. mentella		SE-Iceland	630080N	0124860W	2001	October	565
S. mentella		SE-Iceland	634610N	0124000W	2001	October	546
S. mentella		Central Irminger Sea	622490N	0280960W	2001	June	640
S. mentella		Central Irminger Sea	602798N	0315167W	2003	June	700
S. mentella		Greenland East	640574N	0361119W	2000	September	-
S. mentella		Greenland East	635189N	0363407W	2001	October	321
S. mentella		Greenland East	634840N	0364037W	2001	October	363
S. mentella		Greenland East	633839N	0372445W	2001	October	255
S. mentella		Greenland West	600914N	0470932W	2001	November	338
S. mentella		Southern Irminger Sea	571483N	0409316W	2001	June	159
S. mentella		NAFO1F	562854N	0474909W	2001	July	425
S. mentella		NAFO2J	519823N	0498223W	2001	September	201
S. mentella		Flemish Cap	465203N	0447613W	2001	July	314
S. mentella		Flemish Cap	465653N	0457677W	2001	July	351
S. mentella		Flemish Cap	465155N	0459890W	2001	July	417
S. mentella		**Total**					
S. fasciatus		Flemish Cap	461100N	0463700W	1997	July	392
S. fasciatus		Flemish Cap	451300N	0462700W	1997	July	415
S. fasciatus		Flemish Cap	469313N	0440967W	2001	July	330
S. fasciatus		Flemish Cap	466842N	0453872W	2001	July	240
S. fasciatus		Flemish Cap	466842N	0453872W	2001	July	463
S. fasciatus		Flemish Cap	479502N	0447565W	2001	July	322
S. fasciatus		Flemish Cap	476687N	0448737W	2001	July	240
S. fasciatus		Flemish Cap	475530N	0452262W	2001	July	253
S. fasciatus		Flemish Cap	469425N	0447085W	2001	July	148
S. fasciatus		Flemish Cap	476645N	0441683W	2001	July	431
S. fasciatus		**Total**					
S. viviparus		Norway	701250N	0030563E	2001	October	150
S. viviparus		Norway	704790N	0029371E	2001	October	110
S. viviparus		Norway	710560N	0028365E	2001	October	170
S. viviparus		Norway	711390N	0027314E	2001	October	270
S. viviparus		SW-Iceland	632100N	0253900W	1996	November	370
S. viviparus		SW-Iceland	635998N	0233273W	2001	March	135
S. viviparus							
Sebastes sp.		**Total**					

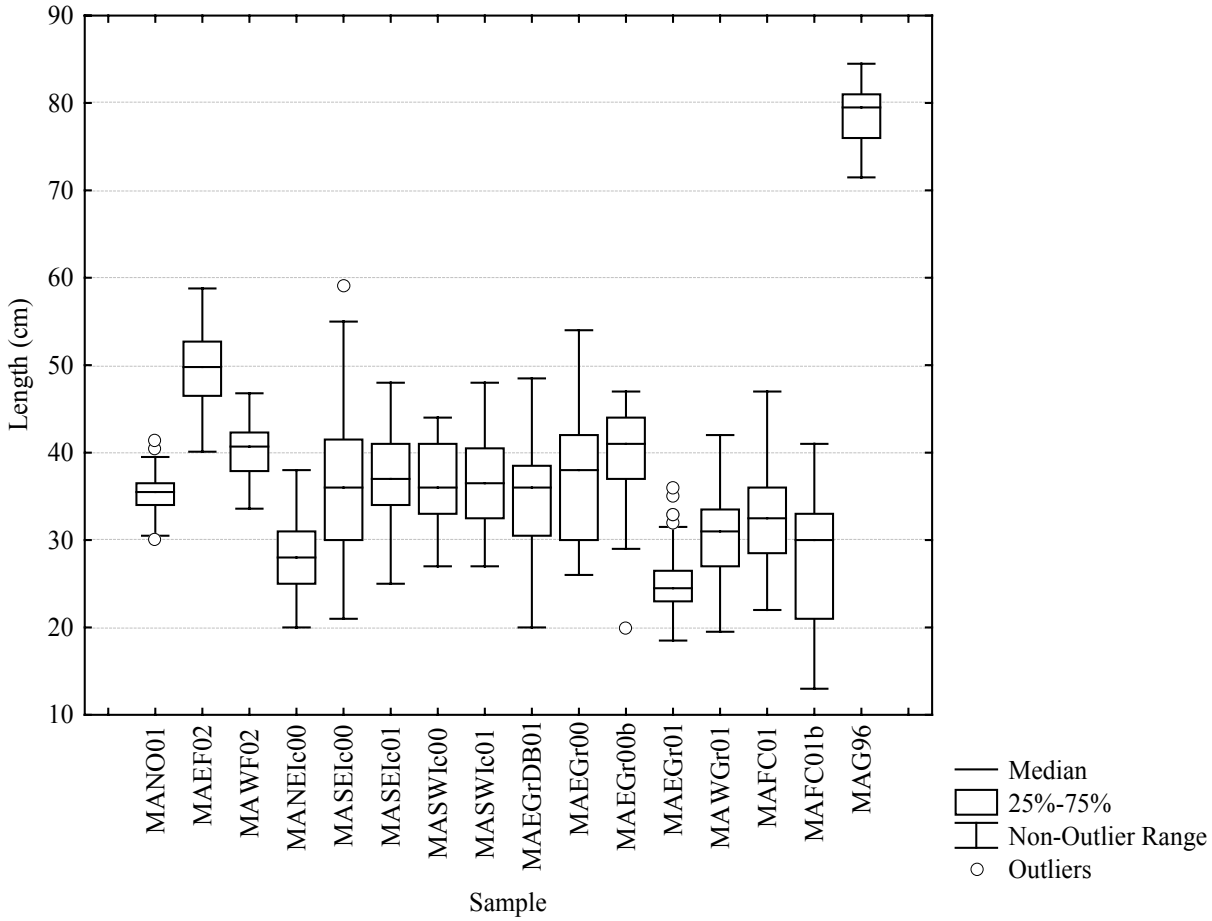
Appendix I (continued)

Station	Nation	Cruise id	Sample code	*n*
574	Iceland[3]	A12-2001	MESEIc01	14
575	Iceland[3]	A12-2001	MESEIc01	16
576	Iceland[3]	A12-2001	MESEIc01	11
583	Iceland[3]	A12-2001	MESEIc01	10
275	Iceland[3]	A8-2001	MECIrm01	57
447	Germany[1]	WH252	MECIrm03	66
705	Germany[1]	WH221	MEEGr00	68
1009	Germany[1]	WH233	MEEGr01	23
1010	Germany[1]	WH233	MEEGr01	20
1011	Germany[1]	WH233	MEEGr01	15
1050	Germany[1]	WH233	MEWGr01	51
704	Germany[1]	WH229	MESIrm01	48
710	Germany[1]	WH229	MENAFO1F01	48
723	Germany[1]	WH229	MENAFO2J01	50
15	Spain[2]	CAFC01	MEFC01	19
24	Spain[2]	CAFC01	MEFC01	25
25	Spain[2]	CAFC01	MEFC01	17
				558
81	Spain[2]	CAFC97	FAFC97	3
88	Spain[2]	CAFC97	FAFC97	10
3	Spain[2]	CAFC01	FAFC01	20
40	Spain[2]	CAFC01	FAFC01	11
59	Spain[2]	CAFC01	FAFC01	3
63	Spain[2]	CAFC01	FAFC01	9
75	Spain[2]	CAFC01	FAFC01	12
79	Spain[2]	CAFC01	FAFC01	6
109	Spain[2]	CAFC01	FAFC01	3
118	Spain[2]	CAFC01	FAFC01	4
			FA	**81**
514	Norway[4]	JH-81508	VINO01	1
516	Norway[4]	JH-81508	VINO01	1
518	Norway[4]	JH-81508	VINO01	4
519	Norway[4]	JH-81508	VINO01	8
10	Iceland[3]	TJB11996	VISWIc96	12
20	Iceland[3]	TBI-02	VISWIc01	16
Total			**VI**	**42**
				1595

[1] Federal Research Centre for Fisheries (BFAFi), Hamburg, Germany; [2] Institute of Marine Research, Vigo, Spain; [3] Marine Research Institute (MRI), Reykjavík, Iceland; [4] Institute for Marine Research (IMR) and the University of Bergen (UIB), Department of Fisheries and Marine Biology, Bergen, Norway; [5] Faroes Fisheries Laboratory, Tórshavn, Faroe Islands

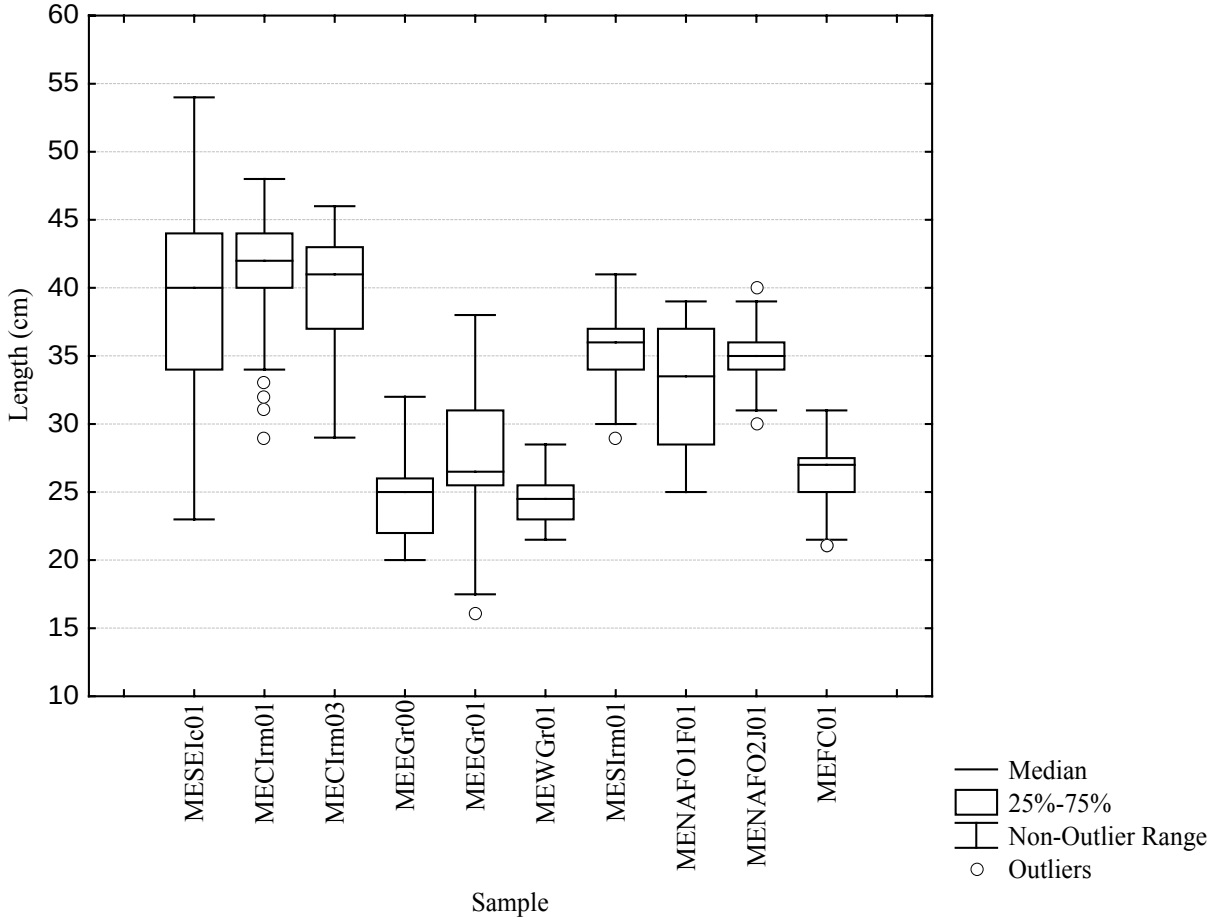
Appendix IIa

Box-and-whisker plot of the length distributions (total length) in the different samples of *S. marinus* used in the microsatellite analysis. Plots were generated in STATISTICA, version 6.1 (StatSoft Inc. 2003). The line within the box represents the median of the distribution. The 25th and 75th percentiles are represented by the top and bottom box edges, respectively. Whiskers represent the lowest and highest values (Non-Outlier range). Sample codes are explained in Table 2.1 and Appendix I.



Appendix IIb

Box-and-whisker plot of the length distributions (total length) in the different samples of *S. mentella* used in the microsatellite analysis. Plots were generated in STATISTICA, version 6.1 (StatSoft, Inc. 2003). The line within the box represents the median of the distribution. The 25th and 75th percentiles are represented by the top and bottom box edges, respectively. Whiskers represent the lowest and highest values (Non-Outlier range). Sample codes are explained in Table 2.1 and Appendix I.



Appendix III

Summary of basic genetic data for redfish samples from different locations and years. Sample codes are explained in Table 2.1 and Appendix I. For each of the eight microsatellite loci analysed, the number of individuals scored (*n*), the number of alleles, the number of "private" alleles (alleles only present in one sample), expected (H_e) and observed (H_o) heterozygosity, results of the tests for deviations from Hardy-Weinberg (HW) proportions ("Probability-test" of Guo and Thompson 1992; P-values and level of significance, H_0 = random union of gametes) and allelic richness (r(50)) are given. Significance levels were adjusted according to the sequential Bonferroni method (Rice 1989). * P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001

	Locus								Average
Sample	SEB9	SEB25	SEB30	SEB31	SEB33	SEB37	SEB45	SEB46	over loci
VI									
n	42	42	41	42	41	41	41	42	
No. of alleles	7	6	26	4	28	29	18	21	17
No. of private alleles	0	1	0	0	2	0	0	2	
He	0.685	0.577	0.936	0.522	0.941	0.944	0.761	0.888	0.782
Ho	0.667	0.619	0.878	0.524	0.951	0.829	0.585	0.738	0.724
HW	0.4286	0.1991	0.0010	0.1404	0.6202	0.0006	0.0029	0.0041	0.0000***
r(50)	6.85	5.79	25.60	3.99	27.60	28.45	17.51	20.55	17.04
FA									
n	81	81	73	81	80	81	81	81	
No. of alleles	10	18	34	9	27	18	14	15	18
No. of private alleles	1	0	1	0	1	0	1	0	
He	0.687	0.880	0.940	0.553	0.900	0.900	0.678	0.835	0.797
Ho	0.605	0.901	0.877	0.494	0.863	0.901	0.691	0.790	0.765
HW	0.2799	0.4541	0.0646	0.0094	0.2436	0.3003	0.0208	0.2441	0.0043*
r(50)	7.78	16.35	27.52	6.12	21.23	15.16	10.77	13.34	14.79
MANO01									
n	75	69	75	77	70	71	73	77	
No. of alleles	14	20	30	13	31	17	13	15	19
No. of private alleles	1	0	0	0	0	0	0	0	
He	0.808	0.906	0.741	0.831	0.954	0.730	0.798	0.899	0.834
Ho	0.720	0.884	0.747	0.844	0.971	0.718	0.808	0.896	0.824
HW	0.0086	0.0964	0.7555	0.4626	0.5051	0.2066	0.0178	0.5050	0.0168
r(50)	12.42	17.72	21.50	11.54	27.01	13.59	11.66	13.34	16.10
MAEF02									
n	59	59	59	59	59	59	59	59	
No. of alleles	13	20	28	13	32	18	19	12	19
No. of private alleles	0	0	0	0	0	0	0	0	
He	0.726	0.905	0.845	0.831	0.958	0.791	0.840	0.862	0.845
Ho	0.678	0.898	0.797	0.763	1.000	0.814	0.915	0.814	0.835
HW	0.4142	0.5033	0.2802	0.2058	0.9876	0.2795	0.4743	0.5489	0.5907
r(50)	11.46	18.38	23.72	11.72	29.10	15.70	16.10	11.62	17.22
MAWF02									
n	39	39	39	39	39	39	39	39	
No. of alleles	9	10	35	16	28	18	23	12	19
No. of private alleles	0	0	0	0	0	0	0	0	
He	0.784	0.730	0.947	0.907	0.935	0.812	0.929	0.812	0.857
Ho	0.821	0.744	0.872	0.949	0.974	0.718	0.846	0.744	0.833
HW	0.0367	0.2069	0.0115	0.4199	0.9493	0.0599	0.0565	0.3061	0.0050*
r(50)	9.00	10.00	35.00	16.00	28.00	18.00	23.00	12.00	18.88

Appendix III (continued)

Sample	Locus								Average over loci
	SEB9	SEB25	SEB30	SEB31	SEB33	SEB37	SEB45	SEB46	
MANEIc00									
n	51	52	51	51	49	52	52	52	
No. of alleles	14	19	34	11	33	18	21	16	21
No. of private alleles	0	0	0	0	1	0	0	0	
He	0.840	0.907	0.864	0.798	0.947	0.764	0.867	0.899	0.861
Ho	0.784	0.865	0.726	0.686	0.918	0.808	0.827	0.808	0.803
HW	0.4740	0.0160	0.0032	0.0509	0.0808	0.0756	0.8423	0.3727	0.0008**
r(50)	13.17	18.08	29.76	10.65	29.98	16.62	18.87	15.06	19.02
MASEIc00									
n	101	101	101	101	101	101	100	101	
No. of alleles	13	22	51	17	41	32	30	17	28
No. of private alleles	0	0	3	0	0	0	0	0	
He	0.872	0.818	0.926	0.862	0.958	0.858	0.932	0.838	0.883
Ho	0.782	0.733	0.822	0.792	0.980	0.634	0.830	0.822	0.799
HW	0.0000***	0.0011	0.0000***	0.0039	0.1454	0.0000***	0.0000***	0.3315	0.0000***
r(50)	11.29	16.43	33.21	14.08	32.37	19.90	23.34	13.45	20.51
MASEIc01									
n	54	54	55	54	50	54	55	55	
No. of alleles	15	18	48	14	37	24	27	17	25
No. of private alleles	0	0	1	0	0	0	0	1	
He	0.830	0.815	0.947	0.889	0.958	0.833	0.932	0.868	0.884
Ho	0.815	0.741	0.891	0.889	0.960	0.611	0.909	0.909	0.841
HW	0.1499	0.0114	0.0062	0.1495	0.1963	0.0000***	0.0032	0.2716	0.0000***
r(50)	13.76	15.80	39.52	13.09	33.83	19.67	24.30	15.21	21.90
MASWIc00									
n	60	60	59	60	60	60	60	60	
No. of alleles	13	19	23	13	30	17	17	15	18
No. of private alleles	0	0	0	0	0	0	0	0	
He	0.810	0.891	0.765	0.833	0.956	0.740	0.813	0.903	0.839
Ho	0.783	0.883	0.848	0.733	0.867	0.733	0.767	0.883	0.812
HW	0.6561	0.0890	0.9450	0.0004	0.0008	0.2562	0.0903	0.0906	0.0000***
r(50)	12.29	16.73	19.18	11.79	27.11	13.68	14.67	13.82	16.16
MASWIc01									
n	52	52	52	52	51	52	52	52	
No. of alleles	11	18	24	10	34	10	15	12	17
No. of private alleles	0	0	0	0	0	0	0	0	
He	0.723	0.902	0.806	0.783	0.956	0.649	0.822	0.881	0.815
Ho	0.750	0.942	0.827	0.808	0.961	0.692	0.769	0.865	0.827
HW	0.6174	0.3181	0.7341	0.6458	0.4953	0.5368	0.5298	0.0212	0.4270
r(50)	9.75	16.79	21.53	9.48	30.66	9.17	13.37	11.44	15.27
MAEGrDB01									
n	50	50	50	50	50	50	50	50	
No. of alleles	16	19	38	14	34	20	17	15	22
No. of private alleles	1	0	0	0	0	0	0	0	
He	0.864	0.897	0.933	0.849	0.956	0.760	0.865	0.846	0.871
Ho	0.840	0.840	0.880	0.700	0.960	0.640	0.800	0.840	0.813
HW	0.0707	0.1047	0.0398	0.0003*	0.6385	0.0017	0.0212	0.0742	0.0000***
r(50)	14.85	18.03	33.61	13.24	31.41	18.02	16.05	13.85	19.88

Appendix III (continued)

Sample	Locus SEB9	SEB25	SEB30	SEB31	SEB33	SEB37	SEB45	SEB46	Average over loci
MAEGr00									
n	71	71	71	71	71	71	71	71	
No. of alleles	15	18	44	16	41	25	24	15	25
No. of private alleles	0	0	2	0	0	1	1	0	
He	0.876	0.867	0.933	0.824	0.964	0.824	0.909	0.833	0.879
Ho	0.803	0.789	0.916	0.859	0.986	0.648	0.873	0.789	0.833
HW	0.0001*	0.0467	0.0292	0.4928	0.8681	0.0000***	0.0002*	0.1338	0.0000***
r(50)	13.19	16.36	32.28	13.67	33.76	18.52	20.38	12.75	20.11
MAEGr00b									
n	47	47	47	47	47	47	47	47	
No. of alleles	11	16	31	13	30	15	17	13	18
No. of private alleles	0	0	0	1	0	0	0	0	
He	0.825	0.806	0.906	0.721	0.941	0.727	0.814	0.822	0.820
Ho	0.830	0.723	0.894	0.809	1.000	0.596	0.702	0.872	0.803
HW	0.0105	0.0444	0.0802	0.8334	0.8564	0.0012	0.0064	0.4324	0.0001**
r(50)	10.77	15.24	28.16	12.26	28.30	14.04	15.55	12.12	17.05
MAEGr01									
n	60	60	58	60	60	59	60	60	
No. of alleles	12	15	42	16	38	18	27	15	23
No. of private alleles	0	0	0	0	0	0	2	0	
He	0.786	0.685	0.947	0.897	0.952	0.834	0.932	0.804	0.855
Ho	0.633	0.650	0.810	0.917	0.967	0.627	0.917	0.833	0.794
HW	0.0014	0.1647	0.0000***	0.3591	0.7797	0.0000***	0.4080	0.2381	0.0000***
r(50)	10.96	12.67	33.30	14.63	31.81	14.26	23.15	12.93	19.21
MAWGr01									
n	50	50	50	50	50	50	50	50	
No. of alleles	14	15	40	14	32	20	20	14	21
No. of private alleles	0	0	3	1	0	0	0	0	
He	0.870	0.853	0.928	0.794	0.957	0.787	0.887	0.861	0.867
Ho	0.800	0.800	0.920	0.760	1.000	0.620	0.800	0.820	0.815
HW	0.0127	0.0627	0.1308	0.4087	0.2289	0.0007	0.0141	0.3158	0.0000***
r(50)	13.06	14.56	34.11	12.76	30.09	17.49	18.48	13.24	19.22
MAFC01									
n	55	54	55	55	54	55	55	55	
No. of alleles	12	18	31	13	28	13	14	14	18
No. of private alleles	0	0	0	0	2	0	0	0	
He	0.838	0.894	0.772	0.874	0.945	0.695	0.855	0.903	0.847
Ho	0.909	0.944	0.818	0.746	0.907	0.582	0.855	0.909	0.834
HW	0.9558	0.2364	0.5292	0.016	0.0014	0.0673	0.4782	0.0603	0.0014*
r(50)	11.76	17.15	26.15	12.10	25.50	11.64	12.43	13.52	16.28
MAFC01b									
n	73	73	73	73	73	73	73	73	
No. of alleles	14	20	34	14	37	19	16	18	22
No. of private alleles	0	0	1	0	1	0	0	1	
He	0.881	0.894	0.859	0.864	0.947	0.733	0.870	0.908	0.869
Ho	0.849	0.877	0.822	0.740	0.904	0.726	0.808	0.863	0.824
HW	0.5720	0.0613	0.2564	0.0234	0.1529	0.1859	0.0536	0.4231	0.0112
r(50)	13.27	17.94	26.86	12.99	29.33	14.47	13.47	15.81	18.02

Appendix III (continued)

Sample	Locus								Average over loci
	SEB9	SEB25	SEB30	SEB31	SEB33	SEB37	SEB45	SEB46	
MAG96									
n	12	12	0	12	12	12	12	12	
No. of alleles	11	8	-	4	11	7	8	6	8
No. of private alleles	0	0	-	0	0	0	0	0	
He	0.861	0.670	-	0.660	0.823	0.625	0.809	0.733	0.740
Ho	0.833	0.667	-	0.500	0.500	0.667	0.750	0.750	0.667
HW	0.7684	0.3846	-	0.0365	0.0008**	0.7205	0.3270	0.3048	0.0116
r(50)	11.00	8.00	-	4.00	11.00	7.00	8.00	6.00	7.86
MESEIc01									
n	51	51	49	51	51	50	51	50	
No. of alleles	9	17	26	17	29	44	18	13	22
No. of private alleles	0	0	0	0	0	5	0	0	
He	0.731	0.888	0.934	0.860	0.946	0.950	0.885	0.677	0.859
Ho	0.686	0.882	0.898	0.784	0.980	0.960	0.961	0.440	0.824
HW	0.2566	0.043	0.1779	0.0213	0.8519	0.7889	0.2086	0.0003*	0.0007**
r(50)	8.91	16.02	24.77	16.11	26.69	37.70	16.46	11.32	19.75
MECIrm01									
n	57	57	56	57	57	57	57	55	
No. of alleles	14	18	33	18	31	39	22	15	24
No. of private alleles	0	0	0	0	0	1	0	2	
He	0.773	0.874	0.950	0.853	0.933	0.949	0.927	0.743	0.875
Ho	0.737	0.930	0.929	0.912	0.947	0.790	0.825	0.582	0.831
HW	0.6424	0.1504	0.1366	0.4611	0.6550	0.0000***	0.0017	0.0012	0.0000***
r(50)	12.83	16.55	28.86	16.19	27.16	32.35	20.33	13.17	20.93
MECIrm03									
n	66	66	66	66	66	66	66	66	
No. of alleles	13	18	40	19	29	43	22	18	25
No. of private alleles	1	1	2	0	0	0	0	0	
He	0.816	0.823	0.954	0.869	0.949	0.946	0.917	0.764	0.880
Ho	0.849	0.712	0.970	0.833	0.924	0.864	0.894	0.621	0.833
HW	0.6100	0.0073	0.7999	0.0948	0.3690	0.0255	0.1670	0.0001*	0.0001**
r(50)	11.28	15.36	32.02	15.88	25.76	31.96	18.41	15.06	20.72
MEEGr00									
n	68	68	68	68	68	68	68	68	
No. of alleles	11	16	31	14	31	43	20	12	22
No. of private alleles	0	0	0	0	0	0	0	0	
He	0.753	0.859	0.910	0.760	0.949	0.943	0.877	0.689	0.842
Ho	0.779	0.824	0.897	0.809	0.985	0.853	0.882	0.456	0.811
HW	0.0721	0.8718	0.0411	0.8135	0.9550	0.0321	0.1747	0.0000***	0.0002**
r(50)	9.28	14.13	25.05	11.82	26.39	33.77	17.62	10.17	18.53
MEEGr01									
n	58	58	57	58	58	57	57	57	
No. of alleles	12	12	30	19	30	29	26	13	21
No. of private alleles	0	0	0	0	0	0	1	1	
He	0.748	0.814	0.923	0.840	0.939	0.930	0.926	0.744	0.858
Ho	0.621	0.707	1.000	0.759	0.897	0.702	0.825	0.474	0.748
HW	0.0013	0.0810	0.0321	0.0281	0.5223	0.0000***	0.0200	0.0001*	0.0000***
r(50)	11.22	11.09	26.32	16.80	26.07	25.41	23.42	11.55	18.98

Appendix III (continued)

Sample	Locus SEB9	SEB25	SEB30	SEB31	SEB33	SEB37	SEB45	SEB46	Average over loci
MEWGr01									
n	51	51	51	51	51	51	51	51	
No. of alleles	10	13	19	14	25	25	20	9	17
No. of private alleles	0	0	0	0	0	0	0	0	
He	0.709	0.822	0.899	0.806	0.947	0.925	0.868	0.636	0.826
Ho	0.667	0.784	1.000	0.824	0.922	0.686	0.922	0.490	0.787
HW	0.1069	0.1334	0.1832	0.5935	0.3407	0.0002*	0.9006	0.0140	0.0006**
r(50)	9.47	12.16	17.64	13.13	24.11	23.48	18.06	8.65	15.84
MESIrm01									
n	48	48	48	48	48	48	47	48	
No. of alleles	10	13	23	13	30	27	21	13	19
No. of private alleles	0	0	0	0	0	0	0	0	
He	0.770	0.847	0.899	0.793	0.953	0.911	0.896	0.711	0.848
Ho	0.875	0.875	0.979	0.771	0.958	0.854	0.894	0.604	0.851
HW	0.3778	0.4553	0.7437	0.0082	0.7160	0.2109	0.0908	0.0019	0.0042*
r(50)	9.62	12.21	21.84	12.40	28.64	24.23	20.14	12.18	17.66
MENAFO1F01									
n	48	48	48	47	48	48	47	46	
No. of alleles	10	11	27	14	26	29	19	11	18
No. of private alleles	1	0	0	0	0	0	0	1	
He	0.728	0.822	0.928	0.799	0.946	0.911	0.891	0.680	0.838
Ho	0.813	0.792	0.979	0.745	0.875	0.833	0.894	0.457	0.798
HW	0.8300	0.3666	0.6599	0.0526	0.0045	0.0114	0.0084	0.0002*	0.0000***
r(50)	9.40	10.88	24.91	13.07	24.75	26.28	18.17	10.39	17.23
MENAFO2J01									
n	48	48	50	50	49	50	48	49	
No. of alleles	11	12	25	12	30	28	19	11	19
No. of private alleles	0	1	0	0	0	0	0	0	
He	0.708	0.833	0.916	0.780	0.944	0.925	0.885	0.687	0.835
Ho	0.771	0.833	0.900	0.720	0.959	0.760	0.854	0.490	0.786
HW	0.5818	0.6469	0.7578	0.0852	0.3761	0.0000***	0.0853	0.0068	0.0000***
r(50)	10.39	11.39	23.21	11.49	27.82	25.14	17.45	10.38	17.16
MEFC01									
n	61	61	61	54	61	61	61	61	
No. of alleles	11	17	28	16	30	34	19	9	21
No. of private alleles	0	0	0	1	0	0	1	0	
He	0.784	0.877	0.893	0.822	0.940	0.928	0.867	0.721	0.854
Ho	0.787	0.902	0.902	0.833	0.967	0.853	0.869	0.607	0.840
HW	0.1452	0.5935	0.8329	0.2945	0.7979	0.2765	0.3720	0.1064	0.3731
r(50)	10.24	15.80	22.48	14.44	25.98	28.62	16.42	8.95	17.87
Total									
No. of alleles	23	28	107	29	55	70	46	33	
No. of private alleles	5	3	13	3	7	7	6	8	
He per locus	0.786	0.840	0.896	0.808	0.947	0.841	0.868	0.804	
HW	0.0000***	0.0000***	0.0000***	0.0000***	0.0275	0.0000***	0.0000***	0.0000***	
r(50)	13.54	17.53	33.45	17.58	30.87	25.27	21.11	15.74	

Appendix IV

Inferred ancestry of unassigned/admixed individuals of North Atlantic *Sebastes* estimated using *STRUCTURE* (Pritchard et al. 2000). All 103 individuals were assigned to more than one cluster (individual proportion of membership $Q < 0.80$). The table shows the proportion of membership (Q) of each admixed individual in each of five inferred clusters (I, II, III, IV and V).

		Inferred cluster				
	Admixed	I	II	III	IV	V
Sample	individual no.	*S. fasciatus*	*S. mentella*	*Sebastes sp.*	*S. viviparus*	*S. marinus*
VI	1	0.013	0.211	0.048	0.717	0.011
FA	1	0.771	0.121	0.057	0.043	0.008
MANO01	1	0.006	0.005	0.343	0.004	0.642
	2	0.004	0.213	0.007	0.090	0.686
	3	0.005	0.274	0.005	0.020	0.696
MAWF02	1	0.183	0.006	0.006	0.008	0.797
MANEIc00	1	0.007	0.201	0.007	0.037	0.748
	2	0.012	0.475	0.042	0.004	0.467
	3	0.014	0.346	0.045	0.042	0.553
	4	0.008	0.067	0.207	0.002	0.716
	5	0.016	0.475	0.133	0.011	0.365
	6	0.014	0.751	0.090	0.003	0.142
	7	0.210	0.008	0.030	0.009	0.743
	8	0.010	0.033	0.005	0.203	0.749
MASEIc00	1	0.055	0.273	0.060	0.007	0.605
	2	0.003	0.320	0.025	0.007	0.645
MASEIc01	1	0.003	0.007	0.739	0.141	0.110
	2	0.046	0.057	0.770	0.124	0.003
	3	0.006	0.008	0.193	0.005	0.788
	4	0.038	0.068	0.737	0.129	0.028
MASWIc00	1	0.203	0.017	0.067	0.699	0.014
	2	0.062	0.012	0.041	0.098	0.787
MASWIc01	1	0.003	0.007	0.277	0.006	0.707
MAEGrDB01	1	0.007	0.245	0.241	0.502	0.005
	2	0.005	0.373	0.016	0.595	0.011
	3	0.005	0.103	0.743	0.005	0.144
	4	0.013	0.131	0.739	0.007	0.110
	5	0.067	0.144	0.747	0.003	0.039
	6	0.010	0.017	0.339	0.004	0.630
	7	0.013	0.009	0.355	0.008	0.615
	8	0.003	0.251	0.481	0.008	0.257
MAEGr00	1	0.017	0.024	0.635	0.035	0.289
	2	0.003	0.230	0.757	0.006	0.004
	3	0.002	0.005	0.772	0.002	0.219

Appendix IV (continued)

Sample	Admixed individual no.	Inferred cluster I *S. fasciatus*	II *S. mentella*	III *Sebastes sp.*	IV *S. viviparus*	V *S. marinus*
MAEGr00b	1	0.008	0.13	0.003	0.103	0.756
	2	0.205	0.006	0.759	0.024	0.006
MAEGr01	1	0.011	0.005	0.637	0.004	0.343
	2	0.006	0.01	0.639	0.006	0.339
	3	0.004	0.22	0.004	0.015	0.757
MAFC01	1	0.029	0.225	0.489	0.041	0.216
	2	0.216	0.655	0.010	0.009	0.110
	3	0.257	0.032	0.035	0.041	0.635
MAFC0b	1	0.191	0.007	0.013	0.008	0.781
	2	0.032	0.246	0.008	0.006	0.708
	3	0.02	0.108	0.006	0.237	0.629
	4	0.006	0.241	0.045	0.02	0.688
	5	0.006	0.077	0.029	0.202	0.686
	6	0.207	0.003	0.002	0.003	0.785
MESEIc01	1	0.004	0.242	0.040	0.596	0.118
MECIrm01	1	0.006	0.391	0.023	0.005	0.575
	2	0.005	0.389	0.567	0.011	0.028
	3	0.294	0.685	0.006	0.006	0.009
	4	0.023	0.786	0.068	0.102	0.021
	5	0.008	0.703	0.011	0.128	0.150
	6	0.045	0.673	0.014	0.261	0.007
	7	0.004	0.659	0.008	0.259	0.070
	8	0.026	0.596	0.042	0.329	0.007
	9	0.008	0.718	0.117	0.010	0.147
	10	0.018	0.741	0.041	0.183	0.017
MECIrm03	1	0.148	0.607	0.112	0.007	0.126
	2	0.013	0.533	0.430	0.010	0.014
	3	0.008	0.755	0.190	0.013	0.034
	4	0.009	0.771	0.132	0.060	0.028
	5	0.004	0.633	0.244	0.061	0.058
	6	0.208	0.766	0.004	0.005	0.017
	7	0.004	0.760	0.122	0.099	0.015
	8	0.006	0.675	0.290	0.022	0.007
	9	0.018	0.788	0.006	0.181	0.007
	10	0.098	0.685	0.045	0.159	0.013
MEEGr00	1	0.029	0.706	0.026	0.230	0.009
	2	0.137	0.395	0.423	0.032	0.013
	3	0.007	0.656	0.005	0.328	0.004
	4	0.006	0.784	0.010	0.004	0.196
	5	0.007	0.731	0.240	0.010	0.012
	6	0.006	0.005	0.616	0.004	0.369
	7	0.008	0.773	0.007	0.004	0.208

Appendix IV (continued)

Sample	Admixed individual no.	Inferred cluster I *S. fasciatus*	II *S. mentella*	III *Sebastes sp.*	IV *S. viviparus*	V *S. marinus*
MEEGr01	1	0.021	0.704	0.020	0.174	0.081
	2	0.010	0.758	0.013	0.009	0.210
	3	0.010	0.713	0.006	0.047	0.224
	4	0.026	0.542	0.391	0.019	0.022
	5	0.013	0.606	0.370	0.006	0.005
	6	0.222	0.632	0.103	0.005	0.038
	7	0.005	0.723	0.029	0.012	0.231
	8	0.004	0.292	0.011	0.687	0.006
	9	0.004	0.124	0.007	0.736	0.129
	10	0.009	0.581	0.007	0.341	0.062
MEWGr01	1	0.007	0.709	0.011	0.225	0.048
MESIrm01	1	0.026	0.534	0.012	0.004	0.424
	2	0.004	0.73	0.246	0.016	0.004
	3	0.004	0.745	0.005	0.184	0.062
MENAFO1F	1	0.006	0.355	0.626	0.006	0.007
	2	0.007	0.796	0.009	0.075	0.113
MENAFO2J	1	0.006	0.709	0.006	0.273	0.006
	2	0.006	0.139	0.004	0.306	0.545
	3	0.020	0.792	0.154	0.004	0.030
	4	0.003	0.574	0.004	0.416	0.003
	5	0.012	0.705	0.005	0.003	0.275
MEFC01	1	0.298	0.692	0.003	0.004	0.003
	2	0.134	0.755	0.042	0.004	0.065
	3	0.006	0.430	0.009	0.462	0.093
	4	0.006	0.421	0.014	0.015	0.544
	5	0.264	0.434	0.276	0.019	0.007
	6	0.208	0.753	0.006	0.021	0.012

CHAPTER 3

A study on the genetic structure of North Atlantic redfish (genus *Sebastes*; Cuvier 1829) based on amplified fragment length polymorphism (AFLP) markers

1. Abstract

The amplified fragment length polymorphism (AFLP) technique was used to investigate the species structure of North Atlantic redfish (genus *Sebastes*) and the population structure of the commercially important redfish species *Sebastes mentella*, with special emphasis on the population structure in the Irminger Sea and on the continental slopes of Greenland and Iceland. The genetic variation within and between 10 samples of *Sebastes mentella* (n = 604), two samples of *S. marinus* (n = 128), a sample of *S. fasciatus* (n = 64) and a sample of *S. viviparus* (n = 42) - collected from different locations throughout the North Atlantic - was determined and assessed. Out of 26 primer pairs initially used to survey polymorphisms within North Atlantic *Sebastes*, seven generated polymorphic AFLP fragments. These seven primer pair combinations produced a total of 85 polymorphic fragments. On average, 48 fragments per sample were polymorphic at the 5% level. Nei's gene diversity within each sample was relatively low (average H_j = 0.21). Nevertheless, Bayesian cluster analysis indicated the prevalence of five genetically distinct clusters among the four species studied. Individuals of *S. viviparus*, *S. fasciatus* and *S. mentella* were grouped into distinct clusters. In contrast, individuals pre-classified as *S. marinus* according to external morphological characters clustered into two genetically distinct groups, whereby one of the two groups most likely represents the species *S. marinus*. UPGMA analysis confirmed the genetic distinctness of these groups. The results obtained from the AFLP analysis provide therefore evidence that the currently recognised species *S. mentella*, *S. fasciatus* and *S. viviparus* are genetically distinct and therefore represent valid species. Furthermore, the results of the AFLP analysis support the existence of cryptic speciation in North Atlantic *Sebastes*, which was indicated by the results of the two preceding studies presented in this thesis employing mitochondrial DNA (see Chapter 1) and microsatellite markers (Chapter 2). The potential of AFLP markers to discriminate individuals from the five genetic clusters was high. On average, it was possible to assign approximately 99% of the individuals to one of the five clusters, without prior information about the sample of origin.

The results of this study also provided insights into the patterns of population structure in *S. mentella*. The main genetic patterns found in *S. mentella* using AFLP analysis were congruent with the findings of the two preceding studies presented in this thesis. However, the significant genetic differences observed between all samples indicated that AFLP markers might be too variable to detect biological meaningful patterns of genetic structure in more weakly differentiated populations of *S. mentella*. Therefore, the findings of the present study suggest that AFLP may be useful for species level investigations of North Atlantic *Sebastes*. Nevertheless, co-dominant markers such as microsatellites are preferable for population level investigations of *S. mentella.*

2. Introduction

North Atlantic redfish (genus *Sebastes*; Cuvier 1829) is a marine, viviparous fish that consists of a complex of four species, *Sebastes mentella* (Travin 1951), *S. marinus* (Linné 1758), *S. fasciatus* (Storer 1854) and *S. viviparus* (Krøyer 1845), which can be found in sympatry and thus contribute to mixed fisheries (Roques et al. 1999a).

The discrimination of these species has been complicated by their morphological similarity (Barsukov et al. 1985) and the low levels of interspecific genetic differentiation observed with proteins and mitochondrial markers (Payne and Ni 1982; McGlade et al. 1983; Trottier et al. 1989; Nedreaas and Nævdal 1989 and 1991a; Rubec et al. 1991; Nedreaas et al. 1994; Bentzen et al. 1998; Sundt and Johansen 1998). The difficulties in species discrimination have also hampered our understanding of the population structure of *Sebastes*.

Consequently, little is known about the population structure of the most widely distributed and commercially important North Atlantic *Sebastes* species, *S. mentella* (ICES 1998a).

Some authors separate *S. mentella* in the central North Atlantic into a so-called "oceanic" and a "deep-sea" phenotype, with differences in certain morphological (Magnússon and Magnússon 1995) and molecular (Johansen et al. 2000b) characters. There has been a strong controversy about whether the different *S. mentella* phenotypes represent different populations (reviewed in ICES 1998a and Saborido-Rey et al. 2005), in particular as a recent study based on microsatellite polymorphism revealed extremely low genetic differences within *S. mentella* over a distance of 6,000 km, from the Grand Banks and Labrador to the Faroe Islands (Roques et al. 2002).

Therefore, the population structure of *S. mentella* remains uncertain (ICES 2002).

The AFLP method (Zabeau and Vos 1993; Vos et al. 1995) is a relatively new molecular tool. It combines the reliability and reproducibility of RFLP (restriction fragment length polymorphism) analysis with the advantage of RAPD (randomly amplified polymorphic DNA) analysis of screening a high number of loci in a single PCR reaction (Liu et al. 1998). The AFLP procedure is based on the detection of DNA restriction fragments by PCR amplification. The technique assays the entire genome for polymorphic markers and requires relatively small amounts of genomic DNA, providing 10-100 times more markers than other molecular techniques (such as allozymes and RFLPs: Lu et al. 1996; Sharma et al. 1996).

In recent years AFLP's are quickly becoming the tool of choice for many applications and organisms, including species and population assignment and population genetics (reviewed in Vos and Kuiper 1997). A recent study has shown that AFLP's may be an alternative to microsatellites in order to enhance resolution in studies of population assignment (Campbell et al. 2003). In redfish, the AFLP technique was successfully applied to identify diagnostic markers for three colour morphotypes of the black rockfish (*Sebastes inermis*; Kai et al. 2002a).

Therefore, the AFLP technique could be a useful tool for detecting inter- and intraspecific genetic structure in North Atlantic *Sebastes*. In the present study, AFLP analysis was used to analyse the species structure of North Atlantic *Sebastes* and the population structure of *S. mentella*, with special emphasis on the population structure in the Irminger Sea and on the continental slopes of Greenland and Iceland.

3. Material and methods

3.1 Sample collection

North Atlantic redfish species and their morphological types were sampled by different scientists from institutes from Norway, Iceland, Spain and Germany (see Appendix I) during research surveys across the North Atlantic from 1996 to 2001.

Individual redfish were pre-classified into species on board using morphometrical characters (e.g. eye diameter, length and inclination of the preopercular spines) and morphological characters (e.g. spines, body shape, symphysial tubercle, colour) usually used for redfish (genus *Sebastes*) identification (Barsukov 1973; Barsukov et al. 1985). Samples from the shelf of Greenland were pre-classified into species using hemoglobin electrophoresis (Nedreaas and Nævdal 1989; 1991a), because the species identification - particularly the differentiation of *S. marinus* and *S. mentella* specimens - using standard characters was ambiguous. Hemoglobin analyses were performed by a Norwegian scientist[1] directly on board. These individuals were then frozen whole and were later analysed in the laboratory by a specialist in redfish morphometrics[2].

Four samples of specimens of *S. mentella* collected in the Irminger Sea were typed as "oceanic" and "deep-sea" type by Icelandic scientists[3], according to characters described by Magnússon and Magnússon (1995), e.g. colour and parasite infestation. In addition, DNA samples of 12 extremely large *S. marinus* (average length 78 cm) - defined as "giant" *S. marinus* by Icelandic scientists[3] - collected in the Irminger Sea in the year 1996 were provided by the cooperating institute from Iceland (see Appendix I).

Gill samples for genetic analyses were taken either directly from the fish on board the research vessels or afterwards in the laboratory at the different institutes (see Appendix I) from fish that had been frozen whole for subsequent morphometrical analyses. The gill samples were preserved in 100% ethanol.

In total 838 adult individuals of the genus *Sebastes* (604 *S. mentella,* 128 *S. marinus*, 64 *S. fasciatus* and 42 *S. viviparus*) were collected from different locations across the North

[1] Hemoglobin analyses were performed by Torild Johansen, University of Bergen (UIB), Norway.

[2] Morphometrical analyses were performed by Dolores Garabana, Institute of Marine Research, Vigo, Spain.

[3] Scientists from the Marine Research Institute (MRI) in Reykjavík, Iceland

Atlantic. Total lengths and sex (if possible) were recorded for all specimens. Table 3.1 gives an overview of the samples collected. In Appendix I the exact sampling information (e.g. exact position, depth) is listed.

The sampling locations are pictured in detail in Figure 3.1 (*S. marinus*, *S. viviparus*, *S. fasciatus*) and 3.2 (*S. mentella*). Size distributions of sampled *S. mentella* are given in Appendix II.

Table 3.1. Summary statistics for redfish samples collected showing species names, morphological types, sampling locations, sampling years, sampling months, sample codes, sex (number of males, number of females and number of individuals with unknown sex) and total sample sizes (*n*). The species names are based on the pre-classification on board using standard characters (Barsukov 1973; Barsukov et al. 1985) and, in case of the samples of *S. marinus* and *S. mentella* from Greenland, on hemoglobin electrophoretic patterns (Hb; Nedreaas and Nævdal 1989; 1991a). Dashes indicate missing information.

						Sample size (*n*)			
Species	Type	Location	Year	Month	Sample code	Males	Females	Unknown	Total
S. mentella		Norway	2001	October	MENO01	-	-	-	59
S. mentella		SE-Iceland	2001	October	MESEIc01	27	26	2	55
S. mentella		SW-Iceland	2001	October	MESWIc01	30	29	1	60
S. mentella Hb		Greenland East	2001	October	MEEGr01	31	20	6	57
S. mentella	deep-sea	Central Irminger Sea	1997	May	MECIrmds97	-	-	-	63
S. mentella	deep-sea	Central Irminger Sea	2001	June	MECIrmds01	39	32	1	72
S. mentella	oceanic	Central Irminger Sea	1997	May	MECIrmoc97	-	-	-	64
S. mentella	oceanic	Southern Irminger Sea	2001	June	MESIrmoc01	37	26	1	64
S. mentella		Southern Irminger Sea	2001	June	MESIrm01	35	13	0	48
S. mentella		Flemish Cap	2001	July	MEFC01	31	30	1	62
S. mentella						230	176	12	**604**
S. marinus		Norway	2001	October	MANO01	10	0	0	10
S. marinus		SE-Iceland	2001	October	MASEIc01	9	1	0	10
S. marinus		SW-Iceland	1997	May	MASWIc97	2	10	0	12
S. marinus Hb		Greenland East/Dohrn Bank	2001	October	MADB01	4	5	1	10
S. marinus Hb		Greenland East	2001	October	MAEGr01	7	3	0	10
S. marinus	giant	Irminger Sea	1996	August	MAG96	1	11	0	12
					MApool	33	30	1	64
S. marinus		Flemish Cap	2001	July	MAFC01	40	24	0	64
S. marinus									**128**
S. fasciatus		Flemish Cap	2001	July	FAFC01	22	42	0	64
S. fasciatus					**FA**				**64**
S. viviparus		Norway	2001	October	VINO01	-	-	-	14
S. viviparus		SW-Iceland	1996	November	VISWIc96	7	5	0	12
S. viviparus		SW-Iceland	2001	March	VISWIc01	7	9	0	16
S. viviparus					**VI**	14	14	0	**42**
Total									**838**

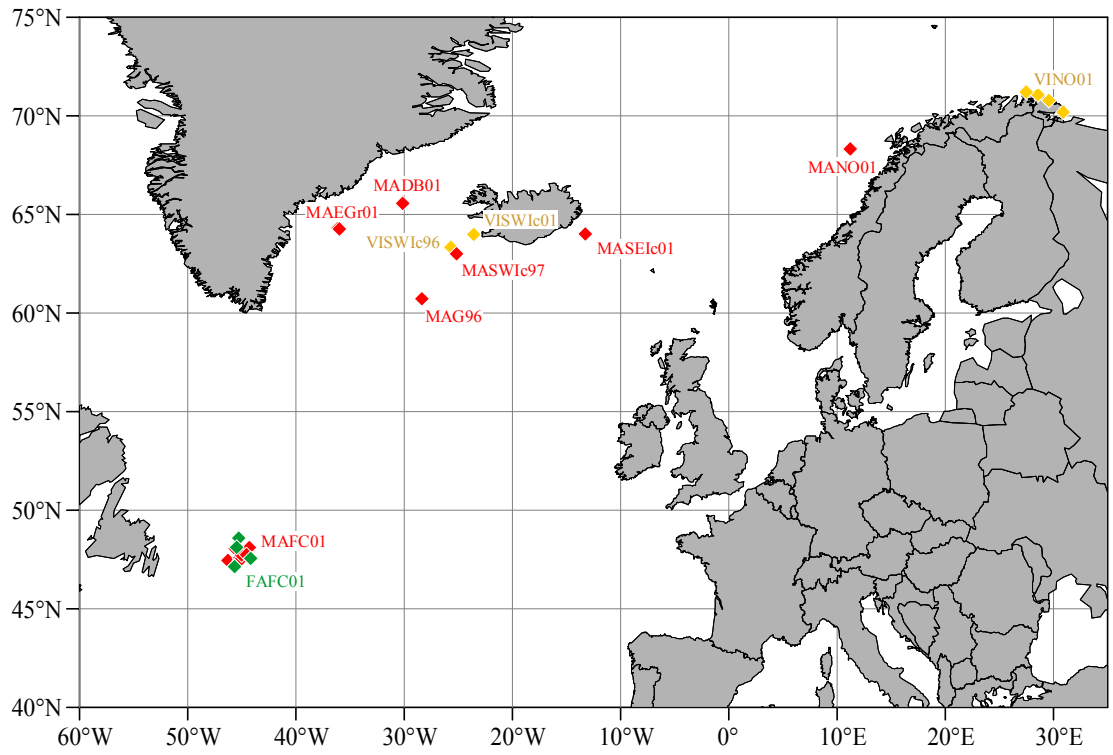


Figure 3.1. Sampling locations for the AFLP analysis of North Atlantic *Sebastes*. Red: S. *marinus*; yellow: *S. viviparus*; green: *S. fasciatus*. Sample codes are explained in Table 3.1.

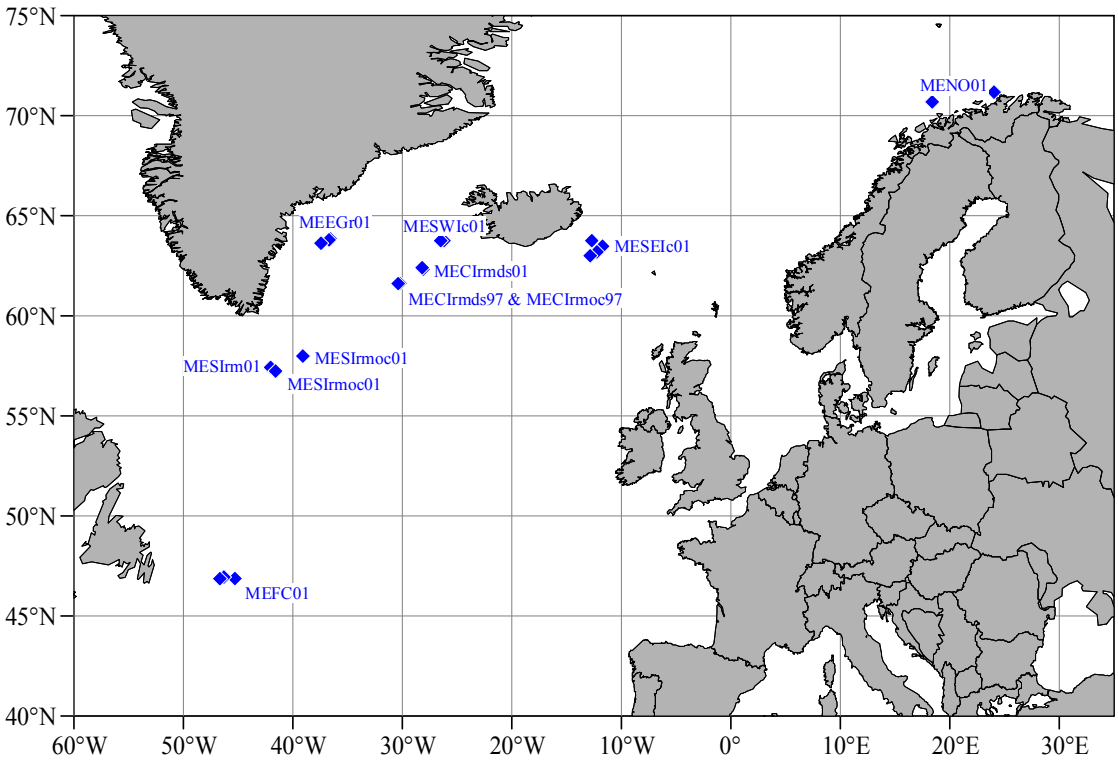


Figure 3.2. Sampling locations for the AFLP analysis of North Atlantic *Sebastes*. Blue: *S. mentella*. Sample codes are explained in Table 3.1.

3.2 Sample analysis

3.2.1 DNA extraction

Total genomic DNA was extracted with the DNeasy™ Tissue Kit (Qiagen) following manufacturers instructions. During this procedure, RNA was digested with RNase purchased from Qiagen following manufacturers instructions.

3.2.2 DNA quantitation

Quantitation of DNA was carried out using the fluorometric DNA assay with Hoechst 33258 Dye as described in Chapter 1.

3.2.3 AFLP analysis

The AFLP procedure includes four steps: Restriction, ligation of adapters, pre-selective amplification (or pre-amplification) and final selective amplification. The principles of the AFLP technique are described in detail in Vos and Kuiper (1997).
The basic principles are outlined below.
During the **restriction step** the purified DNA is digested with two different restriction enzymes, one with an average cutting frequency (e.g. *Hind*III) and the second one with a higher cutting frequency (e.g. *Mse*I). The double digestion of DNA results in three classes of restriction fragments with sticky ends: Fragments flanked only by *Mse*I sites, fragments flanked only by *Hind*III sites and fragments flanked by both *Hind*III and *Mse*I sites (Vos and Kuiper 1997).
In the **ligation step**, synthetic double stranded adapters (linkers) with matching sticky ends are ligated on the DNA fragments. The adapter and restriction site sequences serve as primer binding sites in the subsequent amplification steps.
Depending on genome characteristics, the restriction-ligation procedure may generate thousands of adapted fragments (Savelkoul et al. 1999). For sufficient fragment resolution after denaturing polyacrylamide gel electrophoresis, the number of fragments must be reduced. Therefore, in the third and fourth step of the AFLP analysis (pre-amplification and final selective amplification, respectively), subsets of restriction fragments are amplified by PCR (polymerase chain reaction; Mullis et al. 1986) using selective AFLP primers to reduce the complexity of the restriction fragment mixture (Vos and Kuiper 1997).
In the **pre-amplification step**, the restriction fragments are selectively pre-amplified using adapter-directed primers, each possessing a single selective nucleotide (e.g. adenine) at the 3' end. This single selective nucleotide will only allow amplification of restriction fragments with the matching nucleotide next to the linker. With one selective nucleotide on both

primers, statistically only 1 out of 16 restriction fragments will be amplified during the pre-amplification, achieving a 16-fold reduction of the complexity of the fragment mixture.
During the subsequent **final selective amplification**, primers containing two to three selective nucleotides (the selective nucleotide used in the pre-amplification and one to two additional nucleotides) are used to selectively amplify a fraction of the fragments. With one to two additional selective nucleotides on both primers the complexity of the restriction fragment mixture is reduced further, e.g. with two additional selective nucleotides only one out of 256 restriction fragments will be amplified during PCR.
However, since *Mse*I-*Mse*I fragments are the most frequent fragments, the number of fragments that are actually detected on a denaturing polyacrylamide gel has to be further limited. This is achieved by labelling only the final amplification primer binding to the *Hind*III (= rare cutter)-adapter with a fluorescent dye like IRD800. Thus, only PCR fragments are visualised, which carry at least one *Hind*III-adapter. The more frequent fragments carrying only *Mse*I-adapters remain undetected. The vast majority of bands detected on AFLP gels are fragments flanked by both restriction sites.

AFLP reactions were carried out according to a protocol by Trautner (2000) with several modifications. Sequences of adapters, pre-amplification primers and final amplification primers are given in Table 3.2. Adapters and primers were purchased from MWG Biotech.

Restriction
Genomic DNA was digested with the two restriction enzymes *Mse*I and *Hind*III (New England BioLabs). Approximately 300 ng genomic DNA, 2 units *Hind*III, 0.8 units *Mse*I and 2 µl 10x RL-buffer (restriction-ligation buffer: 10 mM Tris-HCl, 10 mM MgAc, 50 mM KAc, 5 mM DTT, pH 7.5) were mixed and added up to a volume of 20 µl with sterile dH_2O (PCR grade). The mixture was incubated at 37°C for 4 hours.

Ligation
Synthetic double stranded *Hind*III and *Mse*I adapters (MWG Biotech) were ligated to the restriction fragments using T_4 DNA ligase. First, *Hind*III and *Mse*I adapter molecules were prepared by hybridisation of the oligonucleotides U (= lower strand) and O (= upper strand). For preparation of a solution of 50 pmol/µl of the adapter *Mse*I-U/O, equal amounts of 100 pmol/µl solutions of *Mse*I-O and *Mse*I-U were mixed. For preparation of a solution of 5 pmol/µl of the adapter *Hind*III-U/O, equal amounts of 100 pmol/µl solutions of each *Hind*III-O and *Hind*III-U were mixed and diluted 10-fold using sterile dH_2O (PCR grade). The sequences of the adapters are listed in Table 3.2.
To ligate the adapters to the restriction fragments, 1.7 pmol *Hind*III-U/O adapter, 16.7 pmol *Mse*I-U/O adapter, 0.7 units T_4 ligase (New England Biolabs), 0.5 µl 10x RL-buffer (restriction-ligation buffer; see above) and 0.4 µl 10 mM ATP (Gibco/Life Technologies)

were added up to a volume of 5 µl with sterile dH_2O (PCR grade) and mixed with the restriction reaction. The ligation reaction was incubated at room temperature in the dark overnight.

Pre-amplification

The restriction-ligation products were selectively pre-amplified by PCR with adapter-directed primers, each possessing a single selective nucleotide, adenine, at the 3' end (*Mse*I+A primer and *Hind*III+A primer; primer sequences are given in Table 3.2).

The reactions were carried out in 25 µl reaction volumes containing 4.7 µl ligation reaction, 37.5 ng of each *Hind*III+A and *Mse*I+A primer, 2.4 units *Taq* polymerase (MWG Biotech), 0.2 mM of each deoxynucleoside triphosphate (dNTPs; Gibco/Life Technologies) and 1x PCR buffer (10 mM Tris-HCl, 50 mM KCl, 2 mM $MgCl_2$, 0.001% gelatin, pH 8.3; as recommended by Williams et al. 1990). A drop of mineral oil was used to cover the sequencing reactions to avoid evaporation.

PCR amplifications were performed in a thermal cycler (T-Gradient; Biometra) with 19 cycles of 30 seconds DNA denaturation at 94°C, 60 seconds primer annealing at 60°C and 60 seconds extension at 72°C.

Final selective amplification

Products of the pre-selective amplification step were diluted 10-fold with sterile dH_2O (PCR grade) and subjected to a second round of amplification using primers containing two to three selective nucleotides (adenine and one to two additional nucleotides, respectively; see Table 3.2).

As different primer combinations will generate different sets of fragments, a preliminary screening was used to choose primer pairs that generate suitable levels of variation for the taxa being studied. For the final selective amplification step, 26 primer pairs were initially used to survey polymorphisms among species, using 12 specimens of each species of North Atlantic *Sebastes* (*S. mentella*, *S. marinus*, *S. fasciatus* and *S. viviparus*). Subsequently, seven primer pairs were selected for the subsequent analysis of the redfish samples listed in Table 3.1, as they gave the clearest banding patterns and the highest levels of polymorphism. The remaining primer pairs generated none or only one polymorphic locus and were therefore not used. All 26 tested primer combinations and the seven primer combinations chosen for subsequent analysis are listed in Table 3.3.

The final selective amplification reactions were carried out in 10 µl reaction volumes containing 1 pmol IRD800-labelled *Hind*III primer, 5 pmol *Mse*I primer, 1x PCR buffer (10 mM Tris-HCl, 50 mM KCl, 2 mM $MgCl_2$, 0.001% gelatin, pH 8.3; as recommended by Williams et al. 1990), 2.5 µl diluted (10-fold) pre-amplification product, 0.2 mM of each deoxynucleoside triphosphate (dNTPs; Gibco/Life Technologies) and 0.05 units *Taq* polymerase (MWG Biotech).

PCR amplifications were performed in a thermal cycler (T-Gradient, Biometra) with 3 minutes of initial denaturation at 94°C, followed by 38 cycles of 30 seconds DNA denaturation at 94°C, 40 seconds primer annealing (see below) and 90 seconds extension at 72°C. In the first cycle the annealing temperature was set to 65°C. Then it was subsequently reduced by 0.7°C for the next 13 cycles and thereafter maintained at 56°C for the remaining 24 cycles. The advantage of this so-called "touchdown" PCR (Don et al. 1991) is that it improves the specificity of primer binding and therefore enriches for products containing correct matches between primers and template.

Table 3.2. Sequences of the adapters, pre-amplification primers and final amplification primers used in the AFLP analysis. * = fluorescent-labelled primers (IRD800-labelled); U = lower strand; O = upper strand

Adapter/Primer	Sequence (5'-3')
Adapters	
Hind III-O	5'-CTC GTA GAC TGC GTA CC -3'
Hind III-U	5'-AGC TGG TAC GCA GTC TAC-3'
Mse I-O	5'-GAC GAT GAG TCC TGA G -3'
Mse I-U	5'-TAC TCA GGA CTC AT-3'
Pre-amplification primers	
Hind III+A	5'-A GAC TGC GTA CCA GCT TA-3'
Mse I+A	5'-GAC GAT GAG TCC TGA GTA AA -3'
Final amplification primers	
Hind III-AGA*	5'-GAC TGC GTA CCA GCT TAG A-3'
Hind III-AGC*	5'-GAC TGC GTA CCA GCT TAG C-3'
Hind III-AGG*	5'-GAC TGC GTA CCA GCT TAG G-3'
Hind III-AGT*	5'-GAC TGC GTA CCA GCT TAG T-3'
Hind III-ACA*	5'-GAC TGC GTA CCA GCT TAC A-3'
Hind III-ACC*	5'-GAC TGC GTA CCA GCT TAC C-3'
Hind III-ACT*	5'-GAC TGC GTA CCA GCT TAC T-3'
Hind III-ATC*	5'-GAC TGC GTA CCA GCT TAT C-3'
Mse I-AAA	5'-GAT GAG TCC TGA GTA AAA A -3'
Mse I-AAC	5'-GAT GAG TCC TGA GTA AAA C -3'
Mse I-AAG	5'-GAT GAG TCC TGA GTA AAA G -3'
Mse I-ACC	5'-GAT GAG TCC TGA GTA AAC C -3'
Mse I-ACG	5'-GAT GAG TCC TGA GTA AAC G -3'
Mse I-ACT	5'-GAT GAG TCC TGA GTA AAC T -3'
Mse I-AC	5'-GAT GAG TCC TGA GTA AAC -3'
Mse I-AGC	5'-GAT GAG TCC TGA GTA AAG C -3'
Mse I-AGG	5'-GAT GAG TCC TGA GTA AAG G -3'
Mse I-AGT	5'-GAT GAG TCC TGA GTA AAG T -3'

Table 3.3. Primer pairs initially tested in the final selective amplification step to survey polymorphisms among species of North Atlantic *Sebastes*. Primer sequences are given in Table 3.2. Primer pairs finally chosen for large-scale analysis are in bold print.

Hind III primer (IRD800-labelled)	*Mse* I primer (unlabelled)
Hind III-AGA	*Mse* I-AGC
***Hind* III-AGA**	***Mse* I-AGG**
***Hind* III-AGC**	***Mse* I-AGG**
***Hind* III-AGG**	***Mse* I-AGG**
Hind III-AGG	*Mse* I-AGT
***Hind* III-AGT**	***Mse* I-AGG**
Hind III-ACA	*Mse* I-ACC
***Hind* III-ACA**	***Mse* I-AGG**
Hind III-ACC	*Mse* I-AAA
Hind III-ACC	*Mse* I-AAG
Hind III-ACC	*Mse* I-ACG
Hind III-ACC	*Mse* I-ACT
Hind III-ACC	*Mse* I-AC
***Hind* III-ACC**	***Mse* I-AGG**
Hind III-ACT	*Mse* I-AAA
Hind III-ACT	*Mse* I-AAC
Hind III-ACT	*Mse* I-AAG
Hind III-ACT	*Mse* I-ACG
Hind III-ACT	*Mse* I-AGC
***Hind* III-ACT**	***Mse* I-AGG**
Hind III-ATC	*Mse* I-AAA
Hind III-ATC	*Mse* I-AAG
Hind III-ATC	*Mse* I-ACG
Hind III-ATC	*Mse* I-ACT
Hind III-ATC	*Mse* I-AC
Hind III-ATC	*Mse* I-AGG

3.2.4 Analysis of AFLP fragments on an automated sequencer

AFLP fragments were separated, visualised and analysed on a denaturing polyacrylamide gel in an automated sequencer (LI-COR™ 4200 GeneReadIR DNA system) according to manufacturers recommendations.

For the fluorescence fragment analysis, the final amplification products were mixed with 10 µl AFLP loading buffer (98% formamide, 10 mM EDTA, 0.025% bromophenol blue,

0.025% xylene cyanol), denaturated for 5 min at 95°C and immediately after denaturation placed on ice.
1 µl of the mixture was loaded quickly on a 25 cm polyacrylamide gel (SequaGel™ XR; Biozym; prepared according to manufacturers recommendations). In the first and last lane of each gel, 0.3 µl of a sizing standard (50-700 bp Sizing Standard; LI-COR™) were run to determine the length of the AFLP fragments.
The data were collected automatically and analysed on computer using the Windows®-based software Saga Version 2.0 (LI-COR™). Multilocus profiles were scored for the presence (1) or absence (0) of fragments within the size range of 100-750 bp.
Data were assembled in a binary matrix. Only AFLP markers that could be scored unambiguously by eye were retained for analysis.

3.3 Data analysis

Due to the small sample sizes, the genetic data of the three *S. viviparus* samples were pooled and also the *S. marinus* samples with sample sizes of 12 individuals and less (MANO01, MASEIc01, MASWIc97, MADB01, MAEGr01, MAG96) were pooled. The two pooled samples were renamed VI and MApool, respectively.

3.3.1 Descriptive statistics and genetic diversity indices

Number and percentage of polymorphic markers per sample were calculated using the software TFPGA (Tools for Population Genetic Analysis) version 1.3 (Miller 1997).
Allele frequencies at each marker locus in each population, the number and percentage of polymorphic markers within each sample and genetic diversity within and between samples were computed using the software AFLP-SURV version 1.0 (Vekemans 2002). Allelic frequencies at AFLP loci were calculated from the observed frequencies of fragments according to the Bayesian approach of Zhivotovsky (1999) for diploid species and assuming Hardy-Weinberg genotypic proportions, as implemented in AFLP-SURV 1.0 (Vekemans 2002). A non-uniform prior distribution of allelic frequencies was assumed with its parameters derived from the observed distribution of fragment frequencies among loci (Zhivotovsky 1999). This procedure has been shown to produce almost unbiased estimates of allelic frequencies in dominant markers (Zhivotovsky 1999).
After estimating allele frequencies, statistics of genetic diversity were computed in AFLP-SURV following strictly the treatment of Lynch and Milligan (1994). For each sample, the number and percentage of polymorphic markers according to the 5% criterion (= markers with the frequency of the marker allele comprised between 0.05 and 0.95) and the expected heterozygosity (or Nei's gene diversity H_j; Nei 1987; analogous to the gene diversity H_e, see Chapter 2) were computed.

3.3.2 Analysis of inter- and intraspecific genetic structure using model-based cluster analysis

The program *STRUCTURE* (Pritchard et al. 2000) Version 2.0 with the Windows® front-end (Pritchard and Wen 2003) was used to infer inter- and intraspecific genetic structure in the data set and to assign individuals to species. *STRUCTURE* is a Bayesian, Markov chain Monte Carlo-based (MCMC) approach that uses model-based clustering to identify the actual (unknown) number of genetic clusters or gene pools (or populations) in a given data set (= *K*) and to assign individuals probabilistically to these *K* genetic clusters. Each test performed in *STRUCTURE* - conditional on a particular value of *K* - yields a log-likelihood value of the data, $\ln\Pr(X \mid Z)$[1], the highest of which indicates which test was closest to the actual number of genetically distinct populations or clusters in the data set (Pritchard and Wen 2003). The algorithms and models are described in Pritchard et al. (2000) and Falush et al. (2003), but the basic principles of *STRUCTURE* are briefly outlined in Chapter 2 (paragraph 3.8.3) of this thesis.

Since AFLP markers are dominant markers, the applied ancestry model differed from the model, which was used in the microsatellite analysis (see Chapter 2). According to Pritchard et al. (2000), *STRUCTURE* can be applied to data of dominant genetic markers only under the "no-admixture" model, by treating each class of genotypes as being a haploid allele.

For this purpose, an individual with a band was designated as (1/-9) and an individual with no band as (2/-9) where –9 is the value used for missing data (Pritchard et al. 2000).

Even though there is less information than if all the genotypes could be distinguished (as with microsatellite data) and genotypic classes are treated rather than alleles, this setting is valid under the "no-admixture" model (Pritchard and Wen 2003). In contrast to the "admixture" model (which allows for the presence of admixed individuals in the sample) used in the microsatellite analysis, the "no admixture" model assumes that individuals are discretely from one gene pool or another. Therefore, under the "no admixture" model, the posterior probability that an individual *i* originates from population *K* is estimated - instead of the membership coefficient estimated under the "admixture" model (*Q*; representing an estimate of the amount of an individual's genome that is derived from one of the inferred parental populations; see Chapter 2).

The *STRUCTURE* analysis of the data set was divided into two steps, (i) analysis of the species structure and identification of misclassified individuals, (ii) estimation of the number of populations within samples of *S. mentella*. In all runs, the "no-admixture" model was used.

In the analysis of species structure, the independent allele model was applied, since this model is a convenient approximation for populations, which are not extremely closely related (Pritchard and Wen 2003). This can be expected when species (and not populations) are analysed.

[1] *X* denotes the genotypes of the sampled individuals, *Z* denotes the unknown clusters or populations of origin of the individuals

In the analysis of *S. mentella* population structure, the model of correlated allele frequencies (the *F* model by Falush et al. 2003) was used. This allows to extract much more information from the data and to perform accurate assignments of individuals in very closely related populations (Falush et al. 2003; the allele frequency models are explained in detail in Chapter 2; paragraph 3.8.3).
Each run consisted of a burn-in period of 100,000 steps followed by 1,000,000 MCMC replicates, as recommended by Pritchard et al. (2000).

Species structure and identification of misclassified individuals: First it was tested whether *STRUCTURE* is able to identify the species as genetically distinct, isolated groups. For this purpose, ten simulations were run. In each simulation, the *S. viviparus* sample, the *S. fasciatus* sample, the *S. marinus* samples and one of the ten samples of *S. mentella* were pooled. This was done to test if the different runs were consistent, even though different *S. mentella* samples were used in the analysis. The number of clusters represented in these samples was estimated by calculating the posterior probability of the data, assuming that these samples originated from one to ten genetic clusters (*K* from 1 to 10) in the study area, as described in Pritchard et al. (2000). To test whether the individuals cluster according to their morphological species classification, the simulations were run without prior information of species origin. The value of *K*, which was showing the highest likelihood in the ten runs, was chosen and the posterior probability that an individual originated from gene pool *K* was estimated for each individual.
Individuals with low posterior probabilities (< 0.80) in each of the *K* inferred clusters were treated as unassigned individuals. If single individuals of a sample were assigned to another genetic cluster as the major part of the individuals of this sample, they were treated as misclassified specimens and were removed from the data set in the subsequent statistical analyses, such as the population analysis of *S. mentella* using *STRUCTURE* (see below) and the estimations of genetic differentiation (see paragraph 3.3.3, 3.3.4 and 3.3.5).

Population structure of *S. mentella*: After excluding misclassified individuals from the *S. mentella* data set, the most probable number of populations (*K*) within the *S. mentella* samples was estimated by pooling all *S. mentella* samples and calculating the probability $\ln\Pr(X|K)$ of the data, assuming that they originated from one to ten populations. Three independent simulations were run iteratively to assess convergence.

3.3.3 Genetic distances and UPGMA (unweighted pair group method with arithmetic means)

The genetic divergence between all pairs of samples was determined using Nei's (1978) unbiased genetic distance. Nei´s (1978) genetic distances were calculated from allele frequencies using the Taylor expansion according to Lynch and Milligan (1994). Nei's (1978) genetic distances were graphically displayed using the UPGMA (unweighted pair group method with arithmetic means) clustering algorithm. UPGMA is a simple distance-based clustering algorithm (see also Chapter 1; paragraph 3.3.4) that assumes equal rates of divergences among lineages. Based on a distance matrix of pairwise genetic distances, UPGMA starts clustering the pair of taxonomic units (in this case samples) with the smallest genetic distance. Then the algorithm recalculates distances between this cluster and the remaining units. Thereafter, it adds the next closest unit to form a new cluster and repeats the process until all units have been integrated in the tree. The algorithm for UPGMA is explained in detail in Nei and Kumar (2000). The confidence of each node in the UPGMA tree was determined using 1000 bootstrap replicates, resampled with replacement over loci. Nei´s (1978) pairwise distances, bootstrapping and UPGMA dendrogram based on Nei's (1978) genetic distances were calculated using the computer program TFPGA (Miller 1997). TFPGA reports the bootstrap support values and the number of loci that support the topology of the tree defined by the original combined data (Miller 1997).

3.3.4 AMOVA and Φ_{ST}

The statistical analyses described in the previous paragraphs are based on estimated allele-frequency data. AFLPs can also be treated as binary traits, with two phenotypes at a locus (the presence and the absence of a band) and subsequently analysed using analysis of molecular variance (AMOVA) and the corresponding ϕ-statistics (Zhivotovsky 1999; see also Excoffier et al. 1992 and Chapter 1 and 2).

By using this approach, the partition of genetic diversity within and among the North Atlantic *Sebastes* samples was analysed by a hierarchical analysis of molecular variance (AMOVA; Weir and Cockerham 1984; Excoffier et al. 1992) as implemented in Arlequin version 2.000 (Schneider et al. 2000). Two different AMOVAs were calculated.

First, the degree of differentiation among the genetic clusters estimated by the *STRUCTURE* analysis (see paragraph 3.3.2) relative to the differentiation among samples within these clusters was quantified. For this purpose, individuals originating from the same sample and assigned to the same cluster were grouped and the proportion of the total genetic variation that can be found among the clusters and among samples within clusters was calculated.

A second AMOVA was performed to quantify the degree of differentiation among temporal samples of *S. mentella* from the same location relative to the degree of differentiation among samples from different locations. Temporal samples from the same location were grouped and

the proportion of the total genetic variation due to temporal differentiation and the proportion of the total genetic variation due to spatial differentiation were calculated. Based on the variance components, pairwise fixation indices (Φ_{ST}; Weir and Cockerham 1984; Excoffier et al. 1992) were calculated using Arlequin. Significance and resulting P-values of the AMOVA variance components and of the pairwise Φ_{ST} values were tested in Arlequin by using non-parametric permutation procedures as described in Excoffier et al. (1992) with 10,000 permutations. Significance levels for simultaneous tests were adjusted by using the sequential Bonferroni approach (Rice 1989).

3.3.5 Isolation by distance

To test if observed genetic differences between samples of *S. mentella* were correlated with geographic distance, isolation by distance (Wright 1943) analyses were performed.
Isolation by distance was examined by comparing Nei's (1978) unbiased genetic distances for each pair of samples, with the geographic distance separating sampling locations measured as shortest distance in km. The geographic distances among samples were obtained from a 1:1,000,000 scale map using the software MapViewer Version 6 (Golden Software Inc.). Significance in the isolation by distance relationship was tested statistically using a Mantel test (see Manly 1994) using a permutation procedure (10,000 iterations) as implemented in the software IBD 2.0 (Bohonak 2002). IBD assesses whether the pairwise genetic distance matrix is correlated with the pairwise geographic distance matrix. A null distribution is generated by randomising rows and columns of one matrix while holding the other matrix constant (Bohonak 2002). IBD 2.0 (Bohonak, 2002) calculates confidence intervals for slopes of IBD-relationships based on reduced major axis regression and bootstrapping over populations.

4. Results

4.1 Levels of polymorphism and genetic diversity

The seven primer pair combinations used for AFLP analysis generated a total of 85 polymorphic AFLP markers among the 838 individuals of the four North Atlantic *Sebastes* species analysed. The number of polymorphic markers per primer pair combination varied between eight (*Hind*III-AGA and *Mse*I-AGG*; Hind*III-ACT and *Mse*I-AGG) and 18 (*Hind*III-AGC and *Mse*I-AGG).
Table 3.4 gives an overview of the number of polymorphic markers generated by each primer pair combination. An example of the amplification products obtained is shown in Figure 3.3.

Table 3.4. Number of polymorphic AFLP markers generated among 838 individuals of North Atlantic *Sebastes sp.* using primer pair combinations of *Mse*I-AGG and seven different *Hind*III primers (*Hind*III+A and two additional selective nucleotides). The sequences of the eight final amplification primers are given in Table 3.2.

Primer pair		No. of polymorphic markers
Hind III-AGA	*Mse* I-AGG	8
Hind III-AGC	*Mse* I-AGG	18
Hind III-AGG	*Mse* I-AGG	9
Hind III-AGT	*Mse* I-AGG	16
Hind III-ACA	*Mse* I-AGG	13
Hind III-ACC	*Mse* I-AGG	13
Hind III-ACT	*Mse* I-AGG	8
Total		85

In Table 3.5 the number and percentage of polymorphic markers and expected heterozygosities (= Nei's gene diversity; H_j) within each sample are given. On average, 55 marker loci per sample were polymorphic. An average number of 48 marker loci per sample were polymorphic at the 5% level. Nei's gene diversity H_j within each sample was relatively low and varied between 0.19 and 0.26, with an average H_j of 0.21.
Visual inspection of the AFLP fragments revealed no species-diagnostic (species-specific) AFLP markers among North Atlantic *Sebastes* species.

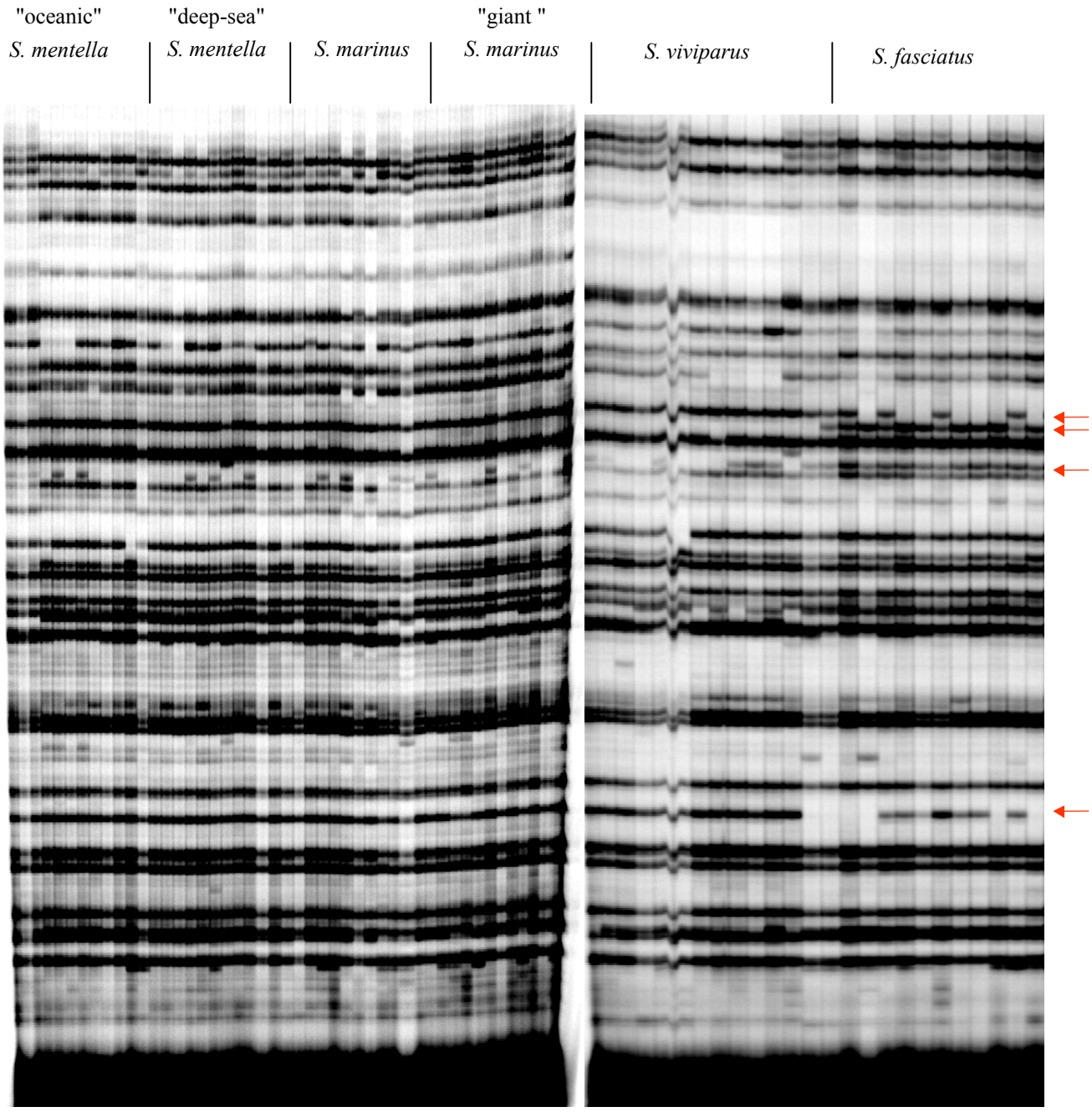


Figure 3.3. Subset of amplification products from twelve individuals of each of the North Atlantic *Sebastes* species and phenotypes generated using the primer pair *Mse*I-AGG and *Hind*III-AGC (from left to right: "oceanic" *S. mentella*, "deep-sea" *S. mentella*, *S. marinus,* "giant" *S. marinus*, *S. viviparus* and *S. fasciatus*). Red arrows: Examples for polymorphic AFLP markers.

Table 3.5. Number and percentage of polymorphic AFLP marker loci and expected heterozygosity within each sample. In total, 85 marker loci were scored for each sample. *n* = sample size; * = loci with allelic frequencies lying within the range 0.05 to 0.95; H_j = expected heterozygosity (Nei's gene diversity; analogous to the average gene diversity H_e); S.E. = standard error of H_j. Sample codes are explained in Table 3.1.

Sample	*n*	No. polymorphic loci	% polymorphic loci	No. of polymorphic loci at the 5% level*	Proportion of polymorphic loci at the 5% level* in %
VI	42	56	66	59	69.4
FA	64	50	59	39	45.9
MApool	64	63	74	66	77.6
MAFC01	64	59	69	53	62.4
MENO01	59	52	61	42	49.4
MESEIc01	55	56	66	43	50.6
MESWIc01	60	50	59	45	52.9
MEEGr01	57	58	68	65	76.5
MECIrmds97	63	53	62	46	54.1
MECIrmds01	72	55	65	44	51.8
MECIrmoc97	64	52	61	42	49.4
MESIrmoc01	64	52	61	42	49.4
MESIrm01	48	51	60	40	47.1
MEFC01	62	58	68	45	52.9

Table 3.5. (continued)

Sample	*n*	H_j	S.E. (H_j)
VI	42	0.19416	0.01756
FAFC01	64	0.18611	0.01918
MApool	64	0.25791	0.02021
MAFC01	64	0.24212	0.01943
MENO01	59	0.20240	0.02001
MESEIc01	55	0.19970	0.01906
MESWIc01	60	0.21463	0.02073
MEEGr01	57	0.22654	0.01868
MECIrmds97	63	0.22236	0.02076
MECIrmds01	72	0.22103	0.02099
MECIrmoc97	64	0.20306	0.02060
MESIrmoc01	64	0.19416	0.01955
MESIrm01	48	0.19675	0.02011
MEFC01	62	0.21697	0.01980

4.2 Results of the model-based cluster analyses

4.2.1 Species structure and species assignment

In all ten simulations performed using *STRUCTURE* (Pritchard et al. 2000) with the pooled samples of the four North Atlantic *Sebastes* species (VI, FA, MApool and MAFC01, combined with one out of ten samples of *S. mentella*, respectively), the natural logarithm of the probability of the data lnPr($X|K$) for each K (= number of genetic clusters) strongly increased from $K = 1$ to $K = 5$ (see Figure 3.4). For $K > 5$, there was relatively little increase in lnPr($X|K$). In situations, where lnPr($X|K$) increases progressively as K increases, the appropriate ("real") value of K (the most likely number of genetic clusters in the data set) is the value of K at which lnPr($X|K$) plateaus, rather than at the highest likelihood (Pritchard et al. 2000; Pritchard and Wen 2003). In all ten simulations, independently of the *S. mentella* sample included in the simulations, the posterior probabilities lnPr($X|K$) reached a plateau between $K = 5$ and $K = 6$ (see Figure 3.4). Furthermore, in all ten simulations assuming $K = 5$, nearly all individuals were strongly assigned to one of the five clusters (see Table 3.6). Therefore, $K = 5$ was chosen as the most appropriate number of gene pools describing the genetic structure in the data set.

The average posterior probabilities of membership of each sample in the five inferred clusters and the number of individuals assigned to each cluster with a posterior probability ≥ 0.80, as well as the number of unassigned individuals (assigned to each of the five inferred clusters with a posterior probability < 0.80), are reported in Table 3.6.

The global performance of *STRUCTURE* allowed assigning 99.4% of the individuals, with 833 out of 838 individuals revealing a posterior probability of membership equal to or larger than 0.80 in one of the five clusters. Only five individuals could not be assigned, as they revealed posterior probabilities of less than 0.80 in each of the five inferred clusters.

Almost all individuals of the *S. viviparus* sample were assigned to cluster II - except three individuals, which were assigned to cluster I. Individuals of the *S. fasciatus* sample were mainly assigned to cluster IV (61 out of 64 individuals), except one individual that was assigned to cluster I and two individuals, which were assigned to cluster V. Most individuals (60 out of 64) of the *S. marinus* sample from the Flemish Cap were assigned to cluster V, whereas the pooled individuals of *S. marinus* collected in different areas (MApool) were only partly (35 out of 64 individuals) assigned to this cluster (cluster V). Of the remaining individuals of sample MApool, 24 individuals were assigned to a separate cluster (cluster III). Five individuals of MApool could not be assigned to one of the five inferred clusters with a posterior probability equal to or larger than 0.80.

About 98% (593 out of 604) of the individuals of the samples of *S. mentella* were assigned to cluster I - with a high posterior probability of membership (≥ 0.80). Of the remaining twelve

individuals, eight individuals (all occurring in the of *S. mentella* sample from East Greenland) were grouped into cluster III (together with the 24 individuals of the sample MApool). Three individuals of the *S. mentella* samples (two individuals of the sample from SE-Iceland and one individuals of the sample from the Flemish Cap) were assigned to cluster V (the cluster into which most of the individuals of the *S. marinus* samples were grouped). The 12 individuals of the *S. mentella* samples that were assigned to cluster III and V, respectively, were excluded from the subsequent analysis of genetic population structure of *S. mentella*.

Due to the clear clustering of the samples into the five inferred clusters, these clusters were subsequently designated as follows: Cluster I was designated as "*S. mentella* cluster", cluster II as "*S. viviparus* cluster", cluster IV was designated as "*S. fasciatus* cluster" and cluster V as "*S. marinus* cluster". Since a part of the individuals of the *S. marinus* sample MApool were grouped into cluster III, separately from the remaining individuals of MApool and the individuals of the *S. marinus* sample from the Flemish Cap (MAFC01), this cluster was designated as "*Sebastes sp.* cluster".

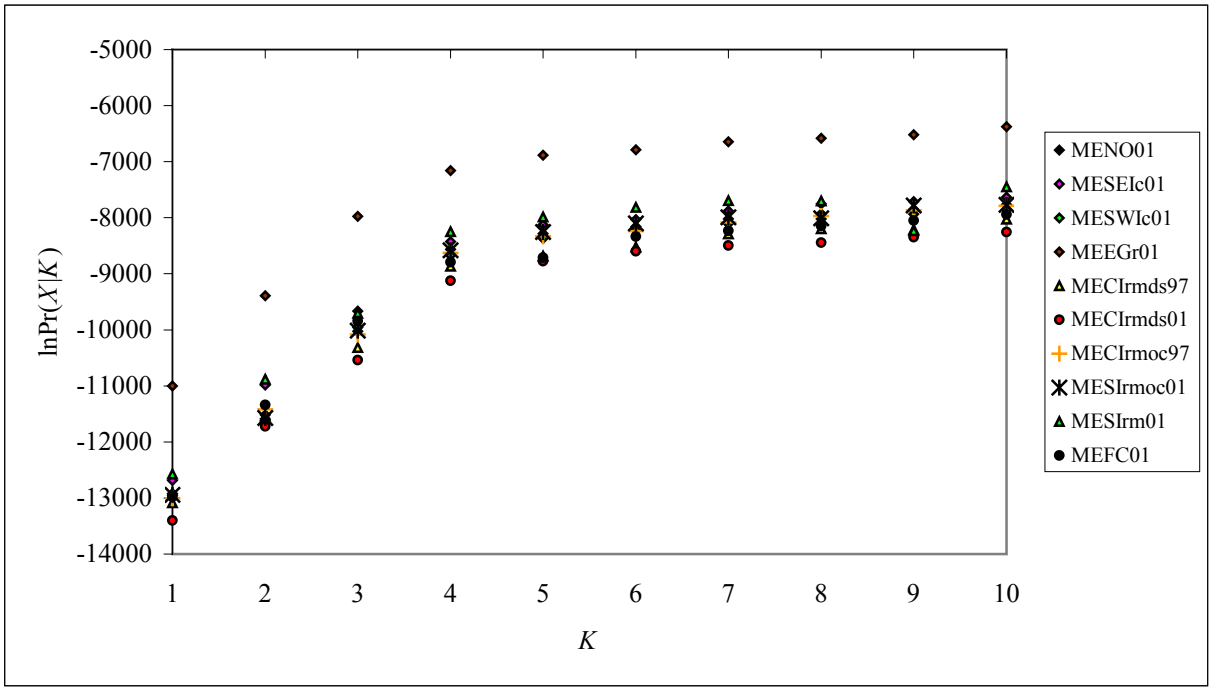


Figure 3.4. Values of lnPr($X|K$), the log-likelihood value of the data, as a function of the number of clusters (K = 1,2,…,10) inferred using *STRUCTURE* (Pritchard et al. 2000) in ten different simulation runs (85 AFLP loci; model assumptions: independent allele frequencies and no admixture). In each of these ten runs, the *Sebastes* samples VI, FA, MApool and MAFC01 were pooled with one of the ten different *S. mentella* samples (see legend) and *STRUCTURE* was run without using prior information of species origin.

Table 3.6. Results of the Bayesian clustering analysis performed using *STRUCTURE* (Pritchard et al. 2000) with the samples VI, FA, MApool, MAFC01 and the ten samples of *S. mentella*. The number of genetic clusters in the sample set was set to $K = 5$ as it was the value at which $\ln\Pr(X|K)$ plateaued (see text). The table shows the mean posterior probabilities that a sample originated from the genetic clusters I, II, III, IV or V. Each individual was assigned to a single cluster if the posterior probability that it originated from that cluster was equal to or larger than 0.80. The numbers of individuals assigned to each cluster with a posterior probability ≥ 0.80 are indicated in parentheses. If the posterior probabilities of membership of an individual in each of the five inferred clusters were smaller than 0.80, the individual was treated as unassigned individual. Sample codes are explained in Table 3.1.

		Inferred cluster					Number of	Percentage
		I	II	III	IV	V	unassigned	assigned
Sample	*n*	*S. mentella*	*S. viviparus*	*Sebastes sp.*	*S. fasciatus*	*S. marinus*	individuals	individuals
VI	42	0.071	0.929	0.000	0.000	0.000	0	100.00
		(3)	(39)	(0)	(0)	(0)		
FAFC01	64	0.016	0.000	0.000	0.953	0.031	0	100.00
		(1)	(0)	(0)	(61)	(2)		
MApool	64	0.000	0.000	0.412	0.000	0.588	5	92.19
		(0)	(0)	(24)	(0)	(35)		
MAFC01	64	0.000	0.000	0.000	0.062	0.937	0	100.00
		(0)	(0)	(0)	(4)	(60)		
MENO01	59	1.000	0.000	0.000	0.000	0.000	0	100.00
		(59)	(0)	(0)	(0)	(0)		
MESEIc01	55	0.964	0.000	0.000	0.000	0.036	0	100.00
		(53)	(0)	(0)	(0)	(2)		
MESWIc01	60	1.000	0.000	0.000	0.000	0.000	0	100.00
		(60)	(0)	(0)	(0)	(0)		
MEEGr01	57	0.860	0.000	0.140	0.000	0.001	0	100.00
		(49)	(0)	(8)	(0)	(0)		
MECIrmds97	63	0.999	0.000	0.001	0.000	0.000	0	100.00
		(63)	(0)	(0)	(0)	(0)		
MECIrmds01	72	1.000	0.000	0.000	0.000	0.000	0	100.00
		(72)	(0)	(0)	(0)	(0)		
MECIrmoc97	64	1.000	0.000	0.000	0.000	0.000	0	100.00
		(64)	(0)	(0)	(0)	(0)		
MESIrmoc01	64	1.000	0.000	0.000	0.000	0.000	0	100.00
		(64)	(0)	(0)	(0)	(0)		
MESIrm01	48	1.000	0.000	0.000	0.000	0.000	0	100.00
		(48)	(0)	(0)	(0)	(0)		
MEFC01	62	0.984	0.000	0.000	0.000	0.016	0	100.00
		(61)	(0)	(0)	(0)	(1)		
S. mentella total	604	(593)	(0)	(8)	(0)	(3)	0	100.00
Total	838	(597)	(39)	(32)	(65)	(100)	5	99.40

As the sample MApool consisted of several samples of *S. marinus* from different locations (see Table 3.1), the assignment results for each single *S. marinus* sample are given in Table 3.7. Table 3.7 shows that most of the individuals assigned to cluster III (the "*Sebastes sp.* cluster") originated from the samples from the shelves of Iceland and Greenland (MASEIc01, MASWIc97, MADB01 and MAEGr01). In contrast, all individuals of the *S. marinus* sample from Norway and of the sample of "giant" *S. marinus* were assigned to the same cluster as the *S. marinus* specimens from the Flemish Cap (cluster V, the "*S. marinus* cluster").

Table 3.7. The number of individuals of the samples of *S. marinus* assigned to each of the five inferred clusters with a posterior probability of membership ≥ 0.80. Assignment analyses were performed using *STRUCTURE* (Pritchard et al. 2000) with the samples VI, FA, MApool, MAFC01 and the ten samples of *S. mentella* (see Table 3.6). The number of unassigned individuals (with posterior probabilities of membership in each of the inferred clusters < 0.80) is given in the last column. Sample codes are explained in Table 3.1.

		Inferred cluster					Number of
		I	II	III	IV	V	unassigned
Sample	*n*	*S. mentella*	*S. viviparus*	*Sebastes sp.*	*S. fasciatus*	*S. marinus*	individuals
MANO01	10					10	
MASEIc01	10			3		7	
MASWIc97	12			4		6	2
MADB01	10			7			3
MAEGr01	10			10			
MAG96	12					12	
MAFC01	64				4	60	
Total	128			24	4	95	5

4.2.2 Population structure of S. mentella

Also the inference of the number of populations K in the samples of *S. mentella* was not straightforward, since log-likelihood values for the data conditional on K increased progressively as K was increased (see Figure 3.5). As already mentioned in the previous paragraph, in such a case the most likely number of populations in the data set is the value of K at which $\ln\Pr(X|K)$ plateaus, rather than at the highest likelihood (Pritchard et al. 2000; Pritchard and Wen 2003). The natural logarithm of the probability of the data $\ln\Pr(X|K)$ for each K strongly increased from $K = 1$ to $K = 4$ (see Figure 3.5). For $K > 4$, there was relatively little increase in $\ln\Pr(X|K)$. In all three independent simulations, the posterior probabilities $\ln\Pr(X|K)$ reached a plateau between $K = 4$ and $K = 5$. Furthermore, Pritchard and Wen (2003) suggested choosing the smallest value of K that captures the major structure in the data. Therefore, $K = 4$ was chosen as the most appropriate number of gene pools describing the genetic structure in the data set.

The average posterior probabilities that a *S. mentella* sample originated from the genetic clusters I, II, III, or IV, as well as the number of individuals assigned to one of these clusters with a posterior probability ≥ 0.80 and the number of unassigned individuals (assigned to each of the five clusters with a posterior probability < 0.80), are reported in Table 3.8.

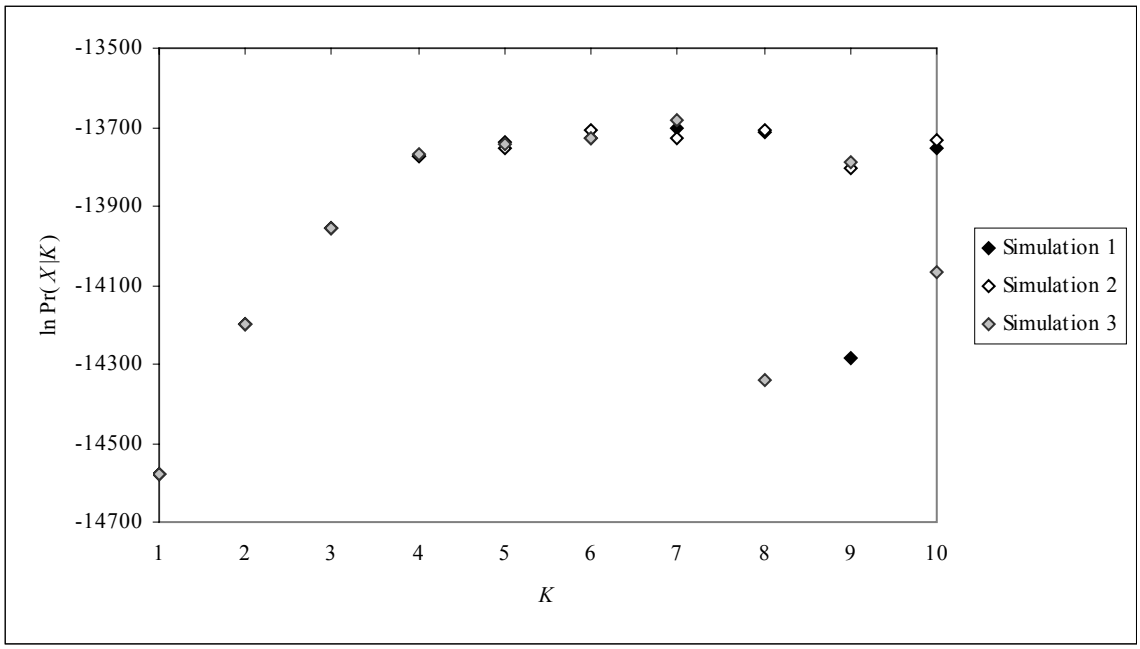


Figure 3.5. Values of lnPr($X|K$), the log-likelihood value of the data, as a function of the number of clusters (K = 1,2,…,10) inferred using *STRUCTURE* (Pritchard et al. 2000) in three independent simulation runs with all samples of *S. mentella* (604 individuals; 85 AFLP loci; model assumptions: correlated allele frequencies and no admixture) without using prior information of the origin of the samples.

Overall, the average posterior probabilities that a *S. mentella* sample originated from one of the four inferred clusters were relatively small, as a relatively high percentage of individuals (about 37%) revealed posterior probabilities of membership less than 0.80 in each of the four inferred clusters. About 63% of the individuals were assigned to one of the four inferred clusters with a posterior probability equal to or higher than 0.80. The highest percentage of individuals (76%) that could be assigned to one of the inferred clusters was observed in one of the samples of "deep-sea" *S. mentella* (MECIrmds97) and in the sample from the Flemish Cap (MEFC01). Table 3.8 shows that the sample from Norway (MENO01) and the sample from the southern Irminger Sea (MESIrm01) revealed some genetic similarities, as both showed a relatively high mean posterior probability of membership in cluster IV (0.68 and 0.71, respectively). Approximately half of the individuals of these two samples were assigned to this cluster. The two samples of "oceanic" *S. mentella*, MECIrmoc97 and MESIrmoc01, showed relatively high mean posterior probabilities of membership in cluster II (0.53 and 0.76, respectively), whereby MECIrmoc97 showed also membership in cluster IV (with a mean posterior probability of membership of 0.37). Several individuals of this sample were therefore assigned to the same cluster as approximately half of the individuals of the samples from Norway and the southern Irminger Sea. Also samples of "deep-sea" *S. mentella* from the central Irminger Sea (MECIrmds97 and MECIrmds01) and the sample from SW-Iceland (MESWIc01) revealed genetic similarities, as they showed relatively high mean posterior

probabilities of membership (0.78, 0.73 and 0.58, respectively) in the same cluster (III) - and a relatively high number of individuals with high posterior probabilities (≥ 0.80) of belonging to this cluster. Most of the individuals of the Flemish Cap sample were grouped into a separate cluster (cluster I). 24 out of 57 individuals of the sample from East Greenland (MEEGr01) were assigned to the same cluster as most of the individuals of the sample from the Flemish Cap. However, MEEGr01 displayed a high number of unassigned individuals, with low posterior probabilities of membership in each of the four inferred clusters. Also the sample from SE-Iceland revealed only small mean posterior probabilities of membership in any of the four clusters.

Table 3.8. Results of the Bayesian clustering analysis of *S. mentella* performed using *STRUCTURE* (Pritchard et al. 2000) with all samples of *S. mentella* pooled, without using prior information of population origin (604 individuals; 85 AFLP loci; model assumptions: correlated allele frequencies and no admixture). The number of genetic clusters in the sample set was set to $K = 4$ as it was the value at which $\ln\Pr(X|K)$ plateaued (see text and Figure 3.5). The table shows the mean posterior probabilities that a sample originated from the genetic clusters I, II, III, or IV. Each individual was assigned to a single cluster if the posterior probability that it originated from that cluster was equal to or larger than 0.80. The numbers of individuals assigned to each cluster with a posterior probability ≥ 0.80 are indicated in parentheses. If the posterior probabilities of membership of an individual in each of the four inferred clusters were smaller than 0.80, the individual was treated as unassigned individual. Sample codes are explained in Table 3.1.

		Inferred cluster				No. of unassigned	% assigned
Sample	*n*	I	II	III	IV	individuals	individuals
MENO01	59	0.125	0.134	0.059	0.682	23	61.02
		(1)	(2)	(1)	(32)		
MESEIc01	55	0.201	0.278	0.294	0.228	40	27.27
		(2)	(5)	(7)	(1)		
MESWIc01	60	0.075	0.230	0.576	0.120	20	66.67
		(0)	(8)	(30)	(2)		
MEEGr01	57	0.525	0.202	0.038	0.235	20	64.91
		(24)	(6)	(0)	(7)		
MECIrmds97	63	0.041	0.112	0.775	0.072	15	76.19
		(0)	(3)	(44)	(1)		
MECIrmds01	72	0.081	0.088	0.730	0.100	23	68.06
		(1)	(1)	(46)	(1)		
MECIrmoc97	64	0.035	0.534	0.058	0.372	22	65.63
		(0)	(27)	(0)	(15)		
MESIrmoc01	64	0.081	0.764	0.034	0.120	20	68.75
		(1)	(42)	(0)	(1)		
MESIrm01	48	0.149	0.096	0.043	0.713	23	52.08
		(0)	(1)	(1)	(23)		
MEFC01	62	0.812	0.119	0.020	0.049	15	75.81
		(43)	(3)	(0)	(1)		
Total	604	(72)	(98)	(129)	(84)	221	63.41

4.3. Genetic distances and UPGMA dendrogram

As the results of the model-based clustering analyses (see paragraph 4.2.1) grouped 24 individuals of the *S. marinus* sample MApool as well as eight individuals of the *S. mentella* sample from East Greenland (MEEGr01) into a distinct cluster, these 32 individuals were pooled and treated as separate sample (termed SEB) in the subsequent analyses.

Table 3.9 shows the pairwise Nei's (1978) genetic distance values between the different *Sebastes* samples. Nei's (1978) genetic distance values for pairwise comparisons between samples of the different species - as well as the sample SEB - ranged from 0.033 (between the pooled sample of *S. marinus*, MApool, and the sample of individuals assigned to the "*Sebastes sp.* cluster", SEB) to 0.200 (between the samples of *S. fasciatus* and *S. viviparus,* FAFC01 and VI).
Nei's (1978) genetic distance values for intraspecific pairwise comparisons between samples of *S. mentella* ranged from 0.003 (between the two samples of "deep-sea" *S. mentella*, MECIrmds97 and MECIrmds01) to 0.027 (between the sample of "deep-sea" *S. mentella* collected in the year 1997, MECIrmds97, and the sample of "oceanic" *S. mentella* collected in the year 2001, MESIrmoc01).

The corresponding UPGMA dendrogram of Nei's (1978) genetic distances (Figure 3.6) showed that the samples of the different North Atlantic *Sebastes* species, as well as the sample SEB, grouped in distinct well-supported clusters, supported by high bootstrap values (between 74 and 99) and 9 to 19 out of 85 polymorphic loci screened. The *S. viviparus* sample clustered apart from all other North Atlantic *Sebastes* samples. The *S. fasciatus* sample grouped with the *S. marinus* and the *S. mentella* samples (supported by 16 loci and a bootstrap value of 74), whereby the latter two clustered closer together (supported by a bootstrap value of 97 and 12 loci). The *S. marinus* sample from the Flemish Cap and the pooled sample of *S. marinus* (MApool) clustered together with high bootstrap support (bootstrap value: 95; 19 loci supported this cluster). The sample of individuals assigned to the "*Sebastes sp.* cluster" in the model-based clustering analysis (see paragraph 4.2.1) grouped with the two *S. marinus* samples, supported by a high bootstrap value of 92 and 9 loci.
All samples of *S. mentella* formed a distinct cluster, supported by 16 loci and a high bootstrap value of 99. Within the *S. mentella* cluster, the two samples of "deep-sea" *S. mentella* from the central Irminger Sea (sampled in the years 1997 and 2001) and the *S. mentella* sample from SW-Iceland were placed on a separate branch (supported by three loci and a high bootstrap value of 98), whereas all other samples of *S. mentella* (samples from Norway, SE-Iceland, Greenland, the central Irminger Sea, the southern Irminger Sea and the Flemish Cap) clustered together, with small branch lengths and no bootstrap support at the corresponding nodes (bootstrap values < 50).

Table 3.9. Pairwise Nei's (1978) genetic distances (above diagonal) and pairwise Φ_{ST} values (below diagonal; see paragraph 4.4) between the different samples of North Atlantic *Sebastes*. Sample codes are explained in Table 3.1. SEB = sample of the 32 individuals assigned to cluster III (the "*Sebastes sp.* cluster") in the *STRUCTURE* analysis (see paragraph 4.2.1). All pairwise Φ_{ST} values were significantly different from zero ($P < 0.001$) after applying the sequential Bonferroni adjustment (Rice 1989).

	VI	FAFC01	MApool	MAFC01	SEB	MENO01	MESEIc01	MESWIc01	MEEGr01	MECIrmds97
VI		0.2004	0.1448	0.1632	0.1585	0.1377	0.1359	0.1298	0.1387	0.1203
FAFC01	0.5860		0.1496	0.1075	0.1700	0.1253	0.1166	0.1099	0.1150	0.1135
MApool	0.4818	0.4884		0.0209	0.0328	0.0787	0.0672	0.0678	0.0796	0.0682
MAFC01	0.4903	0.4224	0.0750		0.0593	0.0931	0.0707	0.0748	0.0876	0.0757
SEB	0.5058	0.5373	0.2004	0.2513		0.0642	0.0608	0.0598	0.0668	0.0630
MENO01	0.5047	0.4668	0.3665	0.3747	0.3451		0.0110	0.0128	0.0083	0.0161
MESEIc01	0.5053	0.4307	0.3285	0.3204	0.3268	0.0710		0.0099	0.0098	0.0148
MESWIc01	0.4838	0.4111	0.3232	0.3186	0.3182	0.0944	0.0528		0.0131	0.0059
MEEGr01	0.4816	0.4281	0.3602	0.3611	0.3249	0.0786	0.0797	0.1015		0.0182
MECIrmds97	0.4464	0.4049	0.3072	0.3093	0.3126	0.1014	0.0661	0.0350	0.1148	
MECIrmds01	0.4685	0.3990	0.3245	0.3225	0.3125	0.1056	0.0532	0.0401	0.1065	0.0284
MECIrmoc97	0.5353	0.4646	0.3530	0.3596	0.3511	0.0936	0.0722	0.1014	0.1149	0.1087
MESIrmoc01	0.5416	0.4581	0.3659	0.3602	0.3583	0.1132	0.0707	0.1040	0.1011	0.1325
MESIrm01	0.5230	0.4910	0.3921	0.4085	0.3878	0.0312	0.0945	0.1224	0.0909	0.1181
MEFC01	0.4781	0.3280	0.3185	0.3015	0.3311	0.0881	0.0696	0.0884	0.1056	0.1057

Table 3.9. (continued)

	MECIrmds01	MECIrmoc97	MESIrmoc01	MESIrm01	MEFC01
VI	0.1284	0.1421	0.1516	0.1319	0.1382
FAFC01	0.1096	0.1174	0.1191	0.1292	0.0847
MApool	0.0694	0.0660	0.0737	0.0780	0.0768
MAFC01	0.0761	0.0789	0.0770	0.0955	0.0751
SEB	0.0612	0.0574	0.0671	0.0697	0.0754
MENO01	0.0164	0.0118	0.0158	0.0072	0.0184
MESEIc01	0.0126	0.0137	0.0088	0.0123	0.0146
MESWIc01	0.0049	0.0176	0.0162	0.0165	0.0203
MEEGr01	0.0160	0.0125	0.0122	0.0080	0.0147
MECIrmds97	0.0032	0.0203	0.0265	0.0175	0.0261
MECIrmds01		0.0197	0.0209	0.0181	0.0234
MECIrmoc97	0.1076		0.0115	0.0094	0.0184
MESIrmoc01	0.1161	0.0669		0.0159	0.0147
MESIrm01	0.1278	0.0855	0.1169		0.0167
MEFC01	0.0985	0.0931	0.0901	0.1020	

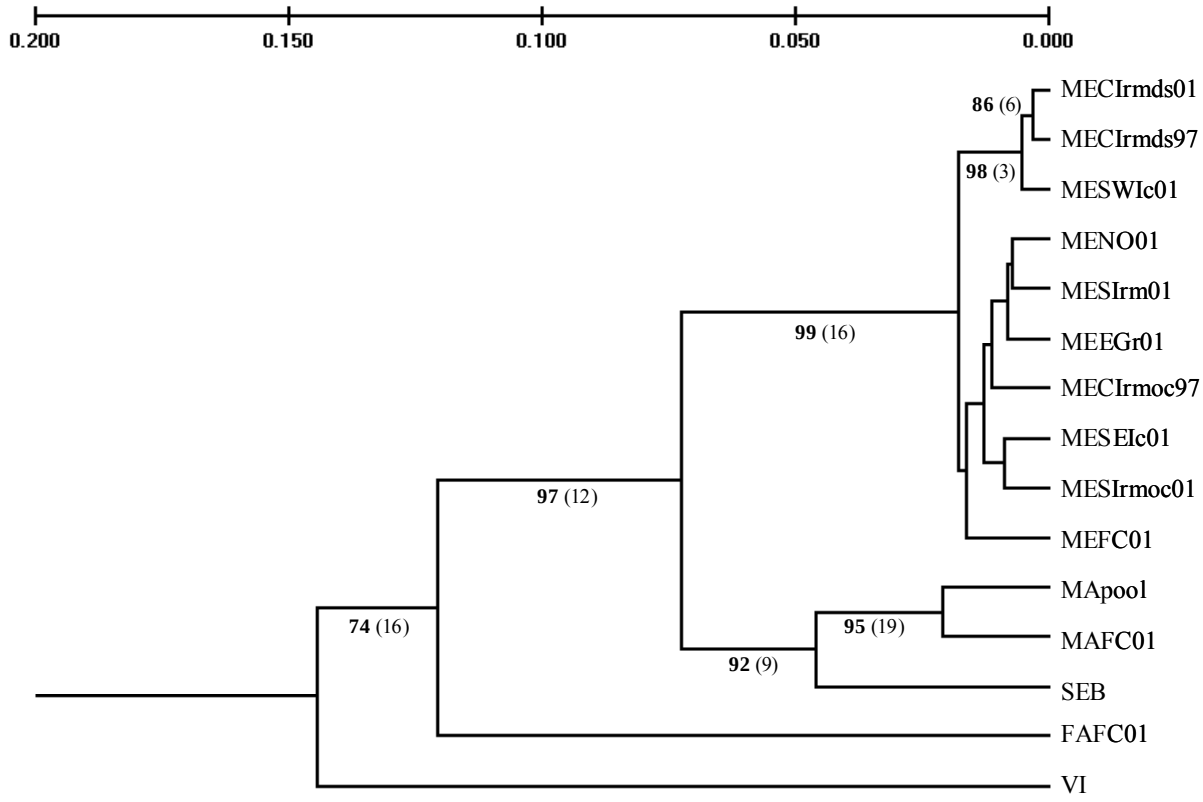


Figure 3.6. UPGMA dendrogram constructed on the basis of Nei's (1978) genetic distances between samples of North Atlantic *Sebastes,* estimated from AFLP data (85 loci). Bold values represent bootstrap support values of 1000 replicates. Only bootstrap values greater than 50 are shown. Numbers in parentheses represent number of loci supporting the corresponding node. Sample codes are explained in Table 3.1. SEB = sample of the 32 individuals assigned to cluster III (the "*Sebastes sp.* cluster") in the *STRUCTURE* analysis (see paragraph 4.2.1).

4.4 AMOVA and pairwise Φ_{ST} values

The results of the different hierarchical analyses of molecular variance (AMOVA) are given in Table 3.10 and 3.11, respectively.

The results of the first AMOVA, analysing the degree of differentiation among the genetic clusters estimated by the *STRUCTURE* analysis (see paragraph 4.3.1) relative to the degree of differentiation among samples within these clusters (see Table 3.10) revealed that 60.97% of the total genetic variance was found within the samples. This variance component was highly significant ($P < 0.001$). The divergence among the five clusters significantly ($P < 0.001$) accounted for 33.79% of the total genetic variance. The smallest but still highly significant part (5.24%) of the total variance was due to differences between samples within clusters.

Table 3.10. Results of the analysis of molecular variance (AMOVA) with genetic variance partitioned among clusters inferred by *STRUCTURE* (Pritchard et al. 2000), samples within clusters and within samples (d.f. = degrees of freedom). *** = highly significant, P < 0.001

Source of variation	d.f.	Variance components	Percentages of variation
Among clusters	3	4.11189	33.79***
Among samples within clusters	10	0.63785	5.24***
Within samples	824	7.42092	60.97***
Total	837	12.1707	100.00

The second AMOVA (Table 3.11), partitioning the variance among temporal samples of *S. mentella* from the same location relative to the variance among *S. mentella* samples from different locations, revealed that the largest part of the total variance (91.99%) was found within the samples. A smaller but still significant part (5.00%) of the total genetic variation was distributed among samples from different locations. 3.01% of the total variation contributed to differences between temporal samples from the same location. This "temporal" variance component was highly significant (P < 0.001).

Table 3.11. Results of the analysis of molecular variance (AMOVA) of the *S. mentella* samples with variance partitioned among locations, among temporal samples from the same locations and within samples (d.f. = degrees of freedom). * = significant, P < 0.05; *** = highly significant, P < 0.001

Source of variation	d.f.	Variance components	Percentages of variation
Among locations	7	0.395	5.00*
Among temporal samples from the same location	2	0.238	3.01***
Within samples	594	7.266	91.99***
Total	603	7.899	100.00

Table 3.9 (below diagonal) shows matrices of pairwise Φ_{ST} values and corresponding probability (P-) values between the North Atlantic *Sebastes* samples. All pairwise Φ_{ST} values between samples of the four species *S. viviparus, S. fasciatus, S. marinus* and *S. mentella*, as well as the sample of individuals assigned to cluster III (the "*Sebastes sp.* cluster") in the *STRUCTURE* analysis (sample SEB; see paragraph 4.2.1), were highly significantly different from zero (P < 0.001) and varied between 0.200 (between the pooled sample of S. *marinus*,

MApool, and the sample SEB) and 0.586 (between the samples of *S. fasciatus* and *S. viviparus,* FAFC01 and VI).

Overall, the degree of genetic differentiation between samples of *S. mentella* from different locations and years was smaller than the degree of genetic differentiation between samples of the different species. Significant Φ_{ST} values varied between 0.028 (between the two samples of "deep-sea" *S. mentella*, MECIrmds97 and MECIrmds01) and 0.133 (between the sample of "deep-sea" *S. mentella* from the central Irminger Sea 1997 and the sample of "oceanic" *S. mentella* from the southern Irminger Sea 2001, MECIrmds97 and MESIrmoc01). All pairwise comparisons were highly significantly different from zero ($P < 0.001$).

4.5. Isolation by distance

No significant correlation ($r = 0.0781$, $P < 0.34$) between geographic and genetic distance (Nei 1978) was observed for samples of *S. mentella* at the spatial scale of this study (from Norway to the Flemish Cap; see Figure 3.7).

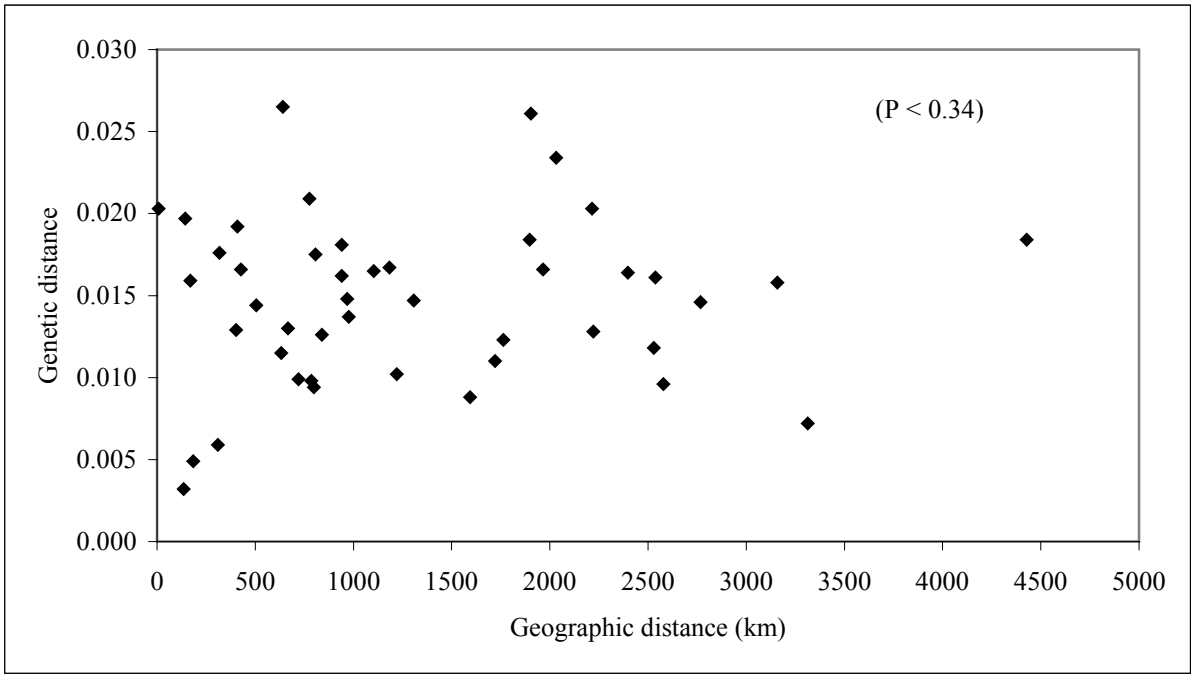


Figure 3.7. Plot of pairwise genetic distances - quantified as Nei's (1978) genetic distance - versus pairwise geographic distances (in km) for 10 samples of *S. mentella*. Significance of the correlation was estimated with Mantel's test (r = standardised Mantel test relationship coefficient).

5. Discussion

5.1 Levels of polymorphism and genetic diversity

In general, the levels of polymorphism observed in the present study were relatively low. Out of 26 primer pairs initially used to survey polymorphisms within samples of North Atlantic *Sebastes*, only seven generated polymorphic AFLP fragments. The remaining primer pairs generated none or only one polymorphic AFLP fragment. Nevertheless, the total number of 85 polymorphic AFLP marker loci - with an average number of 48 polymorphic marker loci at the 5% level per sample - obtained by these seven primer pairs was comparable to the results of a study by Kai et al (2002a), who studied genetic differences among three morphotypes of the black rockfish, *Sebastes inermis*, using six primer pair combinations (*Eco*RI and *Mse*I primers). Furthermore, the number of polymorphic AFLP markers observed in the present study is relatively high compared to other AFLP-based studies of genetic polymorphisms in marine fish species such as *Thynnus* and *Morone* (Han and Ely 2002). It has to be noted that a comparison of the results of the different AFLP studies is not unproblematic, because the restriction enzymes, primer pairs and markers used differ among studies and laboratories.

All samples of North Atlantic *Sebastes* revealed similar - relatively low - levels of gene diversity with an average H_j of 0.21. In contrast, the average gene diversity H_e (analogous to H_j) observed in the microsatellite analysis (see Chapter 2) was much higher (average expected mean heterozygosity within each sample of North Atlantic *Sebastes* ranged from 0.74 to 0.88). This is - at least partly - caused by the fact that AFLP loci have a much lower number of alleles, with two alleles at any locus (Yan et al. 1999), than microsatellite loci. The microsatellite loci used in this thesis revealed between 23 and 107 alleles per locus in samples of North Atlantic *Sebastes* (see Chapter 2).

5.2 Species structure of North Atlantic redfish

The results of the Bayesian cluster analysis of the AFLP data set using *STRUCTURE* (Pritchard et al. 2000) revealed that the redfish specimens, pre-classified morphologically into the four currently known North Atlantic *Sebastes* species, clustered into five genetically distinct groups, even though inference of the number of genetic clusters (K) was not straightforward with the AFLP data set (as for example with the microsatellite data set; see Chapter 2). However, the posterior probabilities $\ln\Pr(X|K)$ reached a plateau between $K = 5$ and $K = 6$ and nearly all individuals were strongly assigned to one of the five clusters (assuming $K = 5$). This indicates an overall pattern of five gene pools in the data set (according to the guidelines proposed by Pritchard and Wen 2003). Most of the individuals of *S. viviparus*, *S. fasciatus* and *S. mentella* grouped into separate distinct clusters, in agreement

with the morphological species identification (see Table 3.1), whereas the samples of *S. marinus* clustered into two genetically distinct groups. As all individuals of the *S. marinus* sample from Norway and the sample of "giant" *S. marinus* were assigned to the same cluster as the *S. marinus* sample from the Flemish Cap, it can be assumed that this cluster represented the species *S. marinus* (the cluster was therefore designated "*S. marinus* cluster"). In contrast, *S. marinus* samples from Iceland and Greenland revealed two groups of individuals: Some individuals were assigned to the "*S. marinus* cluster", but several individuals were grouped in a separate cluster, together with eight individuals of the *S. mentella* sample from Greenland. Due to its genetic distinctness, this cluster was designated "*Sebastes sp.* cluster". The high degree of genetic differentiation among the five clusters inferred by *STRUCTURE* is reflected in high pairwise Nei's (1978) genetic distances and high pairwise Φ_{ST} values. Also the results of the UPGMA cluster analysis supported the genetic distinctness of the five clusters, as the samples of individuals assigned to the "*S. viviparus* cluster", the "*S. fasciatus* cluster", the "*S. mentella* cluster", the "*S. marinus* cluster" and the "*Sebastes sp.* cluster" clustered in distinct clades, supported by high bootstrap values, several AFLP loci and long branch lengths.

The high degree of genetic differentiation among the five clusters was reflected in the results of the corresponding AMOVA (Table 3.10), which revealed that a very high amount of the total genetic variance was imputable to differences between the genetic clusters. This variance component was about six times higher than the genetic variance among samples within clusters.

The results of the AFLP analysis are in congruence with the results of the analyses of mitochondrial ND3 gene sequence variation and microsatellite variation presented in Chapter 1 and Chapter 2 of this thesis, respectively. The results of these analyses indicated that the samples of *S. marinus* collected on the shelves of Greenland, Iceland and the Faroe Islands consisted of two genetically distinct groups of individuals - separated by similar levels of genetic differentiation as the four known North Atlantic *Sebastes* species. This points to the existence of a fifth cryptic *Sebastes* species in these areas.

Furthermore, the genetic relationships between the species revealed by the UPGMA analysis of the AFLP data support the results of the ND3 gene sequence analysis, which indicated that within the North Atlantic *Sebastes* species complex, *S. viviparus* is the most distant species and that *S. mentella* is closely related to *S. marinus*. This is further supported by the results of other previous genetic studies (e.g. McGlade et al. 1983; Sundt and Johansen 1998).

However, the AFLP analysis revealed smaller levels of genetic differentiation between the samples comprising the "*S. marinus* cluster" (the samples MApool and MAFC01) and the sample comprising the "*Sebastes sp.* cluster" (sample SEB) - representing presumably different species of *Sebastes* - than the ND3 gene analysis and the microsatellite analysis. Consequently, the UPGMA analysis grouped the sample SEB together with the two *S. marinus* samples. This could have been caused by the fact that only few individuals of the "*Sebastes sp.* cluster" were present in the AFLP sample set (= too low sample size). The

smaller genetic differences and the clustering in the UPGMA dendrogram can also be explained by the fact that AFLP markers are anonymous markers. The genetic basis of the genetic differences is unknown, because AFLP fragments are identified by their length and not by their base composition. Even though it is generally assumed that co-migrating bands are homologous, there is no *a priori* reason to accept this (Robinson and Harris 1999). Therefore, non-identical fragments of equal length can mistakenly be scored as identical, thus underestimating genetic differentiation. This could explain the smaller genetic differences between the samples comprising the "*S. marinus* cluster" and the sample comprising the "*Sebastes sp.* cluster" observed in the AFLP analysis. However, even though the genetic differences found between the "*Sebastes sp.* cluster" and the "*S. marinus* cluster" were not as high as the genetic differences (pairwise Nei's genetic distances as well as Φ_{ST} values) observed between the other four clusters (the "*S. viviparus* cluster", the "*S. fasciatus* cluster", the "*S. marinus* cluster" and the "*S. mentella* cluster"), they were much higher than the intraspecific differences observed between samples of *S. marinus* and *S. mentella*, respectively.

The results obtained from the AFLP analysis therefore provide evidence that the currently recognised species *S. mentella*, *S. fasciatus* and *S. viviparus* are genetically distinct and represent valid species. Furthermore, the results of the AFLP analysis strongly support the existence of cryptic speciation in North Atlantic *Sebastes*, as indicated by the results of the mitochondrial ND3 gene sequence analysis (see Chapter 1) and the microsatellite analysis (see Chapter 2) presented in this thesis.

5.3 Precision of species assignment

The five genetic clusters were genetically so distinct from each other that it was possible to assign 99% (average over all samples) of the individuals to one of the five clusters with a posterior probability of membership ≥ 0.80, using Bayesian cluster analysis without using prior information about the species and sample of origin. Only the individuals of the *S. marinus* sample MApool were assigned less accurately on average, as five individuals could not be unambiguously assigned to one of the five clusters. Therefore, AFLP analysis was found to be highly effective in distinguishing species of North Atlantic *Sebastes*.

A comparison of the assignment results using AFLP data, microsatellite data and the pre-classification into species on board using external morphological, morphometrical and meristical traits (see paragraph 3.1) revealed that in case of the samples of *S. viviparus*, *S. fasciatus* and *S. mentella* the results of the assignment analyses based on AFLP data were in congruence with the results of the microsatellite analysis and the pre-classification on board. Most of the specimens sampled and pre-classified as *S. viviparus*, *S. fasciatus* and *S. mentella*, respectively, were assigned to the same cluster (the "*S. viviparus* cluster", the "*S. fasciatus* cluster" and the "*S. mentella* cluster", respectively). Only very few individuals of these

samples were assigned to other genetic clusters (see Table 3.6). Comparisons with the results of the assignment analyses using microsatellite data support the assignment using AFLP data, indicating that these few individuals were most likely misclassified or mislabelled during sampling: E.g. the four individuals of *S. marinus* collected on the Flemish Cap, which were assigned to the "*S. fasciatus* cluster", as well as the three individuals of the *S. fasciatus* sample, which were assigned to the "*S. marinus* cluster" (two individuals) and the "*S. mentella* cluster" (one individual), were most likely misclassified during sampling, supported by the results of microsatellite analysis (see Chapter 2). The same applies to the 12 individuals pre-classified as *S. mentella* that were not assigned to the "*S. mentella* cluster": Also these individuals were most likely misclassified on board, as the results of microsatellite analysis revealed the same assignment results as the AFLP analysis[1].

However, there were a few examples, where AFLP-based species assignment was not supported by the microsatellite-based species assignment: The three individuals of the *S. viviparus* sample assigned to the "*S. mentella* cluster" (see Table 3.6) and five of the seven individuals of the *S. marinus* sample from the Dohrn Bank (MADB01; see Table 3.7) assigned to the "*Sebaste*s *sp.* cluster" in the AFLP-based assignment analysis were assigned to the "*S. viviparus* cluster" and the "*S. marinus* cluster", respectively, when microsatellite data were used in the assignment analysis (see Table 2.8 in Chapter 2). Furthermore, a single individual of the sample MESIrm01 was assigned to the "*S. mentella* cluster" in the AFLP analysis, but to the "*Sebaste*s *sp.* cluster" in the microsatellite analysis.

In particular the three individuals of the *S. viviparus* sample assigned to the "*S. mentella* cluster" were most likely misassigned in the AFLP analysis, as morphological pre-classification, microsatellite (see Table 2.8 in Chapter 2) and ND3 gene analyses (see Table 1.7 in Chapter 1) did not support the results of the AFLP analysis. The same applies to the other individuals mentioned above, as microsatellite and/or ND3 gene analysis did not support the results of the AFLP analysis. The misassignment was most likely caused by the fact that in the AFLP analysis too few specimens of *S. viviparus* and the "*Sebaste*s *sp.* cluster" (the newly identified potential cryptic species of North Atlantic *Sebastes*) were analysed.

However, except these few presumably misassigned individuals and 1% of unassigned individuals, the high assignment success observed in the present study demonstrates the high potential of AFLP analysis in North Atlantic *Sebastes* species identification.

5.4 Patterns of intraspecific genetic structure in S. mentella

The results of the AFLP analysis of the *S. mentella* samples from different locations throughout the North Atlantic indicated a possible genetic structure in *S. mentella*.

[1] Only the single individual of *S. mentella* from the Flemish Cap assigned to the "*S. marinus* cluster" was not analysed using microsatellite analysis. Nevertheless, this individual represented a mislabelled individual sample of *S. marinus* (Garabana pers. comm. 2003).

However - as in the analysis of interspecific genetic structure - precise inference of the number of populations *K* in the *S. mentella* data set using model-based cluster analysis (*STRUCTURE*; Pritchard et al. 2000) was difficult. The fact that the posterior probabilities lnPr(*X*|*K*) reached a plateau between *K* = 4 and *K* = 5 (see Figure 3.5) indicated that *K* = 4 was the smallest value of *K* that captured the major structure in the data and therefore the most appropriate number of populations in the data set (according the guidelines given by Pritchard and Wen 2003). However, in contrast to the analysis of interspecific structure, only 63% of the individuals could be assigned to one of the four intraspecific clusters/populations inferred by *STRUCTURE* (Table 3.8), most likely because the level of genetic differentiation was too low. Furthermore, as AFLP markers usually must be treated as dominant markers (Vos and Kuiper 1997), the "no admixture" model had to be applied (according to Pritchard and Wen 2003), which assumes no gene flow between populations. This model assumption may be appropriate on the interspecific level, but may not be appropriate on the intraspecific level, as it is generally assumed that populations of *S. mentella* are closely related with presumably high levels of gene flow (Roques et al. 2002). Therefore, care should be taken in the interpretation of the inferred value of *K* = 4, as clusters may not necessarily correspond to the "real" population structure.

Nevertheless, the observation of genetic structure in *S. mentella* was supported by the results of the AMOVA, indicating that a significant part of the total genetic variation was due to genetic differentiation between samples from different locations. Furthermore, the overall pattern of genetic structure indicated by the results of the *STRUCTURE* analysis was reflected in the results of the UPGMA cluster analysis, which indicated that the two samples of "deep-sea" *S. mentella* from the central Irminger Sea (sampled in the years 1997 and 2001) and the *S. mentella* sample from SW-Iceland were genetically similar and distinct from all other samples of *S. mentella*. This suggests that these three samples represent a distinct gene pool of *S. mentella*. In contrast, the genetic differentiation between the remaining samples of *S. mentella* was less pronounced, as they all grouped together in the UPGMA tree, without bootstrap support at any of the corresponding nodes. Nevertheless, some groupings were consistent among UPGMA cluster analysis and *STRUCTURE* analysis: The separate clustering of the Flemish Cap sample in the UPGMA tree (although not supported by the bootstrap analysis) could indicate that also the sample from the Flemish Cap was genetically distinct from the other samples. This was also indicated by the results of the *STRUCTURE* analysis. Also the genetic similarity of the samples from Norway and the southern Irminger Sea was indicated by both UPGMA and *STRUCTURE* analysis.

The major genetic pattern revealed by the results of the AFLP analysis, indicating that two samples of "deep-sea" *S. mentella* and the sample of *S. mentella* collected off SW-Iceland were genetically different from the other *S. mentella* samples analysed, was supported by the results of the ND3 gene analysis. The results of the ND3 gene analysis revealed that a sample of "deep-sea" *S. mentella* differed significantly in haplotype frequencies from all other

samples of *S. mentella*. Even though it has to be noted that the ND3 gene analysis was based on a relatively small number of individuals of "deep-sea" *S. mentella* and these results must therefore be interpreted with caution, the congruence between the two methods could indicate that the genetic signal is relatively robust (as it was displayed by a mitochondrial marker and a nuclear marker system). This is further supported by a study by Johansen et al. (2000b), who studied allozyme and hemoglobin variation in *S. mentella* and found indication that "deep-sea" *S. mentella* could represent a genetically isolated population of *S. mentella*. However, as outlined in detail in Chapter 1 (paragraph 5.3.4), there has been a strong controversy about whether the different *S. mentella* phenotypes represent "valid" phenotypes, as the differences are subtle and only few experts are able to distinguish between them (ICES 1998a; Johansen et al. 2000b; Saborido-Rey et al. 2005). As discussed in Chapter 1, it cannot be ruled out that "deep-sea" *S. mentella* simply represent an older age group. However, when the results of the different studies presented in this thesis are compared without considering the different phenotypes, the results of the ND3 gene analysis and the AFLP analysis are in congruence with the results of the microsatellite analysis (see Chapter 2). The microsatellite analysis revealed that two samples of *S. mentella* from the central Irminger Sea - consisting of relatively large and probably also old individuals - were genetically different from samples of *S. mentella* from other regions throughout the North Atlantic.

Therefore, there are two possible explanations for the observed genetic structure in *S. mentella*: Differences between different year-classes (due to sweepstakes chance effects; Hedgecock 1994; see also Saborido-Rey et al. 2005) and population separation due to restricted adult dispersal and/or larval transport (Magnússon and Magnússon 1995). Both hypotheses are discussed in detail in Chapter 2 and in the FINAL DISCUSSION chapter.

The results of the AFLP analysis also indicated that *S. mentella* on the Flemish Cap could represent a distinct population. This is in congruence with the result of the microsatellite analysis (see Chapter 2) - and with the general assumption that the redfish populations of the Flemish Cap are relatively isolated (Templeman 1976).

However, some of the results of the AFLP analysis are difficult to interpret. The fact that the analysis of pairwise Φ_{ST} values revealed highly significant differences between all samples of *S. mentella* suggests that AFLP markers may be too variable to detect biological meaningful patterns of genetic structure in more weakly differentiated populations of *S. mentella*. This is supported by the fact that high genetic differences were found between several samples from closely located sampling locations (e.g. SW- and SE-Iceland), whereas only small genetic differences where found between samples from very distant locations (e.g. Norway and the southern Irminger Sea).

Furthermore, the results of the AMOVA revealed temporal instability, as a significant part of the total variation was due to differences among temporal samples from the same location. Due to the observed temporal instability it cannot be ruled out that the observed population structure is an ephemeral phenomenon. The number of temporal genetic samples in this study

was too limited, in particular considering the longevity of the species and its high potential for dispersal. Therefore, larger sample sizes, sampling over a longer period of time and additional molecular markers (such as microsatellites) are needed before definite conclusions about the population structure of *S. mentella* in the Irminger Sea and adjacent waters can be drawn.
Nevertheless, the major genetic patterns found in *S. mentella* using AFLP analysis are congruent with the patterns found in other genetic studies. However, the AFLP technique has a significant drawback. The dominance of the AFLP markers significantly decreases the genetic information available and makes it impossible to estimate standard population parameters such as deviations from Hardy-Weinberg expectations and observed heterozygosity (Vos and Kuiper 1997).
Therefore, co-dominant markers such as microsatellites are presumably more informative in studies on the population structure of *S. mentella*.

6. General conclusions and perspectives

The present study has demonstrated the usefulness of polymorphic AFLP markers for gaining insight into the species structure of North Atlantic *Sebastes*. The AFLP analysis has provided evidence that the currently recognised species *S. mentella*, *S. fasciatus* and *S. viviparus* are genetically distinct and therefore indeed represent valid species. The results of the AFLP analysis also support the existence of cryptic speciation in North Atlantic *Sebastes*, as indicated by the results of the two previous studies presented in this thesis employing mitochondrial DNA (see Chapter 1) and microsatellite markers (Chapter 2).
Furthermore, the results of this study have provided insights into the patterns of population structure in *S. mentella*. Even though the significant genetic differences observed between almost all samples suggest that AFLP markers may be too variable to detect biological meaningful patterns of genetic structure in more weakly differentiated populations of *S. mentella*, the major genetic patterns identified in *S. mentella* using AFLP analysis are congruent with the patterns revealed by other methods, indicating that AFLP analysis could indeed be a useful tool in population genetic studies of *S. mentella*.
However, as the dominant nature of AFLP markers clearly hampers their use for estimating standard population genetic parameters (e.g. deviations from Hardy-Weinberg expectations, observed heterozygosity), AFLP markers are unlikely to outcompete co-dominant markers, such as microsatellites, which clearly allow more powerful population genetic analyses (Mueller and Wolfenbarger 1999). Therefore, the results of the present study suggest that AFLP may be useful for species level investigations of North Atlantic *Sebastes*, but for population level investigations co-dominant markers such as microsatellites are preferable.

Appendix I

Summary statistics for redfish samples collected for AFLP analysis showing species names, morphological types, sampling locations, position of the sampling locations, sampling years, sampling months, average sampling depths, station numbers, coordinating institutes/nations, cruise identification numbers, sample codes and sample sizes (*n*).

			Position (mean)			
Species	**Type**	**Location**	**Longitude**	**Latitude**	**Year**	**Month**
S. mentella		Norway	711180N	0240550E	2001	October
S. mentella		Norway	704260N	0182660E	2001	October
S. mentella		SE-Iceland	6353000	0115200W	2001	October
S. mentella		SE-Iceland	632940N	0114120W	2001	October
S. mentella		SE-Iceland	631270N	0121140W	2001	October
S. mentella		SE-Iceland	630080N	0124860W	2001	October
S. mentella		SE-Iceland	634610N	0124000W	2001	October
S. mentella		SW-Iceland	6328700	0261700W	2001	October
S. mentella		SW-Iceland	6339800	0261000W	2001	October
S. mentella		SW-Iceland	6344200	0262760W	2001	October
S. mentella		Greenland East	635189N	0363407W	2001	October
S. mentella		Greenland East	634840N	0364037W	2001	October
S. mentella		Greenland East	633839N	0372445W	2001	October
S. mentella	deep-sea	Central Irminger Sea	613810N	0301160W	1997	May
S. mentella	deep-sea	Central Irminger Sea	622000N	0280500W	2001	June
S. mentella	deep-sea	Central Irminger Sea	622490N	0280960W	2001	June
S. mentella	oceanic	Central Irminger Sea	613720N	0302090W	1997	May
S. mentella	oceanic	Southern Irminger Sea	575970N	0390360W	2001	June
S. mentella	oceanic	Southern Irminger Sea	572580N	0420040W	2001	June
S. mentella		Southern Irminger Sea	571483N	0409316W	2001	June
S. mentella		Flemish Cap	465203N	0447613W	2001	July
S. mentella		Flemish Cap	465653N	0457677W	2001	July
S. mentella		Flemish Cap	465155N	0459890W	2001	July
S. mentella		Total				
S. marinus		Norway	682030N	0111550E	2001	October
S. marinus		SW-Iceland	630100N	0250800W	1997	May
S. marinus		SE-Iceland	640143N	0131375W	2001	October
S. marinus		Dohrn Bank	653457N	0300652W	2001	October
S. marinus		Greenland East	641929N	0360459W	2001	October
S. marinus		Greenland East	641664N	0355754W	2001	October
S. marinus	giant	Irminger Sea	604400N	0282000W	1996	August
S. marinus	giant	Irminger Sea	604400N	0282000W	1996	August

Appendix I (continued)

[1] Federal Research Centre for Fisheries (BFAFi), Hamburg, Germany; [2] Institute of Marine Research, Vigo, Spain; [3] Marine Research Institute (MRI), Reykjavík, Iceland; [4] Institute for Marine Research (IMR) and the University of Bergen (UIB), Department of Fisheries and Marine Biology, Bergen, Norway

Depth (m)	Station	Institute	Cruise id	Sample code	*n*
340	527	Norway[4]	JH-81516	MENO01	45
360	535	Norway[4]	JH-81516	MENO01	14
371	573	Iceland[3]	A12-2000	MESEIc01	2
385	574	Iceland[3]	A12-2001	MESEIc01	16
412	575	Iceland[3]	A12-2001	MESEIc01	16
565	576	Iceland[3]	A12-2001	MESEIc01	12
546	583	Iceland[3]	A12-2001	MESEIc01	9
784	410	Iceland[3]	A12-2001	MESWIc01	23
460	411	Iceland[3]	A12-2001	MESWIc01	10
613	412	Iceland[3]	A12-2001	MESWIc01	27
321	1009	Germany[1]	WH233	MEEGr01	23
363	1010	Germany[1]	WH233	MEEGr01	20
255	1011	Germany[1]	WH233	MEEGr01	14
750	9	Iceland[3]	BTH11997	MECIrmds97	63
775	274	Iceland[3]	A8-2001	MECIrmds01	23
640	275	Iceland[3]	A8-2001	MECIrmds01	49
725	11	Iceland[3]	BTH11997	MECIrmoc97	64
200	308	Iceland[3]	A8-2001	MESIrmoc01	39
250	311	Iceland[3]	A8-2001	MESIrmoc01	25
159	704	Germany[1]	WH229	MESIrm01	48
314	15	Spain[2]	CAFC01	MEFC01	20
351	24	Spain[2]	CAFC01	MEFC01	25
417	25	Spain[2]	CAFC01	MEFC01	17
					604
195	560	Norway[4]	JH-81549	MANO01	10
630	1	Iceland[3]	BTH11997	MASWIc97	12
308	881	Iceland[3]	B132001	MASEIc01	10
364	974	Germany[1]	WH233	MADB01	10
256	995	Germany[1]	WH233	MAEGr01	2
278	996	Germany[1]	WH233	MAEGr01	8
732	49	Iceland[3]	TJ11996	MAG96	10
786	50	Iceland[3]	TJ11996	MAG96	2
				MApool	64

Appendix I (continued)

Summary statistics for redfish samples collected for AFLP analysis showing species names, morphological types, sampling locations, position of the sampling locations, sampling years, sampling months, average sampling depths, station numbers, coordinating institutes/nations, cruise identification numbers, sample codes and sample sizes (*n*).

			Position (mean)			
Species	**Type**	**Location**	**Longitude**	**Latitude**	**Year**	**Month**
S. marinus		Flemish Cap	472712N	0457550W	2001	July
S. marinus		Flemish Cap	476687N	0448737W	2001	July
S. marinus		Flemish Cap	476315N	0449568W	2001	July
S. marinus		Flemish Cap	469425N	0447085W	2001	July
S. marinus		Flemish Cap	472643N	0447045W	2001	July
S. marinus		Flemish Cap	474342N	0445983W	2001	July
S. marinus		Flemish Cap	476645N	0441683W	2001	July
S. marinus		Flemish Cap	474452N	0443257W	2001	July
S. marinus		Total				
S. fasciatus		Flemish Cap	469313N	0440967W	2001	July
S. fasciatus		Flemish Cap	466842N	0453872W	2001	July
S. fasciatus		Flemish Cap	466842N	0453872W	2001	July
S. fasciatus		Flemish Cap	479502N	0447565W	2001	July
S. fasciatus		Flemish Cap	476687N	0448737W	2001	July
S. fasciatus		Total				
S. viviparus		Norway	701250N	0305630E	2001	October
S. viviparus		Norway	704790N	0293710E	2001	October
S. viviparus		Norway	710560N	0283650E	2001	October
S. viviparus		Norway	711390N	0273140E	2001	October
S. viviparus		SW-Iceland	632100N	0253900W	1996	November
S. viviparus		SW-Iceland	635998N	0233273W	2001	March
S. viviparus		Total				
Total						

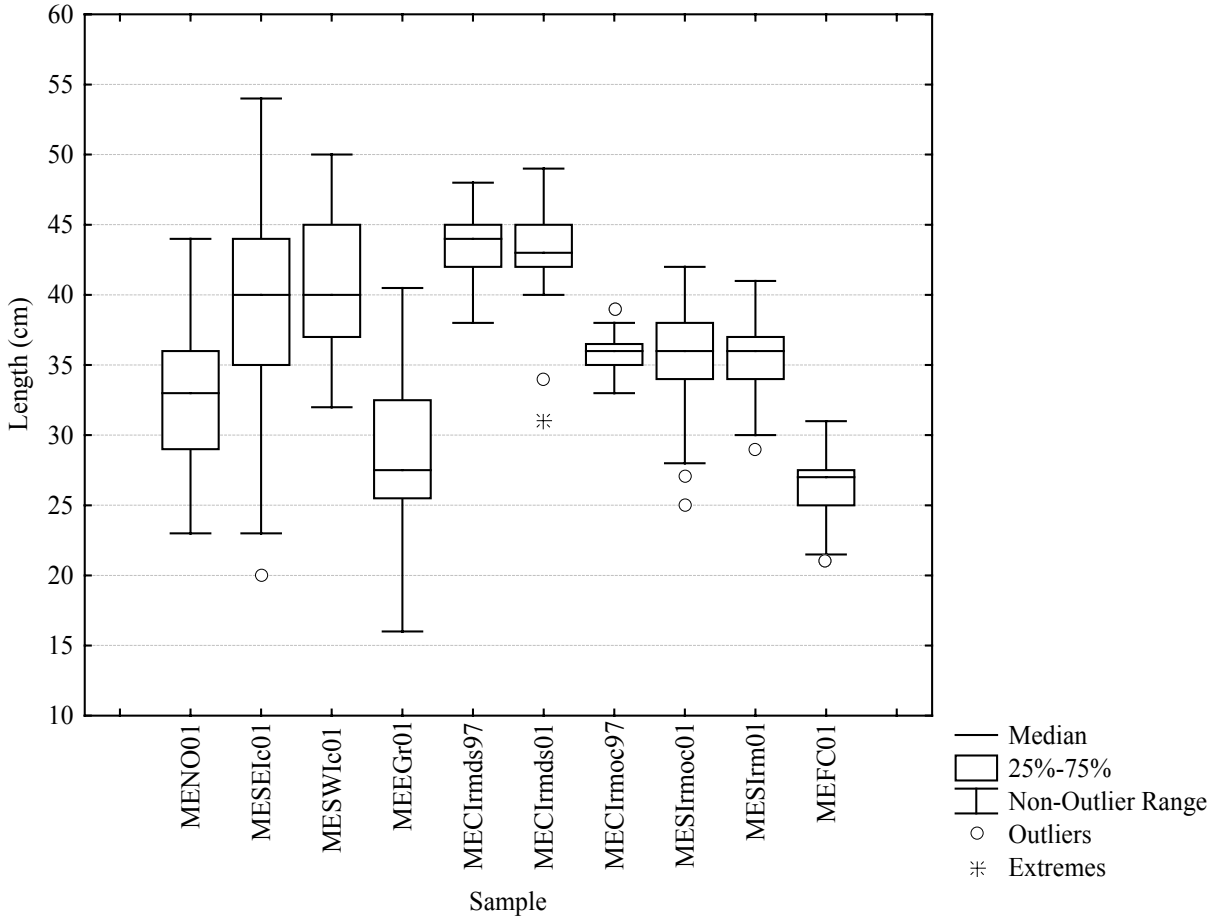
Appendix I (continued)

[1] Federal Research Centre for Fisheries (BFAFi), Hamburg, Germany; [2] Institute of Marine Research, Vigo, Spain; [3] Marine Research Institute (MRI), Reykjavík, Iceland; [4] Institute for Marine Research (IMR) and the University of Bergen (UIB), Department of Fisheries and Marine Biology, Bergen, Norway

Depth (m)	**Station**	**Institute**	**Cruise id**	**Sample code**	***n***
293	45	Spain[2]	CAFC01	MAFC01	17
240	75	Spain[2]	CAFC01	MAFC01	5
247	76	Spain[2]	CAFC01	MAFC01	13
148	109	Spain[2]	CAFC01	MAFC01	2
172	110	Spain[2]	CAFC01	MAFC01	10
208	113	Spain[2]	CAFC01	MAFC01	12
431	118	Spain[2]	CAFC01	MAFC01	3
269	122	Spain[2]	CAFC01	MAFC01	2
				MAFC01	64
					128
330	3	Spain[2]	CAFC01	FAFC01	20
240	40	Spain[2]	CAFC01	FAFC01	16
463	59	Spain[2]	CAFC01	FAFC01	6
322	63	Spain[2]	CAFC01	FAFC01	15
240	75	Spain[2]	CAFC01	FAFC01	7
				FA	**64**
150	514	Norway[4]	JH-81508	VINO01	1
110	516	Norway[4]	JH-81508	VINO01	1
170	518	Norway[4]	JH-81508	VINO01	4
270	519	Norway[4]	JH-81508	VINO01	8
370	10	Iceland[3]	TJB11996	VISWIc96	12
135	20	Iceland[3]	TBI-02	VISWIc01	16
				VI	**42**
					838

Appendix II

Box-and-whisker plots of the length distributions (total length) in the different samples of *S. mentella* used in the AFLP analysis. Plots were generated in STATISTICA version 6.1 (StatSoft, Inc. 2003). The line within the box represents the median of the distribution. The 25th and 75th percentiles are represented by the top and bottom box edges, respectively. Whiskers represent the lowest and highest values (Non-Outlier range). Sample codes are explained in Table 3.1 and Appendix I.



FINAL DISCUSSION

In the present thesis, mitochondrial DNA sequence analysis, microsatellite analysis and AFLP analysis were used to gain insight into the genetic structure of the North Atlantic redfish (genus *Sebastes*) species complex.
In particular, the following aspects were investigated:

(1) The genetic structure of North Atlantic *Sebaste*s species and their evolutionary relationship
(2) The possibility of introgressive hybridisation between the species
(3) The genetic population structure of *S. marinus* occurring on the continental slopes of Greenland, Iceland and the Faroe Islands, in Norwegian waters and on the Flemish Cap
(4) The genetic population structure of *S. mentella* occurring in the Irminger Sea, on the continental slopes of Greenland and Iceland and on the Flemish Cap
(5) The determination of the most appropriate molecular markers for reliable species identification and detection of population structure

The main findings will be summarised in the following paragraphs and will be complemented with an outline of conclusions and perspectives.

The genetic structure of North Atlantic Sebastes species: Cryptic species and possible mechanisms of speciation

The results obtained from the three molecular methods used in the present study provided evidence that the currently recognised species *S. mentella*, *S. fasciatus* and *S. viviparus* are genetically distinct and, therefore, indeed represent valid species.
However, the most salient finding of this thesis was that samples of specimens pre-classified as *S. marinus* according to external morphological characters - collected on the continental slopes of Greenland, Iceland and the Faroe Islands - were genetically extremely heterogeneous. The results of the ND3 gene sequence analysis, the microsatellite and the AFLP analyses provided evidence of two genetically isolated groups of individuals within these samples.
One of these groups was found in all areas studied, from Norway to the Flemish Cap, and most likely represents the species *S. marinus*. The occurrence of the second genetically distinct group of individuals, however, suggested that another cryptic species of *Sebastes* occurs on the continental slopes of Greenland, Iceland and the Faroe Islands - and maybe also in other areas which have not been covered by the present study.

In particular, the results of the ND3 gene sequence analysis and the microsatellite analysis indicated the occurrence of cryptic speciation due to (i) the prevalence of two distinct ND3 gene haplotype lineages with a relatively high level of intraspecific genetic divergence in the *S. marinus* samples from Iceland and Greenland and (ii) the occurrence of two genetically isolated groups of individuals within samples from Greenland, Iceland and the Faroe Islands, identified using microsatellite polymorphisms. These two groups revealed genetic differences of the same order of magnitude as the genetic differences observed between all other North Atlantic *Sebastes* species. The genetic distinctness of the two groups of individuals - on the level of mitochondrial as well as nuclear DNA - provides strong evidence of species separation and cryptic speciation, in particular as the groups occurred sympatrically.
Alternate hypotheses that could explain the observed genetic pattern - such as selection or introgressive hybridisation - are less likely. Selection is likely to differentiate only a minority of functional coding loci (reviewed in Carvalho and Hauser 1995). The strong concordance among different loci and marker systems, however, indicated a robust signal of genetic differentiation between the different groups of individuals. Furthermore, the two distinct lineages were observed in samples collected in different years, indicating the long-term stability of the genetic structure.
As outlined in the next paragraph, also hybridisation is an unlikely cause of the observed genetic pattern.
Even though the discovery of cryptic speciation within North Atlantic *Sebastes* may be surprising, the occurrence of cryptic species is not uncommon for the genus *Sebastes*: Cryptic speciation has been observed in the South Atlantic and in the North Pacific (Rocha-Olivares 1999b; 1999c; Kai et al. 2002a and 2002b; Kai and Nakabo 2004; Gharrett et al. 2005). Previous genetic studies focusing on the North Atlantic may have failed to detect cryptic speciation due to (i) the limited resolution of the protein markers which dominated previous genetic studies (e.g. Payne and Ni 1982; Nedreaas and Nævdal 1991a; Rubec et al. 1991), (ii) the geographic restriction of these studies and/or (iii) the failure to include all species, making it impossible to detect the genetic variation of North Atlantic *Sebastes* over its range of distribution (e.g. McGlade et al. 1983; Desrosiers et al. 1999; Roques et al. 1999a).

In the context of cryptic species it is important to note that one of the most characteristic attributes of the genus *Sebastes* is its species richness with about 110 species worldwide (Love et al. 2002). Furthermore, *Sebastes* appears to speciate quickly, perhaps on the scale of a few tens of thousands of years (Rocha-Olivares et al. 1999a; 1999b; 1999c). How can it be explained that rapid acquisition of reproductive barriers appears to be a common phenomenon for *Sebastes* throughout the range of the genus - including the North Atlantic?
Some authors have proposed that after the colonisation of the North Atlantic by an ancestor species of North Atlantic *Sebastes* about 3.5 to 3.0 million years ago, vicariant events (although not defined by Briggs 1995) eventually broke the distribution range of the ancestral

species and the isolated populations evolved into the four currently known North Atlantic *Sebastes* species (Roques et al. 2001). According to this scenario, the area Faroe Islands/Iceland/Greenland could represent a secondary contact zone of already reproductively isolated species. However, even though allopatric speciation seems to have occurred on a number of occasions in the Pacific, it has been proposed that is likely that many species of *Sebastes* have formed in sympatry (Love et al. 2002). This could also apply to the North Atlantic *Sebastes* species, in particular as *S. marinus*, *S. mentella* and the newly identified potential cryptic species occur sympatrically in the area Faroe Islands/Iceland/Greenland.
The identification of the mechanisms in North Atlantic *Sebastes* speciation is difficult, as our knowledge about biology and behaviour of these species is still limited. However, a particular aspect of the life history of the genus may explain the occurrence of sympatric speciation and cryptic species: One of the most characteristic features of the genus *Sebastes* is the unusual mode of reproduction. All species are viviparous with internal fertilisation (reviewed by Wourms 1991). Internal fertilisation may permit more restrictive mate selection than is possible in fishes with external fertilisation (Love et al. 2002). This may allow several species to occur sympatrically, and also for speciation to take place more rapidly than in fish with external fertilisation. Very little is known about the courtship behaviour of North Atlantic redfish species, but in some Pacific species (e.g. *S. inermis* and *S. mystinus*) it involves the performance of a sequence of stereotypical behaviours, performed mostly by the male, that are required to ensure successful copulation (Helvey 1982; Shinomiya and Ezaki 1991; reviewed in Rocha-Olivares et al. 1999b). In addition, *Sebastes* species are capable of producing sounds using the gas bladder musculature (Hallacher 1974) and sound production may play a role in mate recognition as in other teleosts (Tavolga 1971). There is also some evidence that the production of pheromones (chemical attraction) is included in these rituals (see Love et al. 2002 and references therein). Therefore, the existence of elaborate courtship behaviours, presumably leading to mate selection and consequent assortative mating, represent some plausible mechanisms for rapid acquisition of reproductive and/or recognition barriers in redfish (Rocha-Olivares et al. 1999b) - and may thus explain the observed genetic patterns indicating cryptic speciation.

Although the results of the present study indicated the occurrence of cryptic speciation in North Atlantic *Sebastes*, they did not support the hypothesis that "giant" *S. marinus* represent a genetically isolated group or separate species of North Atlantic *Sebastes*, as proposed by Johansen et al. (2000a). There is indication that a part of the individuals of "giant" *S. marinus* analysed in this study were the product of an ancient introgressive hybridisation event between *S. mentella* and *S. marinus* (see next paragraph), but not all individuals were introgressed and therefore represented "ordinary" *S. marinus*, on the level of both nuclear and mtDNA.
Even though the five genetic groups - representing the four currently recognised species and the newly identified cryptic species - were genetically distinct, the results of the ND3 gene

analysis revealed low levels of sequence divergence and a star-like phylogeny. This indicates that the North Atlantic *Sebastes* species are closely related and most likely derived from an explosive expansion after a population bottleneck or other demographic effects like founder events (according to Slatkin and Hudson 1991; Grant and Bowen 1998; Avise 2000). This is in congruence with the general perception that the North Atlantic *Sebastes* species are closely related and share a common ancestor (Bentzen et al. 1998; Sundt and Johansen 1998; Rocha-Olivares 1999a; Love et al. 2002). Furthermore, the evolutionary scenario revealed by the phylogenetic analysis of the ND3 data suggested that *S. mentella* is probably the representative of the most basal lineage of North Atlantic *Sebastes*. The results of the ND3 gene sequence analysis further indicated a direct descendance of the different genetic lineages from the basal *S. mentella* lineage, whereby the *S. viviparus* lineage probably originated from an earlier split from this basal lineage. The *S. mentella* lineage could therefore represent the radiation origin that gave rise to all other North Atlantic *Sebastes* species.

Is there indication of natural hybridisation between the species of North Atlantic Sebastes?

Even though the results of the ND3 gene analysis indicated that the North Atlantic *Sebastes* species are closely related and evolved relatively recently, no evidence of recent broad-scale hybridisation between the species was found. The small number of admixed genotypes identified by microsatellite analysis indicated that there is only restricted, if any, hybridisation between *S. marinus*, *S. mentella* and the potential cryptic species in the areas off the Faroe Islands, Iceland, East and West Greenland, as well as between *S. fasciatus*, *S. marinus* and *S. mentella* on the Flemish Cap. Furthermore, the degree of genetic variability in the ND3 gene was large enough to recognise diagnostic mutations characteristic for several of the species analysed. This presence of diagnostic mutations can also be interpreted as evidence of reproductive isolation between the species.

Therefore, extensive broad-scale hybridisation between *S. marinus* and *S. mentella* as an explanation for the wide overlap in morphological and morphometrical characters and in electrophoretic mobility patterns of hemoglobins and blood serum proteins frequently observed on the shelves of Greenland and Iceland (Altukhov and Nefyodov 1968; Nedreaas and Nævdal 1991a; Nedreaas et al. 1994) can be ruled out. According to Nedreaas et al. (1994), the observed variation in the hemoglobin pattern could be explained either by hybridisation between *S. marinus* and *S. mentella*, or variation within *S. marinus*. The results of the present study indeed support the latter hypothesis, indicating that the variation in the hemoglobin pattern found in samples of *S. marinus* in these areas is not the result of hybridisation between *S. mentella* and *S. marinus*, but is rather due to intraspecific variation or, more likely, interspecific variation: As outlined in the previous paragraph, there is indication that a fifth - apparently cryptic - species of *Sebastes* occurs on the continental slopes of Greenland, Iceland and the Faroe Islands. The fact that this cryptic species was

found mainly in samples of *S. marinus* indicates that it is - even though very distinct genetically - morphologically similar or indistinguishable from *S. marinus*, at least with the morphological and morphometrical characters currently used to discriminate the North Atlantic *Sebastes* species. The occurrence of this cryptic species in samples of *S. marinus* from Greenland and Iceland could explain the variation in the hemoglobin pattern observed by Nedreaas et al. (1994) in these areas.

In general, hybridisation in *Sebastes* appears to be a rare event (Love et al. 2002). Nevertheless, broad-scale introgressive hybridisation between *S. mentella* and *S. fasciatus* in the Gulf of St. Lawrence and adjacent areas has been observed by Roques et al. (2001). This extensive asymmetrical hybridisation resulted in the co-existence of two introgressed populations of *S. mentella* and *S. fasciatus*, which remained reproductively isolated in the main zone of sympatry, most likely due to selection favouring the differential survival of certain hybrids (Roques et al. 2001).
In the Pacific, however, where a high number of *Sebastes* species exists, which are often closely related and live in close proximity, hybrids have rarely been seen or at least recognised (reviewed in Love et al. 2002).
The extent of hybridisation in *Sebastes* is probably limited by factors such as the mode of reproduction (internal fertilisation) as well as temporally and spatially separated mating activities (Desrosiers et al. 1999). As outlined in the previous paragraph, the former is most likely the most important factor that limits hybridisation between the *Sebastes* species.

Even though no evidence of recent hybridisation between North Atlantic *Sebastes* species was found in the present study, the results indicate that introgressive hybridisation has occurred in the past - at least in some cases:
The comparison of the results of the ND3 gene sequence analysis and the microsatellite analysis revealed that the individuals of *S. marinus* from the Flemish Cap resembled *S. marinus* on the level of nuclear DNA, but showed mitochondrial haplotypes identical to haplotypes found in *S. fasciatus*. This incorporation of *S. fasciatus* specific ND3 gene haplotypes into *S. marinus* on Flemish Cap - without any evidence of nuclear introgression - indicates ancient introgressive hybridisation events between *S. marinus* and *S. fasciatus*.
The hypothesis of ancient introgressive hybridisation events (in this case between *S. mentella* and *S. marinus*) is also supported by the appearance of *S. mentella* specific haplotypes in several individuals of "giant" *S. marinus* analysed in the present study - again without any evidence of nuclear introgression - even though in contrast to *S. marinus* from the Flemish Cap not all individuals were introgressed.
In conclusion, the results of the present thesis indicated that the North Atlantic *Sebastes* species are genetically isolated and recent hybridisation is unlikely or restricted, but that there is evidence of ancient introgression events among certain species in certain areas.

Patterns of population structure in North Atlantic Sebastes

The general patterns of population structure in *S. mentella*, *S. marinus* and the newly identified genetic group and potential cryptic species of North Atlantic *Sebastes* - revealed by microsatellite polymorphism - were congruent with the weak genetic structuring usually reported for populations of marine fish species (Ward et al. 1994). The apparent genetic homogeneity of many marine species is commonly thought to be due to large effective population sizes that limit genetic drift (DeWoody and Avise 2000) and life history characteristics that favour dispersal (e.g. planktonic eggs and larvae, juvenile and adult vagility) in continuous dynamic oceanic environments (Ward et al. 1994; Waples 1998).
However, differential patterns of population structure were observed for *S. marinus* and *S. mentella* and the newly identified cryptic species of North Atlantic *Sebastes*. Whereas the results of the microsatellite analysis did not indicate population structure within the latter, indication of population structure was found in *S. marinus* and *S. mentella*. Even though the levels of genetic differentiation observed in *S. marinus* and *S. mentella* were subtle, they were nevertheless significant and comparable to estimates found in populations of other marine fish species (Ward et al. 1994). However, *S. mentella* revealed lower levels of population structure than *S. marinus*, which may be partly explained by their respective life histories: In general, lower population structure is expected for pelagic and widely distributed marine species - such as *S. mentella* - compared to benthic and/or more geographically restricted species - such as *S. marinus* (Avise et al. 1987; reviewed in Roques et al. 2001).
The observed patterns of population structure found within *S. marinus* and *S. mentella* will be discussed in the following two paragraphs.

The genetic population structure of S. marinus

The results of the analysis of microsatellite variation presented in Chapter 2 revealed population structure in *S. marinus*, indicating that there are at least three genetically distinct populations of *S. marinus* in the geographical area covered by this study. The samples from (i) the Flemish Cap, (ii) Greenland and (iii) Norway, Iceland and the Faroe Islands stemmed from distinct populations, whereby the latter two populations apparently overlap in some areas, e.g. on the Dohrn Bank located between Iceland and Greenland and in some areas on the Icelandic shelf. The comparison of the samples from Norway, the Faroe Islands and Iceland revealed extremely small, if any, genetic differences. This suggested that these samples could belong to the same population, or that high levels of gene flow or recent ancestry connect *S. marinus* populations occurring in these areas.
The genetic distinctness of Flemish Cap *S. marinus* is in congruence with the general perception that the redfish populations of the Flemish Cap are relatively isolated (reviewed in Templeman 1976). However, the observed population structure of *S. marinus* in the central

and eastern North Atlantic is only partly in congruence with the current management units. Even though *S. marinus* in these areas are believed to constitute two stocks (ICES 1998a) - which is consistent with the present results - the general assumed population boundaries are not supported by the results of the microsatellite analysis. *S. marinus* in the Northeast Arctic (ICES subareas I and II) is considered to represent a single stock, separated from the stock on the continental slopes of Greenland, Iceland and the Faroe Islands (ICES Divisions V, VI, XII, XIV and NAFO 1; ICES 1998a). In contrast, the present results indicate that *S. marinus* from Norwegian waters, Iceland and the Faroe Islands could belong to the same population, which, however, seems to be separated from *S. marinus* occurring off Greenland. The current separation into the two management units - Northeast Arctic and Greenland/Iceland/Faroe Islands - is mainly based on the generally assumed locations of the major areas of larval extrusion (located off the Norwegian coast and southwest of Iceland, respectively) and the nursery areas (located in the Barents Sea and off Greenland and Iceland, respectively; ICES 1998a). However, previous analyses of hemoglobin and allozyme polymorphisms already raised some doubts about this stock division and pointed to a closer relationship between *S. marinus* from Norwegian and Faroe waters (Nedreaas and Nævdal 1991a). Furthermore, gene flow caused by extensive migration between Norway, the Faroe Islands and Iceland cannot be precluded, considering the high dispersal potential of the species (Reinert and Lastein 1992).

It should be noted that there is also the possibility that *S. marinus* from Norway and the area Faroe Islands/Iceland represent two reproductively isolated populations. However, due to recent common ancestry of the two populations or due to high levels of gene flow, no signficant genetic differences were detected with the microsatellite markers used in the present study. The lack of genetic differentiation could also represent an ephemeral phenomenon, as only one sample from Norway was analysed, which could have represented a group of migrants from the Icelandic/Faroe population. More temporal replicates are needed to verify whether the observed genetic pattern is stable over time. The same applies to the finding that the samples from Greenland and Iceland represent two genetically distinct populations of *S. marinus*. Even though the results are strongly supported by the fact that the structure was observed in two subsequent years, indicating temporal stability, it is crucial to investigate the population structure over a longer time period, especially given the extreme longevity of the species.

Nevertheless, the present study - representing the first analysis of *S. marinus* population structure using polymorphic microsatellite markers - has provided new insights into the genetic structure of *S. marinus* populations in the North Atlantic.

The genetic population structure of S. mentella

In this thesis, the genetic structure of *S. mentella* was analysed using three different methods. Microsatellite analysis was used to analyse the population structure of *S. mentella*. The ND3 gene sequence analysis was used to determine, whether the morphological phenotypes of *S. mentella* ("oceanic" and "deep-sea" type) are also genetically distinct, as proposed by some authors (e.g. Johansen et al. 1998 and 2000b). AFLP analysis was used to analyse both issues: *S. mentella* population structure and differentiation between phenotypes.

Overall, the results of the microsatellite analysis presented in Chapter 2 revealed weak but significant genetic structure in *S. mentella*, even though the levels of genetic differentiation were smaller than those observed in *S. marinus*. In summary, the results of the microsatellite analysis indicated significant genetic differences between the sample from the Flemish Cap, the samples from the central Irminger Sea and the remaining samples from East Greenland (2001), West Greenland, the southern Irminger Sea and the two samples collected in the NAFO areas 1F and 2J, which were genetically homogeneous. However, it cannot be ruled out that also migrants from other areas/populations were present in the samples, as the geographic area covered in the microsatellite analysis was limited. This is also indicated by the temporal instability observed in samples from Greenland waters.

Even though the observed genetic distinctness of Flemish Cap *S. mentella* is not as prominent as in *S. marinus* and there is indication of gene flow or a relatively recent isolation, the results indicated that *S. mentella* on the Flemish Cap represents a distinct population - in congruence with the general perception that the redfish populations of the Flemish Cap are relatively isolated (Templeman 1976). The pattern of genetic structure observed in *S. mentella* in the area Irminger Sea/Iceland/Greenland is at least partly congruent with the results of a previous microsatellite-based study by Roques et al. (2002). Roques et al. (2002) found no or only weak genetic differences between samples from Iceland, the Irminger Sea and East and West Greenland. Also the present study found no or only extremely weak genetic differences between samples from East Greenland, West Greenland, the southern Irminger Sea and the NAFO areas 1F and 2J. In contrast to the findings by Roques et al. (2002), however, the two samples from the central Irminger Sea revealed significant levels of genetic differentiation from the other samples collected in these areas. The different outcomes of the two studies can be explained by the larger sample size from the Irminger Sea analysed in the present study. Furthermore, the results of the present study indicated temporal instability, most likely due to a different distribution of population units in different years, which could explain the differences to the study by Roques et al. (2002).

The results of the ND3 gene analysis and the AFLP analysis reflect - at least to some extent - the results of the present microsatellite analysis, indicating genetic structure in *S. mentella* in the Irminger Sea and adjacent areas: Samples from the central Irminger Sea, which were additionally classified as "deep-sea" type of *S. mentella*, were genetically different from

samples classified as "oceanic" *S. mentella* collected in the Irminger Sea, as well as from samples of *S. mentella* from other regions throughout the North Atlantic.

As outlined in detail in Chapter 1 (paragraph 5.3.4), there has been a strong controversy about whether the different *S. mentella* phenotypes represent "valid" phenotypes, as the differences are subtle and only a few experts are able to distinguish between them (ICES 1998a; Johansen et al. 2000b; Saborido-Rey et al. 2005). A very recent study has revealed no morphometrical differences between the two morphological types (Garabana 2005), indicating that the sorting of individuals into phenotypes was based on what seems to be mainly personal experience and - at least to some extent - size (ICES 2004b; Nielsen 2004). Nevertheless, both ND3 gene analysis and AFLP analysis revealed genetic differences between samples of so-called "deep-sea" *S. mentella* collected in the Irminger Sea and all other samples of *S. mentella*.

The population structure of *S. mentella* in the Irminger Sea and on the continental slopes of Greenland and Iceland - and the three hypotheses that have been put forward to describe this structure (the single stock hypothesis, the two stock hypothesis and the three stock hypothesis; ICES 1998a; ICES 2004b; see the GENERAL INTRODUCTION chapter for details) - has been discussed for many years.

The results of the present study as well as the study by Roques et al. (2002) do not support a separation of demersal *S. mentella* living on the shelves and the so-called "oceanic" *S. mentella* occurring pelagically in shallower waters of the Irminger Sea, as proposed by the two and the three stock hypothesis. The results of the present study, however, also do not support the single stock hypothesis either, as they indicated some genetic structure in *S. mentella*.

The genetic structure observed in the present study is indeed difficult to explain, in particular as it has been proposed that *S. mentella* release their larvae in a single and wide spawning (larval release) area in an area southwest of Iceland and above the Reykjanes Ridge, with a wide depth distribution and with a single spawning peak from April to June - supporting the existence of a single stock (Saborido-Rey et al. 2005). One might expect *S. mentella* to be panmictic if there is indeed a single and wide spawning (= larval release) area and one single nursery area for *S. mentella* (located on the shelf of Greenland), as proposed by Saborido-Rey et al. (2005). However, panmixia is inconsistent with our data, which indicate genetic structure.

There are two possible explanations for the observed genetic structure in *S. mentella*: Differences between different year-classes due to sweepstakes chance effects (Hedgecock 1994; see also Saborido-Rey et al. 2005) and population separation due to restricted adult dispersal and/or larval transport (e.g. Magnússon and Magnússon 1995).

Assuming that larval drift, juvenile dispersal, and/or adult migration are prevalent, one possible explanation for genetic heterogeneity, particularly for long-lived and highly fecund species, is a sweepstakes chance effect, as proposed for Alaskan shortraker rockfish (*Sebastes borealis*; Matala et al. 2004a) and bocaccio (*Sebastes paucispinis;* Matala et al. 2004b). Under

a sweepstakes model, genetic composition can vary from year-class to year-class and from location to location just by random chance (Hedgecock 1994). Samples with different age-class compositions may exhibit divergence that reflects variation among cohorts rather than geographically influenced divergence (Matala et al. 2004b). This could also apply to *S. mentella*: According to Saborido-Rey et al. (2005) the population of *S. mentella* in the Irminger Sea and adjacent areas could be sustained by few strong year-classes, even if there is recruitment failure over long periods of time. This is supported by the fact that after a severe depletion of *S. mentella* off East Greenland during the 1980's and early 1990's, a sharp increase in abundance was observed in 1995-1997, mainly due to a single strong year-class of young fish, which later migrated into the central Irminger Sea recruiting to the adult stock there (Stransky 2000). Furthermore, no significant correlation between geographic and genetic distances was observed in *S. mentella* at the spatial scale of the present study. Such genetic heterogeneity without a clear geographic pattern is expected under the sweepstakes chance effect (Matala et al. 2004a). In addition, a comparison of the total lengths of all *S. mentella* specimens used in the microsatellite analysis revealed that the total length of the two genetically distinct samples from the central Irminger Sea collected in greater depths consisted of much larger individuals (see Appendix IIb in Chapter 2) as for example the samples from Greenland and thus most likely represent a different (older) age group. Also the length distribution of the *S. mentella* specimens used in the ND3 gene analysis as well as in the AFLP analysis revealed that the total length of "deep-sea" *S. mentella* was on average higher than that of "oceanic" S. *mentella* and untyped specimens collected in other areas (see Appendix II in Chapter 1 and Appendix II in Chapter 3). This could indicate that "deep-sea" *S. mentella* could represents an older age group, as studies have shown that the distribution of the age composition fully reflects the dynamics of the length composition in North Atlantic redfish (Saborido-Rey et al. 2005).

However, cohort effects would require that there is a non-trivial level of genetic differentiation among redfish cohorts, requiring relatively little overlap (gene flow) among generations and small effective population sizes N_e (Nielsen 2004).

Small effective population sizes are not implicitly indicated by the levels of variability observed at the microsatellite loci analysed, with between 23 and 107 alleles at individual loci. Therefore, to validate cohort differentiation as a possible explanation for the observed genetic structure, more investigations, e.g. of the qualitative estimation of the required effective population sizes and the overlap among cohorts, are needed (Nielsen 2004).

The observed genetic structure could also be caused by population separation due to restricted adult dispersal and larval transport. Even though the assumed existence of a single and wide spawning (larval release) area and one single nursery area (Saborido-Rey et al. 2005) supports the single stock hypothesis and sweepstakes chance effects, it cannot be ruled out that there are factors that restrict adult and also juvenile dispersal. The findings of the present study - in particular the existence of a fifth cryptic *Sebastes* species in the area Greenland/Iceland/Faroe

Islands - have shown that our knowledge about the whole North Atlantic *Sebastes* species complex is very limited. The larval and juvenile ecology of *S. mentella* is poorly known (Roques et al. 2002). The same applies to adult behaviour, e.g. migration and mating behaviour. Population structure could therefore develop as a result of unknown behavioural characteristics regulating vagility at various life history stages (still poorly described; Roques et al. 2002). As outlined above, it is generally believed that assortative mating has played an important role in redfish speciation (reviewed in Love et al. 2002), but it could also represent an important factor for the evolution of *Sebastes* population structure.

Both hypotheses, differences between year-classes due to sweepstakes chance effects and population separation due to restricted adult dispersal and/or larval transport are consistent with the finding of this study, namely the significant genetic differentiation between *S. mentella* samples collected in the central Irminger Sea and *S. mentella* samples collected in Greenland waters and in shallower waters in the southern Irminger Sea. However, it is relatively unlikely that the observed genetic differentiation is caused by differences between cohorts alone, due to the high levels of genetic diversity and the high number of alleles revealed by the microsatellite analysis.

Furthermore, there is strong indication of temporal instability of the structure, suggesting that geographic patterns could result from movement of population components across putative population boundaries. The number of temporal genetic samples in this study is limited, in particular considering the longevity of the species and its high potential for dispersal. Therefore, more temporal replicates are needed before definite conclusions about the population structure of *S. mentella* in the Irminger Sea and adjacent waters can be drawn.

The usefulness of mitochondrial DNA sequences, microsatellite and AFLP data for species discrimination and detection of population structure of North Atlantic Sebastes

Species discrimination

The three genetic methods used in this thesis - ND3 gene sequence analysis, microsatellite and AFLP analysis - have proven useful in detecting interspecific patterns of genetic structure in North Atlantic *Sebastes*.

Even though the degree of genetic variability of the ND3 gene was relatively low, diagnostic mutations characteristic for several of the North Atlantic *Sebastes* species could be recognised. However, there was a strong indication of introgression in certain species and areas (e.g. the incorporation of *S. fasciatus* specific haplotypes into *S. marinus* on the Flemish Cap). This emphasises the importance of covering the entire geographic species or taxon ranges and the inclusion of additional nuclear markers when conducting mtDNA sequence analysis (Ballard and Whitlock. 2004). Furthermore, DNA sequence analysis is too expensive and time consuming for a routine application in large-scale surveys, unless it reveals species-

specific restriction sites in the gene analysed, which would allow the fast and relatively inexpensive PCR-RFLP analysis[1]. However, as no restriction sites were found within the ND3 gene, the method is not applicable to large-scale screening of individuals.

Nevertheless, the ND3 gene analysis has provided the first hypothesis for the evolutionary relationship between the North Atlantic *Sebastes* species based on quantitative phylogenetic analyses of molecular sequence data. Furthermore, the combination of ND3 gene analysis and microsatellites has provided new insights into possible ancient introgressive hybridisation events within North Atlantic *Sebastes* that would have remained unnoticed if only a single genetic analysis method had been used.

The present study of microsatellite variation showed that microsatellites are a useful tool for North Atlantic *Sebastes* species identification. The results demonstrated that on average 94% of the individuals could be successfully assigned to species using only six microsatellite loci. A higher number of loci will probably further improve assignment success. Furthermore, by using Bayesian cluster analysis it was possible to rule out extensive broad-scale hybridisation between *S. marinus* and *S. mentella* as cause for the wide overlap in morphological and morphometrical characters frequently observed for example in Greenland waters. The potential usefulness and simplicity of redfish species discrimination using microsatellite analysis was already demonstrated by Roques et al. (1999a). Microsatellite analysis should therefore be considered as reliable standard method for species identification during scientific surveys in areas were several species occur sympatrically and species identification is particularly difficult, e.g. on the continental slopes of Greenland, Iceland and the Faroe Islands.

AFLP analysis was also found to be highly effective in distinguishing species of North Atlantic *Sebastes*. About 98% of the individuals analysed were successfully assigned to species. However, a few individuals were assigned to the wrong species. This was most likely caused by limited sample sizes for some species (*S. viviparus* and the newly identified cryptic species of *Sebastes*). Nevertheless, the high assignment success demonstrates the high potential of AFLP analysis in North Atlantic *Sebastes* species identification. Furthermore, the genetic relationship between the species revealed by the AFLP analysis supported the results of the ND3 gene sequence analysis, indicating that within the North Atlantic *Sebastes* species complex, *S. viviparus* is the most distant species and that *S. mentella* is closely related to *S. marinus*, as indicated by other studies (McGlade et al. 1983; Sundt and Johansen 1998).

However, the AFLP technique has two drawbacks: Firstly, data usually must be treated as dominant markers (= heterozygous genotypes with a given band cannot be directly distinguished from homozygotes with the band), thus significantly decreasing the genetic

[1] amplification of a DNA fragment using PCR and subsequent digestion with restriction endonucleases, resulting in species-specific restriction fragment length polymorphisms (RFLP)

information available (Vos and Kuiper 1997). Therefore, it was not possible to investigate hybridisation between the North Atlantic *Sebastes* species using AFLP analysis. Furthermore, the AFLP analysis is more time consuming and expensive than the microsatellite analysis.
In conclusion, microsatellite analysis is probably the easiest and the most reliable method to assign individual redfish to the correct species (Roques et al. 1999a) - and to investigate hybridisation. Furthermore, microsatellite analysis was successfully used both on inter- and on intraspecific level (see below).

Detection of population structure

Only microsatellite and AFLP analysis will be discussed here, since the ND3 gene revealed relatively low levels of genetic variability and DNA sequence analysis is too time-consuming and expensive given the larger sample sizes required for analyses of population structure.
The results of the microsatellite analysis provided new insights into the population structure of North Atlantic *Sebastes* species, in particular of *S. marinus*, but also of *S. mentella*. The method was sufficiently sensitive to detect highly significant genetic differences between samples of *S. mentella* and *S. marinus* from different locations, even though one to two loci (in the population genetic analysis of *S. mentella* and *S. marinus*, respectively) had to be excluded from the statistical analyses due to null alleles. A higher number of loci will probably increase the resolution of population structure analysis.
In contrast, the results of the AFLP analysis of population structure of *S. mentella* were more or less inconclusive. Significant genetic differences were observed between all samples, suggesting that AFLP markers may be too variable to detect biological meaningful patterns of genetic structure in weakly differentiated populations of *S. mentella*. However, the major genetic pattern revealed by the results of the AFLP analysis indicated that "deep-sea" *S. mentella* was genetically different from all other samples of *S. mentella*. This finding was supported by the results of the ND3 gene analysis. Even though the ND3 gene analysis was based on small sample sizes and results must therefore be interpreted with caution, the congruence between the two methods indicated that the genetic signal is relatively robust (as it was displayed by a mitochondrial marker with low variability and a nuclear marker system with high variability).
As mentioned above, however, AFLP markers have the disadvantage that they are dominant markers, which restricts the use of population genetic theory and statistical framework that can be applied. Consequently, microsatellites clearly allow more powerful population genetic analyses than AFLPs (Mueller and Wolfenbarger 1999).
In conclusion, this study has demonstrated that the application of several molecular markers - in particular the combination of mtDNA sequences and microsatellite data - can provide valuable insights into the species and population structure of North Atlantic *Sebastes*.

Consequently, microsatellites should be the preferred marker in large-scale analyses of North Atlantic *Sebastes*, as they represent an easy and reliable method on the inter- as well as on the intraspecific level.

Conclusions and perspectives

The results of this thesis provide new insights into the genetic structure of the North Atlantic *Sebastes* species, into their evolutionary relationship, as well as into their genetic population structure.

The most relevant finding of this study was the observation of a group of individuals that was genetically distinct from the four currently known species of North Atlantic *Sebastes* (*S. mentella*, *S. marinus*, *S. fasciatus* and *S. viviparus*) - pointing to the existence of a fifth cryptic North Atlantic *Sebastes* species, which occurs on the continental slopes of Greenland, Iceland and the Faroe Islands.

The addition of a new cryptic species to the North Atlantic *Sebastes* species complex underscores the fact that we still know very little about these species - and emphasises the need for further investigations to understand the distribution and the biology of this species complex.

The indication of cryptic speciation has also wide-ranging implications for the management and conservation of North Atlantic *Sebastes*, as it reveals the existence of biological diversity that will go unnoticed without the careful examination of genetic variation. Therefore, a clear understanding of species differentiation and distribution of North Atlantic *Sebastes* - essential for a proper understanding of their biology and for an adequate management of the exploited stocks - requires a reliable technique for accurate species identification (Gascon 2003). As demonstrated in the present study and by Roques et al. (1999a), microsatellite analysis could be considered as a reliable standard method for redfish species identification, e.g. on surveys in areas where several species occur sympatrically and where species identification is particularly difficult. Moreover, microsatellite markers can easily be used at any life history stage, which may be of interest in studies of recruitment and larval ecology (Roques et al. 1999a).

The results of this genetic study also provided new insights into the patterns of population structure in *S. marinus* and *S. mentella,* although the exact determination of intraspecific population structure will require a more detailed temporal and geographical sampling scheme, especially since the species are long-lived and the distribution of potential populations can vary temporally.

The findings of the present study would therefore indicate the need for a more intensive study with more loci (microsatellite and perhaps also mtDNA loci), larger sample sizes and sample collections covering a broader area and in particular temporal replicates over a longer period of time. Also further investigations of the biology of the species are needed (e.g. larval and

juvenile ecology, migration patterns, mating behaviour, identification of mating areas, etc.) to provide a more detailed picture of the *Sebastes* species complex on the continental slopes of Greenland, Iceland and the Faroe Islands and in the Irminger Sea.

REFERENCES

Adams, B. K., and J. A. Hutchings. 1999. Microgeographic population structure of brook charr: a comparison of microsatellite and mark-recapture data. J. Fish Biol. 62 (3): 517-533

Albertson, R. C., Markert, J. A., Danley, P. D., and T. D. Kocher. 1999. Phylogeny of a rapidly evolving clade: The cichlid fishes of Lake Malawi, East Africa. Proc. Natl. Acad. Sci. 96: 5107-5110

Alesandrini, S., and G. Bernardi. 1999. Ancient species flocks and recent speciation events: What can rockfish teach us about cichlids (and vice versa)? J. Mol. Evol. 49 (6): 814-818

Allendorf, F. W., Ryman, N., and F. M. Utter. 1987. Genetics and fishery management past, present, and future. Pp. 1-19 *in* Ryman, N., and F. M. Utter, eds. Population genetics and fishery management. Washington Sea Grant Program, Seattle: 420 pp.

Altukhov, Ju. P., and G. N. Nefyodov. 1968. A study of blood serum protein composition by agar-gel electrophoresis in types of redfish (genus *Sebastes*). ICNAF Res. Bull. 5: 86-90

Altukhov, Ju. P., Nefyodov, G. N., and A. N. Payusova. 1968. Thermostability of isolated muscle in determining the taxonomic relationship of the *marinus*- and *mentella*-types of redfish (*Sebastes*). ICNAF Res. Bull 5: 130-136

Altukhov, Ju. P. 1990. Population genetics - diversity and stability. Harwood Academic Publishers, London, UK: 352pp.

Arnold, M. L., Hamrick, J. L., and B. D. Bennett. 1993. Interspecific pollen competition and reproductive isolation in *Iris*. J. Hered. 84: 13-16

Atkinson, D. B. 1986. The redfish resources off Canada's east coast. Proc. Int. Rockfish Symp. Oct. 1986, Anchorage, Alaska: 15-33

Ávila de Melo, A. M., Saborido-Rey, F., and R. Alpoim, 1998. An assessment of redfish in NAFO Division 3M, including an approach to precautionary management based on spawning biomass and growth. NAFO SCR Doc. 98/53 Ser. N3044: 51 pp.

Avise, J. C., Reeb, C. A., and N. C. Saunders. 1987. Geographic population structure and species differences in mitochondrial DNA of mouthbrooding marine catfishes (Ariidae) and demersal spawning toadfishes (Batrachoididae). Evolution 41: 991-100

Avise, J. C. 1994. Molecular markers, natural history and evolution. Chapman & Hall, Inc., London: 511 pp.

Avise, J. C. 1998. The history and purview of phylogeography: a personal reflection. Mol. Ecol. 7 (4): 371-379

Avise, J. C. 2000. Phylogeography: the history and formation of species. Harvard University Press, Cambridge, MA, USA: 447 pp.

Ballard, J. W. O., and M. Kreitman. 1995. Is mitochondrial DNA a strictly neutral marker? Trends Ecol. Evol. 10 (12): 485-488

Ballard, J. W. O. 2000. When one is not enough: Introgression of mitochondrial DNA in *Drosophila*. Mol. Biol. Evol. 17 (7): 1126-1130

Ballard, J. W. O., and M. C. Whitlock. 2004. The incomplete natural history of mitochondria. Mol. Ecol. 13 (4): 729-744

Balloux, F., and J. Goudet. 2002. Statistical properties of population differentiation estimators under stepwise mutation in a finite island model. Mol. Ecol. 11 (4): 771-783

Balloux, F., and N. Lugon-Moulin. 2002. The estimation of population differentiation with microsatellite markers. Mol. Ecol. 11 (2): 155-165

Bargelloni, L., Marcato, S., and T. Patarnello. 1998. Antarctic fish hemoglobins: evidence for adaptive evolution at subzero temperature. Proc. Natl. Acad. Sci. USA 95: 8670-8675

Barsukov, V. V. 1973. Systematics of the Atlantic redfishes. Fish. Res. Bd. Can. Trans. Ser. 2531: 33 pp.

Barsukov, V. V., and G. P. Zakharov. 1973. Morphological and biological characteristics of the American redfish. Fish. Res. Bd. Can. Trans. Ser. 2488: 66 pp.

Barsukov, V. V. 1981. A brief review of the subfamily Sebastinae. J. Ichthyol. 21: 3-27

Barsukov, V. V., Litvinenko, N. I., and V. P. Serebryakov. 1985. Manual for the identification of redfish species of the North Atlantic and adjacent areas. Can. Transl. Fish. Aquat. Sci. 5168: 25 pp.

Barsukov, V. V., Oganin, I. A., and A. I. Pavlov. 1990. Morphological and ecological differences between *Sebastes fasciatus* and *S. mentella* on the Newfoundland Shelf and Flemish Cap Bank. Vopr. Ikhtiol. 30: 791-803

Beaumont, M., Barratt, E. M., Gottelli, D., Kitchener, A. C., Daniels, M. J., Pritchard, J. K., and M. W. Bruford. 2001. Genetic diversity and introgression in the Scottish wildcat. Mol. Ecol. 10 (2): 319-336

Bekkevold, D., Carvalho, G., Koljonen, M.-L., Hansen, M. M., Nielsen, E. E., Hutchinson, B., and H. Wilcock. 2004. Provide recommendations on the applications for the estimation of effective population size in wild populations of marine fish and shellfish. *In* Report of the Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM). ICES C.M. 2004/F:04: 44 pp.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., and F. Bonhomme. 1996-2002. GENETIX 4.04 Logiciel sous Windows TM, pour la Génétique des Populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France

Bentzen, P., Taggart, C. T., Ruzzante, D. E., and D. Cook. 1996. Microsatellite polymorphism and the population structure of Atlantic cod (*Gadus morhua*) in the northwest Atlantic. Can. J. Fish. Aquat. Sci. 53: 2706-2721

Bentzen, P., Wright, J. M., Bryden, L. T., Sargent, M., and K. C. T. Zwanenburg. 1998. Tandem repeat polymorphism and heteroplasmy in the mitochondrial control region of redfishes *(Sebastes*: Scorpaenidae). J. Hered. 89: 1-7

Benzécri, J. P. 1973. L'analyse des donnees. Dunod, Paris: 616 pp.

Bernatchez, L., Glémet, H., Wilson, C. C., and G. Danzmann. 1995. Introgression and fixation of Arctic charr (*Salvelinus alpinus*) mitochondrial genome in an allopatric population of brook trout (*Salvelinus fontinalis*). Can. J. Fish. Aquat. Sci. 52: 179-185

Billington, N. 2003. Mitochondrial DNA. Pp. 59-100 *in* Hallerman, E. M., ed. Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, Maryland, USA: 458 pp.

Birky, C. W. J., Maruyama, T., and P. Fuerst. 1983. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts and some results. Genetics 103: 513-527

Birky, C. W. J., Fuerst, P., and T. Maruyama. 1989. Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. Genetics 121: 613-627

Blier, P. U., and H. Lemieux. 2001. The impact of the thermal sensitivity of cytochrom c oxidase on the respiration rate of Arctic charr red muscle mitochondria. J. Comp. Physiol. B 171: 247-253

Bohonak, A. J. 2002. IBD (Isolation By Distance): a program for analyses of isolation by distance. J. Hered. 93: 153-154

Bone, Q., Marshall, N. B., and J. H. S. Blaxter. 1995. Biology of Fishes. Blackie, London, UK: 332 pp.

Briggs, J. C. 1970. A faunal history of the North Atlantic Ocean. Syst. Zool. 19: 19-34

Briggs, J. C. 1995. Global Biogeography. Elsevier Science B. V., Amsterdam: 452 pp.

Broughton, R. E., and J. R. Gold. 2000. Phylogenetic relationships in the North American cyprinid genus *Cyprinella* (Actinopterygii: Cyprinidae) based on sequences of the mitochondrial ND2 and ND4L genes. Copeia 1 (1): 1-10

Brown, W. M., Prager, E. M., Wang, A., and A. C. Wilson. 1982. Mitochondrial DNA sequences of primates: Tempo and mode of evolution. J. Mol. Evol. 18: 225-239

Buonaccorsi, V. P., Kimbrell, C. A., Lynn, E. A., and R. D. Vetter. 2002. Population structure of copper rockfish (*Sebastes caurinus*) reflects postglacial colonization and contemporary patterns of larval dispersal. Can. J. Fish. Aquat. Sci. 59: 1374-1384

Campana, S. E., Zwanenburg, K. C., and J. N. Smith. 1990. $^{210}Pb/^{226}Ra$ determination of longevity in redfish. Can. J. Fish. Aquat. Sci. 47: 163-165

Campbell, D., Duchesne, P., and L. Bernatchez. 2003. AFLP utility for population assignment studies: analytical investigation and empirical comparison with microsatellites. Mol. Ecol. 12 (7): 1979-1992

Campton, D. E. 1987. Natural hybridisation and introgression in fishes: methods of detection and genetic interpretations. Pp. 161–192 *in* Ryman, N., and F. Utter, eds. Population Genetics and Fishery Management. University of Washington Press, Seattle, USA: 420 pp.

Carvalho, G. R., and L. Hauser. 1995. Molecular genetics and the stock concept in fishes. Pp. 53-79 *in* Carvalho, G.R., and T. J. Pitcher, eds. Molecular genetics in fisheries. Chapman & Hall, London, UK: 141 pp.

Carvalho, G. R., and L. Hauser. 1998. Advances in the molecular analysis of fish population structure. Ital. J. Zool. 65 (Suppl.): 21-33

Castelloe, J., and A. R. Templeton. 1994. Root probabilities for intraspecific gene trees under neutral coalescent theory. Mol. Phylogenet. Evol. 3: 102-113

Cavalli-Sforza, L. L., and A. W. F. Edwards. 1967. Phylogenetic analysis: models and estimation procedures. Evolution 21: 550-570

Claridge, M. F., Dawah, H. A., and M. R. Wilson, eds. 1997. Species. The units of biodiversity. Chapman & Hall, London: 439 pp.

Clement, M., Posada, D., and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9 (10): 1657-1659

Colbourne, E. B., and D. R. Senciall. 1996. Temperature, salinity and sigma-t along the standard Flemish Cap transect. Can. Tech. Rep. Hydrogr. Ocean Sci. 172: 222 pp.

Colbourne, E., and K. D. Foote. 2000. Variability of the stratification and circulation on the Flemish Cap during the decades of the 1950s – 1990s. J. Northwest Atl. Fish. Sci. 26: 103-122

Congiu, L., Dupanloup, I., Patarnello, T., Fontana, F., Rossi, R., Arlati, G., and L. Zane. 2001. Identification of interspecific hybrids by amplified fragment length polymorphism: the case of sturgeon. Mol. Ecol. 10 (9): 2355-2359

Cope, J. M. 2004. Population genetics and phylogeography of the blue rockfish (*Sebastes mystinus*) from Washington to California. Can. J. Fish. Aquat. Sci. 61: 332-342

Cornuet, J. M., Piry, S., Luikart, G., Estoup, A., and M. Solignac. 1999. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics 153: 1989-2000

Crandall, K. A., and A. R. Templeton. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics 134: 959-969

Crozier, R. H. 1997. Preserving the information content of species: Genetic diversity, phylogeny, and conservation worth. Annu. Rev. Ecol. Syst. 28: 243-268

Cuvier, G. 1829. Le Règne animal, distribuè d'après son organisation, pour servir de base à l'histoire naturelle des animaux et d'introduction à l'anatomie comparèe. Edition 2, v. 2: I–xvii + 1-532

Delaroche, F. E. 1809. Suite du mémoire sur les espèces de poissons observées à Iviça. Observations sur quelques-uns des poissons indiqués dans le précédent tableau et descriptions des espèces nouvelles ou peu connues. Ann. Mus. Hist. Nat. Paris: 313-361

Desrosiers, B., Sévigny, J.-M., and J.-P. Chanut. 1999. Restriction fragment length polymorphism of rDNA in the redfishes *Sebastes fasciatus* and *S. mentella* (Scorpaenidae) from the Gulf of St. Lawrence. Can. J. Zool. 77: 267-277

DeWoody, J. A., and J. C. Avise. 2000. Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. J. Fish Biol. 56: 461-473

Di Rienzo, A., Peterson, A. C., Garza, J. C., Valdes, A. M., Slatkin, M., and N. B. Freimer. 1994. Mutational processes of simple-sequence repeat loci in human populations. Proc. Natl. Acad. Sci. USA 91 (8): 3166-3170

Docker, M. F., Youson, J. H., Beamish, R. J., and R. H. Devlin. 1999. Phylogeny of the lamprey *Lampetra* inferred from mitochondrial cytochrom b and ND3 gene sequences. Can. J. Fish. Aquat. Sci. 56: 2340-2349

Doiron, S., Bernatchez, L., and P. U. Blier. 2002. A comparative mitogenomic analysis of the potential adaptive value of Arctic charr mtDNA introgression in brook charr populations (*Salvelinus fontinalis* Mitchill). Mol. Biol. Evol. 19 (11): 1902-1909

Domanico, M. J., and R. B. Phillips. 1995. Phylogenetic analysis of Pacific salmon (genus *Oncorhynchus*) based on mitochondrial DNA sequence data. Mol. Phylogenet. Evol. 4: 366-371.

Domanico, M. J., Phillips, R. B., and T.H. Oakley. 1997. Phylogenetic analysis of Pacific salmon (genus *Oncorhynchus*) using nuclear and mitochondrial DNA sequences. Can. J. Fish. Aquat. Sci. 54: 1865-1872

Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., and J. S. Mattick. 1991. Touchdown PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res. 19: 4008

Donnelly, P., and S. Tavaré. 1986. The ages of alleles and a coalescent. Adv. Appl. Probab. 18: 1-19

Dowling, T. E., and W. S. Moore. 1985. Evidence for selection against hybrids in the family Cyprinidae (genus *Notropis*). Evolution 39: 152-158

Dowling, T. E., and W. M. Brown. 1989. Allozymes, mitochondrial DNA, and levels of phylogenetic resolution among four minnow species (*Notropis*: Cyprinidae). Syst. Zool. 38 (2): 126-143

Downs, T. R., and W. W. Wilfinger. 1983. Fluorometric Quantification of DNA in Cells and Tissue. Anal. Biochem. 131: 538-547

Dushchenko, V. V. 1987. Polymorphism of NADP-dependent Malate-Dehydrogenase in *Sebastes mentella* (Scorpaenidae) from the Irminger Sea. J. Ichthyol. 27 (1): 129-131

Edmands, S. 1999. Heterosis and outbreeding depression in interpopulation crosses spanning a wide range of divergence. Evolution 53: 1757-1768

Eschmeyer, W. N., and J. C. Hureau. 1971. *Sebastes mouchezi*, a senior synomym of *Helicolenus tristanensis*, with comments on *Sebastes capensis* and zoo-geographical considerations. Copeia 3: 576-579

Estoup, A., and B. Angers. 1998. Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. Pp. 55-85 *in* Carvalho, G.R., ed. Advances in Molecular Ecology. IOS Press, Amsterdam, Netherlands: 314 pp.

Estoup, A., Jarne, P., and J. M. Cornuet. 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetic analysis. Mol. Ecol. 11 (9): 1591-1604

Excoffier, L., Smouse, P. E., and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplogroups: applications to human mitochondrial DNA restriction data. Genetics 131: 479-491

Excoffier, L., and P. E. Smouse. 1994. Using allele frequencies and geographic subdivision to reconstruct gene trees within a species: molecular variance parsimony. Genetics 136: 343-359

Falush, D., Stephens, M., and J. K. Pritchard. 2003. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics 164: 1567-1587

Farias, I. P., Ortí, G., Sampaio, I., Schneider, H., and A. Meyer. 1999. Mitochondrial DNA phylogeny of the family Cichlidae: monophyly and fast molecular evolution of the neotropical assemblage. J. Mol. Evol. 48: 703-711

Farris, J. S. 1989a. The retention index and homoplasy excess. Syst. Zool. 38: 406-407

Farris, J. S. 1989b. The Retention Index and the Rescaled Consistency Index. Cladistics 5: 417-419

Felsenstein, J. 1973. Maximum-likelihood estimation of evolutionary trees from continuous characters. Am. J. Hum. Genet. 25: 471-492

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17: 368-376

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39: 783-791

Felsenstein, J. 1988. Phylogenies from molecular sequences: Inference and reliability. Annu. Rev. Genet. 22: 521-565

Ferris, S. D., and W. J. Berg. 1987. The utility of mitochondrial DNA in fish genetics and fishery management. Pp. 277-301 *in* Ryman, E., and E. Utter, eds. Population genetics and fishery management. Univ. Washington Press, Seattle: 420 pp.

Fitch, W. M. 1971. Towards defining the course of evolution: Minimum change for a specific tree topology. Syst. Zool. 20: 406-416

Gaggiotti, O. E., Lange, O., Rassman, K., and C. Gliddon. 1999. A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. Mol. Ecol. 8 (9): 1513-1520

Garabana, D. 2005. The genus *Sebastes* Cuvier, 1829 (Pisces, Scorpaenidae) in the North Atlantic: Species and stock discrimination using traditional and geometric morphometrics. Ph.D. Thesis, University of Vigo, Spain: *In prep.*

Gascon, D., ed. 2003. Redfish multidisciplinary research zonal program (1995-1998): final report. Can. Tech. Rep. Fish. Aquat. Sci. 2462: 139 pp.

Gaudeul, M., Taberlet, P., and I. Till-Bottraud. 2000. Genetic diversity in an endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), inferred from amplified fragment length polymorphism markers. Mol. Ecol. 9 (10): 1625-1637

Gauthier, S., and G. A. Rose. 2002. Acoustic observation of diel vertical migration and shoaling behaviour in Atlantic redfishes. J. Fish Biol. 61: 1135-1153

Geiger, H. J., Smoker, W. W., Zhivotovsky, L. A., and A. J. Gharrett. 1997. Variability of family size and marine survival in pink salmon (*Oncorhynchus gorbuscha*) has implications for conservation biology and human use. Can. J. Fish. Aquat. Sci. 54: 2684-2690

Gharrett, A.J., Gray, A.K., and J. Heifetz. 2001. Identification of rockfish (*Sebastes* spp.) by restriction site analysis of the mitochondrial ND-3/ND-4 and 12S/16S gene regions. Fish. Bull. 99: 49-62

Gharrett, A.J., Matala, A. P., Peterson, E. L., Gray, A. K, Li, Z., and J. Heifetz. 2005. Two genetically distinct forms of Rougheye Rockfish (*Sebastes aleutianus*) are different species. Trans. Amer. Fish. Soc. 134: 242-260

Gilbert, C. H. 1890. A preliminary report on the fishes collected by the steamer Albatross on the Pacific coast of North America during the year 1889, with descriptions of twelve new genera and ninety-two new species. Proc. U. S. Natl. Mus. 13: 49-126

Gmelin, J. F. 1789. Caroli Linne Systema naturae. Systema Naturae per regna tria naturae, secundum classes, ordines, genera, species; cum characteribus, differentiis, synonymis, locis. Editio decimo tertia, aucta, reformata. 3 volumes in 9 parts. Lipsiae, 1788-93. Systema Naturae Linné: 1033-1516

Goldstein, D. B., and D. D. Pollock. 1997. Launching microsatellites: A review of mutation processes and methods of phylogenetic inference. J. Hered. 88 (5): 335-342

Goudet, J. 1995. FSTAT version 1.2: a computer program to calculate F-statistics. J. Hered. 86: 485-486

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995).

Grant, W. S., and B. W. Bowen. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. J. Hered. 89 (5): 415-426

Grant, W. S., Garcia-Marin, J. L., and F. M. Utter. 1999. Defining Population Boundaries for Fishery Management. Pp. 27-72 *in* Mustafa, S., ed. Genetics in sustainable fisheries management. Fishing News Books, Blackwell Science, Oxford, UK: 223 pp.

Griffith, R., Orr, K. J., Adam, A., and I. Barber. 2000. DNA sex identification in the three-spined stickleback. J. Fish Biol. 57: 1331-1334

Guinand, B. 1996. Use of a multivariate model using allele frequency distributions to analyse patterns of genetic differentiation among populations. Biol. J. Linn. Soc. 58: 173-195

Guo, S. W., and E. A. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48: 361-372

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41: 95-98

Hallacher, L. E. 1974. The comparative morphology of extrinsic swimbladder musculature in the scorpionfish genus *Sebastes* (Pisces: Scorpaenidae). Proc. Calif. Acad. Sci. 40: 59-86

Han, K., and B. Ely. 2002. Use of AFLP analyses to assess genetic variation in *Morone* and *Thunnus* species. Mar. Biotechnol. 4: 141-145

Hansen, M. M., Kenchington, E., and E. E. Nielsen. 2001. Assigning individual fish to populations using microsatellite DNA markers: Methods and applications. Fish and Fisheries 2: 93-112

Hansen, M. M. 2003. Application of molecular markers in population and conservation genetics, with special emphasis on fishes. Doctoral thesis (DSc thesis, Dr. scient. afhandling), University of Aarhus, Denmark

Hardy, O. J., Charbonnel, N., Freville, H., and M. Heuertz. 2003. Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. Genetics 163 (4): 1467-1482

Hartl, D. E., and A. G. Clark. 1997. Principles of population genetics (2nd edition). Sinauer Associates, Inc., Sunderland: 682 pp.

Hasegawa, M., Kishino, M., and T. Yano. 1985. Dating the human-ape split by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22: 160-174

Hauge, X. Y., and M. Litt. 1993. A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphism by the PCR. Hum. Mol. Genet. 2 (4): 411-415

Hauser, L., Adcock, G. J., Smith, P. J., Ramirez, J. H. B., and G. R. Carvalho. 2002. Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). Proc. Natl. Acad. Sci. USA 99: 11742-11747

Hedgecock, D. 1994. Does variance in reproductive success limit effective population sizes of marine organisms? Pp. 122–134 *in* Beaumont, A. R., ed. Genetics and Evolution of Aquatic Organisms. Chapman & Hall, London: 539 pp.

Hedgecock, D., Hutchinson, E. S., Li, G., Sly, F. L., and K. Nelson. 1994. The central stock of northern anchovy (*Engraulis mordax*) is not a randomly mating population. Calif. Coop. Oceanic Fish. Invest. Rep. 35: 121-136

Helvey, M. 1982. First observations of courtship behavior in rockfish, genus *Sebastes*. Copeia 4: 763-770

Hillis, D. M., and J. J. Bull. 1993. An empirical test of bootstrapping as a method assessing confidence in phylogenetic analysis. Syst. Biol. 41: 182-192

Hoarau, G., Holla, S., Lescasse, R., Stam, W. T., and J. L. Olsen. 2002a. Heteroplasmy and evidence for recombination in the mitochondrial control region of the flatfish *Platichthys flesus*. Mol. Biol. Evol. 19: 2261-2264

Hoarau, G., Rijnsdorp, A. D., Van Der Veer, H. W., Stam, W. T. and J. L. Olsen. 2002b. Population structure of plaice (*Pleuronectes platessa* L.) in northern Europe: microsatellites revealed large-scale spatial and temporal homogeneity. Mol. Ecol. 11 (7): 1165-1176

Hutchings, J. A., Bishop T. D., and C. R. McGregor-Shaw. 1999. Spawning behaviour of Atlantic cod *Gadus morhua*: evidence of mate competition and mate choice in a broadcast spawner. Can. J. Fish. Aquat. Sci. 56: 97-104

Hutchinson, W. F., van Oosterhout, C., Rogers, S. I., and G. R. Carvalho. 2003. Temporal analysis of archived samples indicates marked genetic changes in declining North Sea cod (*Gadus morhua*). Proc. R. Soc. Lond. Ser. B 270: 2125 - 2132

ICES. 1983. Report on the joint NAFO/ICES study group on biological relationships of the West Greenland and Irming er Sea redfish stocks. ICES C.M. 1983/G:3: 13 pp.

ICES. 1992. Report of the study group on redfish stocks. ICES C.M. 1992/G:14: 7 pp.

ICES. 1998a. Report of the study group on redfish stocks. ICES C.M. 1998/G:3, 30 pp.

ICES. 1998b. Report of the North-Western Working Group. ICES C.M. 1998/ACFM:19: 350 pp.

ICES. 2001. Report of the North-Western Working Group. ICES C.M. 2001/ACFM:20: 11 pp.

ICES. 2002. Report of the North-Western Working Group. ICES C.M. 2002/ACFM:20: 405 pp.

ICES. 2004a. Report of the Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM). ICES C.M. 2004/F:04: 44 pp.

ICES. 2004b. Report of the study group on stock identity and management units of redfishes (SGSIMUR). ICES C.M. 2005/ACFM:10: 334pp.

ICES. 2004c. Report of the North-Western Working Group. ICES C.M. 2004/ACFM: 25: 476 pp.

Ihssen, P. E., Booke, H. E., Casselman, J. M., McGlade, J. M., Payne, N. R., and F. M. Utter. 1981. Stock identification: materials and methods. Can. J. Fish. Aquat. Sci. 38: 1838-1855

Jarne, P., and P. J. L. Lagoda. 1996. Microsatellites, from molecules to populations and back. Trends Ecol. Evol. 11 (10): 424-429

Jensen, J. M. 1979. Redfish at West Greenland, Subarea 1. ICNAF Res. Doc. 79/VI/54: 22 pp.

Jiggins, C. S., and J. Mallet. 2000. Bimodal hybrid zones and speciation. Trends Ecol. Evol. 15 (6): 250-255

Joensen, H., and O. Grahl-Nielsen. 2001. The redfish species *Sebastes viviparus*, *Sebastes marinus* and *Sebastes mentella* have different composition of their tissue fatty acids. Comp. Biochem. Physiol. Part B 129: 73-85

Johansen, T., Daníelsdóttir, A. K., Meland, K., and G. Nævdal. 1998. Progress report on the studies of the genetic relationship of deep-sea and oceanic *Sebastes mentella* in the Irminger Sea. ICES C.M. 1998/K:18: 18pp.

Johansen, T., Nævdal, G., Daníelsdóttir, A. K., and N. R. Hareide. 2000a. Genetic characterisation of giant *Sebastes* in the deep water slopes in the Irminger Sea. Fish. Res. 45: 207-216

Johansen, T., Danielsdottir, A. K., Meland, K., and G. Naevdal. 2000b. Studies of the genetic relationship between deep-sea and oceanic *Sebastes mentella* in the Irminger Sea. Fish. Res. 49: 179-192

Johansen, T. 2003. Genetic study of genus *Sebastes* (redfish) in the North Atlantic with emphasis on the stock complex in the Irminger Sea. Ph.D. Thesis, University of Bergen, Norway

Johns, G. C., and J. C. Avise. 1998. Tests for ancient species flocks based on molecular phylogenetic appraisals of *Sebastes* rockfishes and other marine fishes. Evolution 52 (4): 1135-1146

Kai, Y., Nakayama, K., and T. Nakabo. 2002a. Genetic differences among three colour morphotypes of the black rockfish, *Sebastes inermis*, inferred from mtDNA and AFLP analysis. Mol. Ecol. 11 (12): 2591-2598

Kai, Y., Yagishita, N., Ikeda, H., and T. Nakabo. 2002b. Genetic differences between two color morphotypes of redfish, *Sebastes scythropus* (Osteichthyes: Scorpaenidae). Species Diversity 7: 371-380

Kai, Y., Nakayama, K., and T. Nakabo. 2003. Molecular phylogenetic perspective on speciation in the genus *Sebastes* (Scorpaenidae) from the northwest Pacific and the position of *Sebastes* within the subfamily Sebastinae. Ichthyol. Res. 50: 239-244

Kai, Y., and T. Nakabo. 2004. A new species of *Sebastes* (Scorpaeniformes: Scorpaenidae) from the Pacific coast of southern Japan. Ichthyol. Res. 51: 5-9

Kelly, G. F., and A. M. Barker. 1961: Vertical distribution of young redfish in the Gulf of Maine. Rapp. P.-v. Réun. Cons. Int. Explor. Mer 150: 220-233

Kenchington, T. J. 1986. Morphological comparison of two Northwest Atlantic redfishes, *Sebastes fasciatus* and *S. mentella*, and techniques for their identification. Can. J. Fish. Aquat. Sci. 43: 781-787

Kendall, A. W., Jr. 1991. Systematics and identification of larvae and juveniles of the genus *Sebastes*. Environ. Biol. Fish. 30: 173-190

Kendall, A. W., Jr. 2000. An historical review of *Sebastes* taxonomy and systematics. Mar. Fish. Rev. 62 (2): 1-23

Kimura, M., and J. Crow. 1964. The number of alleles that can be maintained in a finite population. Genetics 49: 725-738

Kimura, M., and T. Ohta. 1978. Stepwise mutation model and distribution of allelic frequencies in finite population. Proc. Natl. Acad. Sci. USA 75: 2868-2872

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16: 111-120

Kluge, A. G., and J. S. Farris. 1969. Quantitative Phyletics and the Evolution of Anurans. Syst. Zool. 18: 1-32

Knowlton, N. 1993. Sibling species in the sea. Annu. Rev. Ecol. Syst. 24: 189-216

Knowlton, N. 2000. Molecular genetic analyses of species boundaries in the sea. Hydrobiologia 420: 73-90

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Paabo, S., Villablanca, F. X., and A. C. Wilson. 1989. Dynamics of mitochondrial DNA evolution in mammals: amplification and sequencing with conserved primers. Proc. Natl. Acad. Sci. USA 86: 6196-6200

Kocher, T. D., and K. L. Carleton. 1997. Base substitution in fish mitochondrial DNA: patterns and rates. Pp. 13–24 *in* Kocher, T. D., and C.A. Stepien, eds. Molecular systematics of fishes. Academic Press, San Diego, CA, USA: 314 pp.

Kocher, T. D. 2003. Fractious phylogenies. Nature 423: 489-491

Kosswig, K. 1974. Age and growth of redfish (type giants) off SW-Iceland. ICES C.M. 1974/F:9: 3 pp.

Kotthaus, A. 1950. Oekologische und fischereiwissenschaftliche Untersuchungen über den Rotbarsch. III. Rassenuntersuchungen am Rotbarsch. Ber. Dt. wiss. Komm. Meeresforsch. 12: 94-114

Kotthaus, A. 1961. Contributions to the race problem in redfish. Rapp. P.-v. Réun. Cons. Int. Explor. Mer 150: 42-44

Krauss, S. L. 2000. Accurate gene diversity estimates from amplified fragment length polymorphism (AFLP) markers. Mol. Ecol. 9 (9): 1241-1245

Krøyer, H. N. 1845. Ichthyologiske Bidrag. Naturhist. Tidsskr. Kjøbenhavn (N. S.): 213-282

Kumar, S., Tamura, K., and M. Nei. 1993. MEGA: Molecular Evolutionary Genetics Analysis. Pennsylvania State University, Pennsylvania and Tempe, Arizona, USA

Kumar, S., Tamura, K., Jakobsen, I. B., and M. Nei. 2001. MEGA2: Molecular Evolutionary Genetics Analysis software. Bioinformatics 17 (12): 1244-1245

Labarca, C., and K. Paigen. 1980. A simple, rapid and sensitive DNA assay procedure. Anal. Biochem. 102: 344-352

Linné, C. 1758. Systema Naturae, Ed. X. (Systema naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Tomus I. Editio decima, reformata.) Holmiae. Systema Nat. ed. 10 i-ii + 1-824

Liu, Z., Nichols, A., Li, P., and R. A. Dunham. 1998. Inheritance and usefulness of AFLP markers in channel catfish (*Ictalurus punctatus*), blue catfish (*I. furcatus*), and their F1, F2, and backcross hybrids. Mol. Gen. Genet. 258 (3): 260-268

Liu, Z. J., and J. F. Cordes. 2004. DNA marker technologies and their applications in aquaculture genetics. Aquaculture 238: 1-37

Long, J. C. 1991. The genetic structure of admixed populations. Genetics 127: 417-428

Love, M. S., Yoklavich, M., and L. Thorsteinson. 2002. The rockfishes of the Northeast Pacific. University of California Press, Berkeley, CA, USA: 405 pp.

Lu, J., Knox , M. R., Ambrose, M. J., Brown, J. K. M., and T. H. N. Ellis. 1996. Comparative analysis of genetic diversity in pea assessed by RFLP and PCR-based methods. Theor. Appl. Genet. 93: 1103-1111

Lu, G., Basley, D. J., and L. Bernatchez. 2001. Contrasting patterns of mitochondrial DNA and microsatellite introgressive hybridisation between lineages of lake whitefish (*Coregonus clupeaformis*); relevance for speciation. Mol. Ecol. 10 (4): 965-985

Lugon-Moulin, N., Brünner, H., Wyttenbach, A., Hausser, J., and J. Goudet. 1999. Hierarchical analysis of genetic differentiation in a hybrid zone of *Sorex araneus* (Insectivora, Soricidae). Mol. Ecol. 8 (3): 419-431

Luikart, G., and P. R. England. 1999. Statistical analysis of microsatellite DNA data. Trends Ecol. Evol. 14 (7): 253-256

Lundy, C. J., Moran, P., Rico, C., Milner, R. S., and G. M. Hewitt. 1999. Macrogeographical population differentiation in oceanic environments: a case study of European hake (*Merluccius merluccius*), a commercially important fish. Mol. Ecol. 8 (11): 1889-1898

Lydeard, C., and K. J. Roe. 1997. The phylogenetic utility of the mitochondrial cytochrome b gene for inferring relationships among Actinopterygian fishes. Pp. 285-303 *in* Kocher, T. D., and C.A. Stepien, eds. Molecular systematics of fishes. Academic Press, San Diego, CA: 314 pp.

Lynch, M., and B. G. Milligan. 1994. Analysis of population genetic structure with RAPD markers. Mol. Ecol. 3 (2): 91-99

Magnússon, J. 1977. Pelagic redfish in the Irminger Sea: distribution and abundance. ICES C.M. 1977/H:43:10 pp.

Magnússon, J. 1983. The Irminger Sea oceanic stock of redfish "spawning" and "spawning" area. ICES C.M. 1983/G:56: 11 pp.

Magnússon, J., Nedreaas, K. H., Magnússon, J. V., Reynisson, P., and T. Sigurdsson. 1994. Report on the joint Icelandic/Norwegian survey on oceanic redfish in the Irminger Sea and adjacent waters, in June/July 1994. ICES C.M. 1994/G:44: 29 pp.

Magnússon, J., and J. V. Magnússon. 1995. Oceanic redfish (*Sebastes mentella*) in the Irminger Sea and adjacent waters. Sci. Mar. 59: 241-254

Maniatis, T., Fritsch, E. F., and J. Sambrook. 1982. Molecular Cloning, A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Manly, B. F. J. 1994. Multivariate statistical methods: a primer (2nd edition). Chapman & Hall, New York: 208pp.

Mantel, N. 1967. The detection of disease clustering and generalized regression approach. Cancer Res. 27: 209-220

Marchuk, D., Drumm, M., Saulino, A., and F. S. Collins. 1991. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. Nucleic Acids Res. 19: 1154

Matala, A. P., Gray, A. K., Heifetz, J., and A. J. Gharrett. 2004a. Population structure of Alaskan shortraker rockfish, *Sebastes borealis*, inferred from microsatellite variation. Environ. Biol. Fish. 69: 201-210

Matala, A. P., Gray, A. K., Gharrett, A. J., and M. S. Love. 2004b. Microsatellite variation indicates population genetic structure of Bocaccio. N. Am. J. Fish. Manage. 24: 1189-1202

McGlade, J. M., Annand, M. C., and T. J. Kenchington. 1983. Electrophoretic identification of *Sebastes* and *Helicolenus* in the northwestern Atlantic. Can. J. Fish. Aquat. Sci. 40: 1861-1870

McKay, S. J., Devlin, R. H., and M. J. Smith. 1996. The phylogeny of Pacific salmon and trout based on growth hormone type-2 (GH2) and mitochondrial NADH dehydrogenase subunit 3 (ND3) DNA sequences. Can. J. Fish. Aquat. Sci. 53: 1165-1174

Meyer, A. 1993. Evolution of mitochondrial DNA in fishes. Pp. 1–38 *in* Hochachka, P. W., and T. P. Mommsen, eds. Biochemistry and molecular biology of fishes. Volume 2. Elsevier Press, Amsterdam

Meyer, A. 1994. Molecular phylogenetic studies of fish. Pp. 219–249 *in* Beaumont, A. R., ed. Genetics and evolution of aquatic organisms. Chapman & Hall, New York: 560 pp.

Mickett, K., Morton, C., Feng, J., Li, P., Simmons, M., Dunham, R. A., Cao, D., and Z. J. Liu. 2003. Assessing genetic diversity of domestic populations of channel catfish (*Ictalurus punctatus*) in Alabama using AFLP markers. Aquaculture 228: 91-105

Miller, M. P. 1997. Tools for population genetic analysis (TFPGA) 1.3: a Windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by the author.

Miller, K. M., Schulze, A. D., and R. E. Withler. 2000. Characterisation of microsatellite loci in *Sebastes alutus* and their conservation in congeneric rockfish species. Mol. Ecol. 9 (2): 240-242

Mindell, D. P., and C. E. Thacker. 1996. Rates of molecular evolution: Phylogenetic issues and applications. Annu. Rev. Ecol. Syst. 27: 279–303

Morris, D. B., Richard, K. R., and J. M. Wright. 1996. Microsatellites from rainbow trout (*Oncorhynchus mykiss*) and their use for genetic study of salmonids. Can. J. Fish. Aquat. Sci. 53: 120-126

Moser, H. G., and G. W. Boehlert. 1991. Ecology of pelagic larvae and juveniles of the genus *Sebastes*. Environ. Biol. Fish. 30: 203-224

Mueller, U. G., and L. L. Wolfenbarger. 1999. AFLP genotyping and fingerprinting. Trends Ecol. Evolut. 14: 389-394

Mullis, K. B., Falcoona, F. A., Scharf, S., Saiki, R. K., Horn, G. T., and H. A. Erlich. 1986. Specific enzymatic amplification of DNA *in vitro* – the polymerase chain reaction. Cold Spring Harb. Symp. Quant. Biol. 51: 263-273

Mullis, K. B., and F. A. Falcoona. 1987. Specific synthesis of DNA *in vitro* via a polymerase-catalysed chain reaction. Methods Enzymol. 155: 335-350

Murray, V. 1989. Improved double-stranded DNA sequencing using linear polymerase chain reaction. Nucleic Acids Res. 17: 8889

Murray, V., Monchawin, C., and P. R. England. 1993. The determination of the sequences present in the shadow bands of a dinucleotide repeat PCR. Nucleic Acids Res. 21: 2395-2398.

Muus, B. J., and J. G. Nielsen. 1999. Sea fish. Scandinavian Fishing Year Book, Hedehusene, Denmark: 340 pp.

Naciri, M., Lemaire, C., Borsa, P., and F. Bonhomme. 1999. Genetic study of the Atlantic/Mediterranean transition in sea bass (*Dicentrarchus labrax*). J. Hered. 90: 591-596

Nævdal, G. 1978. Differentiation between *marinus* and *mentella* types of redfish by electrophoresis of haemoglobins. Fiskeridir. Skr. Ser. Havunders. 16: 359-308

Nakamura, K., Ozaki, A., Akutsu, T., Iwai, K., Sakamoto, T., Yoshizaki, G., and N. Okamoto. 2001. Genetic mapping of the dominant albino locus in rainbow trout (*Oncorhynchus mykiss*). Mol. Genet. Genomics 265 (4): 687-693

Nedreaas, K., and G. Nævdal. 1989. Studies of Northeast Atlantic species of redfish (genus *Sebastes*) by protein polymorphism. J. Cons. int. Explor. Mer 46: 76-93

Nedreaas, K. 1990. Age determination of Northeast Atlantic *Sebastes* species. J. Cons. int. Explor. Mer 47: 208-230

Nedreaas, K., and G. Nævdal. 1991a. Genetic studies of redfish (*Sebastes* spp.) along the continental slopes from Norway to East Greenland. ICES J. Mar. Sci. 48: 173-186

Nedreaas, K., and G. Nævdal. 1991b. Identification of 0- and 1-group redfish (genus *Sebastes*) using electrophoresis. ICES J. Mar. Sci. 48: 91-99

Nedreaas, K., Johansen, T., and G. Nævdal. 1994. Genetic studies of redfish (*Sebastes* spp.) from Icelandic and Greenland waters. ICES J. Mar. Sci. 51: 461-467

Nefyodov, G. N. 1971. Serum haptoglobins in the *marinus* and *mentella* types of North Atlantic redfish. Rapp. P-v. Reun. Cons. Int. Explor. Mer 161: 126–129

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583-590

Nei, M., and F. Tajima. 1981. DNA polymorphism detectable by restriction endonucleases. Genetics 105: 207-217.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York: 512 pp.

Nei, M., and S. Kumar. 2000. Molecular Evolution and Phylogenetics. Oxford University Press, New York: 333pp.

Ni, I.-H. 1981. Separation of the sharp-beaked redfish, *S. fasciatus* and *S. me*ntella, from Northeastern Grand Bank by morphology of extrinsic gasbladder musculature. J. Northw. Atl. Fish. Sci. 2: 7-12

Nielsen, E. E., Hansen, M. M., and V. Loeschke. 1997. Analysis of microsatellite DNA from old scale samples of Atlantic salmon *Salmo salar*: a comparison of genetic composition over 60 years. Mol. Ecol. 6 (5): 487-492

Nielsen, E. E., Hansen, M. M., Schmidt, C., Meldrup, D., and P. Grønkjær. 2001. Determining the population of origin of individual cod in the Northeast Atlantic. Nature 413: 272

Nielsen, E. E., Hansen, M. M., Ruzzante, D. E., Meldrup, D., and P. Grønkjær. 2003. Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea, revealed by individual admixture analysis. Mol. Ecol. 12 (6): 1497-1508

Nielsen, E. E., Nielsen, P. H., Meldrup, D., and M. M. Hansen. 2004. Genetic population structure of turbot (*Scophthalmus maximus* L.) supports the presence of multiple hybrid zones for marine fishes in the transition zone between the Baltic Sea and the North Sea. Mol. Ecol. 13 (3): 585-595

Nielsen, E. E. 2004. Review of the genetic data of *S. mentella* in the Irminger Sea and adjacent waters. Working Document No. 8 in "Report of the ICES Study Group on Stock Identity and Management Units of Redfishes (SGSIMUR), 31rst of August – 3rd of September 2004, Bergen, Norway". ICES C.M. 2005/ACFM:10: 334pp.

O'Connell, M., and J. M. Wright. 1997. Microsatellite DNA in fishes. Rev. Fish Biol. Fish. 7: 331-363

O'Reilly, P. T., and J. M. Wright. 1995. The evolving technology of DNA fingerprinting and its application to fisheries and aquaculture. J. Fish Biol. 47 (Suppl. A): 29-55

O'Reilly, P. T., Canino, M. F., Bailey, K. M., and P. Bentzen. 2004. Inverse relationship between F_{ST} and microsatellite polymorphism in the marine fish, walleye pollock (*Theragra chalcogramma*): implications for resolving weak population structure. Mol. Ecol. 13 (7): 1799-1814

O'Rourke, F. J. 1961. An immunological and chromatographic study of *S. marinus* (L.) and *S. mentella* Travin. Rapp. P.-v. Réun. Cons. Int. Explor. Mer 150: 100-103

Orr, J. W., and J. E. Blackburn. 2004. The dusky rockfishes (Teleostei: Scorpaeniformes) of the North Pacific Ocean: resurrection of *Sebastes variabilis* (Pallas, 1814) and a rediscription of *Sebastes ciliatus* (Tilesius, 1813). Fish. Bull. 102: 328-348

Ovenden, J. R., Smolenski, A. J., and E. W. G. White 1989. Mitochondrial DNA restriction site variation in Tasmanian populations of orange roughy (*Haplostethus atlanticus*), a deep-water marine teleost. Austr. J. Mar. Freshw. Res. 40: 1-9

Ovenden J. R. 1990. Mitochondrial DNA and marine stock assessment: A review. Aust. J. Mar. Freshw. Res. 41 (6): 835-53

Palumbi, S. R. 1994. Genetic divergence, reproductive isolation and speciation in the sea. Annu. Rev. Ecol. Syst. 25: 547-72

Payne, R. H., and I.-H. Ni. 1982. Biochemical population genetics of redfishes (*Sebastes*) off Newfoundland. J. Northw. Atl. Fish. Sci. 3: 169-172

Pemberton, J. M, Slate, J., Bancroft, D. R., and J. A. Barrett. 1995. Non-amplifying alleles at microsatellite loci: a caution for parentage and population studies. Mol. Ecol. 4 (2): 249-252

Petersen, M. F., and J. F. Steffensen. 2003. Preferred temperature of juvenile Atlantic cod *Gadus morhua* with different haemoglobin genotypes at normoxia and moderate hypoxia. J. Exp. Biol. 206: 359-364

Phillips, R. B., and T. H. Oakley. 1997. Phylogenetic relationships among the Salmoninae based on nuclear and mitochondrial DNA sequences. Pp. 145–162 *in* Kocher, T. D., and C.A. Stepien, eds. Molecular systematics of fishes. Academic Press, San Diego, CA: 314 pp.

Pierpaoli, M., Birò, Z. S., Herrmann, M., Hupe, K., Fernandes, M., Ragni, B., Szemethy, L., and E. Randi. 2003. Genetic distinction of wildcat (*Felis silvestris*) populations in Europe, and hybridisation with domestic cats in Hungary. Mol. Ecol. 12 (10): 2585-2598

Posada, D., and K. A. Crandall. 2001. Intraspecific gene genealogies: trees grafting into networks. Trends Ecol. Evol. 16: 37-45

Power, D. J., and I.-H. Ni. 1982. Morphology of the extrinsic gasbladder musculature in golden redfish, *Sebastes marinus*. NAFO Sci. Counc. Res. Doc. 82/IX/110: 8 pp.

Power, D. J., and I.-H. Ni. 1985. Morphometric differences between golden redfish (*Sebastes marinus*) and beaked redfishes (*S. mentella* and *S. fasciatus*). J. Northw. Atl. Fish. Sci. 6: 1-7

Pritchard, J. K., Stephens, M., and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945-959

Pritchard, J. K., and W. Wen. 2003. Documentation for *STRUCTURE* software, Vers. 2. Department of Human Genetics, The University of Chicago. Available at http://pritch.bsd.uchicago.edu/structure.html

Ramon, M. L., Lobel, P. S., and M. D. Sorenson. 2003. Lack of mitochondrial genetic structure in hamlets (*Hypoplectrus spp*.): recent speciation or ongoing hybridisation? Mol. Ecol. 12 (11): 2975-2980

Rand, D. M. 1994. Thermal habit, metabolic rate and the evolution of mitochondrial DNA. Trends Ecol. Evol. 9 (4): 125-131

Rand, D. M. 2001a. The units of selection on mitochondrial DNA. Annu. Rev. Ecol. Syst. 32: 415-448

Rand, D. M. 2001b. Mitochondrial genomics flies high. Trends Ecol. Evol. 16: 2-4

Randi, E., Tabarroni, C., Rimondi, S., Lucchini, V., and A. Sfougaris. 2003. Phylogeography of the rock partridge (*Alectoris graeca*). Mol. Ecol. 12 (8): 2201-2214

Ransom, D. G., and L. I. Zon. 1999. Mapping zebrafish mutations by AFLP. Methods Cell. Biol. 60: 195-211

Raymond, M., and F. Rousset. 1995a. An exact test for population differentiation. Evolution 49 (6): 1280-1283

Raymond, M., and F. Rousset. 1995b. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J. Hered. 86: 248-249

Rehbein, H. 1983. Differentiation of redfishes from the Northeast Atlantic *(Sebastes marinus* L., *S. mentella* Travin, *S. viviparus* Krøyer and *Helicolenus dactylopterus* D. Delaroche 1809) by isoelectric focusing of sarcoplasmic proteins. ICES C.M. 1983/G:40: 12 pp.

Reinert, J., and Lastein, L. 1992. Stock identification of *S. marinus* L. and *S. mentella* Travin in the Northeast Atlantic based on meristic counts and morphometric measurements. ICES C.M. 1992/G:29: 21 pp.

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43 (1): 223-225

Riginos, C., and M. W. Nachman. 2001. Population subdivision in marine environments: the contributions of biogeography, geographical distance and discontinuous habitat to genetic differentiation in a blennioid fish, *Axoclinus nigricaudus*. Mol. Ecol. 10 (6): 1439-1453

Robinson, J. P., and S. A. Harris. 1999. Amplified fragment length polymorphisms and microsatellites: A phylogenetic perspective. Chapter 12 *in* Gillet E. M., ed. Which DNA Marker for Which Purpose? Final Compendium of the Research Project Development, optimisation and validation of molecular tools for assessment of biodiversity in forest trees in the European Union DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity

Rocha-Olivares, A., and R. D. Vetter. 1999. The effects of oceanographic circulation on the gene flow, genetic structure, and phylogeography of the rosethorn rockfish (*Sebastes helvomaculatus*). Can. J. Fish. Aquat. Sci. 56: 803-813

Rocha-Olivares, A., Kimbrell, C. A., Eitner, B. J., and R. D. Vetter. 1999a. Evolution of a mitochondrial cytochrome b gene sequence in the species-rich genus *Sebastes* (Teleostei, Scorpaenidae) and its utility in testing the monophyly of the subgenus *Sebastomus*. Mol. Phylogen. Evol. 11 (3): 426-440

Rocha-Olivares, A., Rosenblatt, R. H., and R. D. Vetter. 1999b. Molecular evolution, systematics, and zoogeography of the rockfish subgenus *Sebastomus* (*Sebastes*, Scorpaenidae) based on mitochondrial cytochrome b and control region sequences. Mol. Phylogen. Evol. 11 (3): 441-458

Rocha-Olivares, A., Rosenblatt, R. H., and R. D.Vetter. 1999c. Cryptic species of rockfishes (*Sebastes*: Scorpaenidae) in the Southern Hemisphere inferred from mitochondrial lineages. J. Hered. 90: 404-411

Rocha-Olivares, A., Moser, H. G., and J. Stannard. 1999d. Molecular identification and description of pelagic young of the rockfishes *Sebastes constellatus* and *Sebastes ensifer*. Fish. Bull. 98: 353-363

Rocha-Olivares, A., Leal-Navarro, R. A., Kimbrell, C., Lynn, E. A., and R. D. Vetter. 2003. Microsatellite variation in the Mexican rockfish *Sebastes macdonaldi*. Sci. Mar. 67 (4): 451-460

Roques, S., Duchesne, P., and L. Bernatchez. 1999a. Potential of microsatellites for individual assignment: the North Atlantic redfish (genus *Sebastes*) species complex as a case study. Mol. Ecol. 8 (10): 1703-1717

Roques, S., Pallotta, D., Sévigny, J.-M., and L. Bernatchez. 1999b. Isolation and characterization of polymorphic microsatellite markers in the North Atlantic redfish (Theleostei: Scorpaenidae, genus *Sebastes*). Mol. Ecol. 8 (4): 685-702

Roques, S., Sévigny, J.-M., Bernatchez, L., and D. Power. 2000. Redfish species distribution and population genetic structure in the Northwest Atlantic: Preliminary results. NAFO Sci. Counc. Res. Doc. 00/48: 18 pp.

Roques, S., Sévigny, J.-M., and L. Bernatchez. 2001. Evidence for broad-scale introgressive hybridisation between two redfish (genus *Sebastes*) in the North-west Atlantic: a rare marine example. Mol. Ecol. 10 (1): 149-165

Roques, S., Sévigny, J.-M., and L. Bernatchez. 2002. Genetic structure of deep-water redfish, *Sebastes mentella*, populations across the North Atlantic. Mar. Biol. 140: 297-307

Rousset, F. 1996. Equilibrium values of measures of population subdivision for stepwise mutation processes. Genetics 142: 1357-1362

Rousset, F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145: 1219-1228

Rubec, P. J., McGlade, J. M., Trottier, B. L., and A. Ferron. 1991. Evaluation for methods for separation of Gulf of Saint-Lawrence beaked redfishes, *Sebastes fasciatus* and *S. mentella*: malate dehydrogenase mobility patterns compared with extrinsic gasbladder muscle passages and anal fin ray counts. Can. J. Fish. Aquat. Sci. 48: 640-660

Ruzzante, D. E., Taggart, C. T., Cook, D., and S. Goddard. 1996. Genetic differentiation between inshore and offshore Atlantic cod (*Gadus morhua*) off Newfoundland: microsatellite DNA variation and antifreeze level. Can. J. Fish. Aquat. Sci. 53: 634-645

Ruzzante, D. E., Taggart, C. T., and D. Cook. 1999. A review of the evidence for genetic structure of cod (*Gadus morhua*) populations in the Northwest Atlantic and population affinities of larval cod off Newfoundland and the Gulf of St. Lawrence. Fish. Res. 43: 79-97

Ruzzante, D. E., Wroblewski, J. S., Taggart, C. T., Smedbol, R. K., Cook, D., and S. V. Goddard. 2000. Bayscale population structure in coastal Atlantic cod in Labrador and Newfoundland, Canada. J. Fish Biol. 56: 431-447

Saborido-Rey, F. 1993. Distribution, abundance and biomass trends of the genus *Sebastes* on Flemish Cap (Division 3M). NAFO Sci. Counc. Res. Doc. 93/24:12 pp.

Saborido-Rey, F. 1994. The Genus *Sebastes* Cuvier, 1829 (Pisces, Scorpaenidae) in the North Atlantic: Species and population identification using morphometric techniques; Growth and reproduction of the Flemish Cap populations (In spanish). Ph.D. Thesis, Univ. Autónoma Madrid, Madrid, Spain: 276 pp.

Saborido-Rey, F., Garabana, L., and C. Stransky. 2001. A review of the population structure and ecology of redfish in the Irminger Sea and adjacent waters. NAFO Sci. Counc. Res. Doc. 01/107:20 pp.

Saborido-Rey, F., Garabana, L., Stransky, C., Melnikov, S., and V. Shibanov. 2005. Review of the population structure and ecology of *S. mentella* in the Irminger Sea and adjacent waters. Rev. Fish Biol. Fish. (*in press*)

Saiki, R. K., Gelfand, D. H., Stoffel, S., Schard, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and H. A. Erlich. 1988. Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487-491

Saitou, N., and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406-425

Sanger, F., Nicklen, S., and A. F. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467

Savelkoul, P. H. M., Aarts, H. J. M., de Haas, J., Dijkshoorn, L., Duim, B., Otsen, M., Rademaker, J. L. W., Schouls, L., and J. A. Lenstra. 1999. Amplified-Fragment Length Polymorphism Analysis: the state of an art. J. Clin. Microbiol. 37(10): 3083-3091

Schlötterer, C., and D. Tautz. 1992. Slippage synthesis of simple sequence DNA. Nucleic Acids Res. 20: 211-215

Schlötterer, C. 2000. Evolutionary dynamics of microsatellite DNA. Chromosoma 109: 365-371

Schmidt, C. 2000. Populationsgenetische Untersuchungen am Ostsee-Dorsch. Diploma thesis, University of Kiel, Germany: 90 pp.

Schneider, S., Roessli, D., and L. Excoffier. 2000. Arlequin. A Software for Population Genetics Data Analysis, Version 2.000. University of Geneva, Geneva

Seeb, L. W. 1998. Gene flow and introgression within and among three species of rockfishes, *Sebastes auriculatus, S. caurinus*, and *S. maliger*. J. Hered. 89: 393-403

Seki, S., Agresti, J. J., Gall, G. A. E., Taniguchi, N., and B. May. 1999. AFLP analysis of genetic diversity in three populations of Ayu *Plecoglossus altivelis*. Fish. Sci. 65 (6): 888-892

Sévigny, J.-M., De Gagné, P., Lafontaine, Y., and J. Dodson. 2000. Identification and distribution of the larvae of redfish species (*Sebastes fasciatus* and *S. mentella*: Scorpaenidae) in the Gulf of St. Lawrence. Fish. Bull. 98: 375–388

Shaklee, J. B., and K. P. Currens. 2003. Genetic stock identification and risk assessment. Pp. 291-328 *in* Hallerman, E. M., ed. Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, Maryland: 458 pp.

Sharma, S. K., Knox, M. R., and T. H. N. Ellis. 1996. AFLP analysis of the diversity and phylogeny of Lens and its comparison with RAPD analysis. Theor. Appl. Genet. 93: 751-758

Shaw, P. W., Pierce, G. J., and P. R. Boyle. 1999. Subtle population structuring within a highly vagile marine invertebrate, the veined squid *Loligo forbesi*, demonstrated with microsatellite DNA markers. Mol. Ecol. 8 (3): 407-417

She, J.-X., Autem, M., Kotoulas, G., Pasteur, N., and F. Bonhomme. 1987. Multivariate analysis of genetic exchanges between *Solea aegyptiaca* and *Solea senegalensis* (Teleosts, Soleidae). Biol. J. Linn. Soc. 32: 357-371

Shields, G. F., and J. R. Gust. 1995. Lack of geographic structure in mitochondrial DNA sequences of Bering Sea walleye pollock, *Theragra chalcogramma*. Mol. Mar. Biol. Biotechnol. 4 (1): 69-82

Shinomiya, A., and O. Ezaki. 1991. Mating habits of the rockfish *Sebastes inermis*. Environ. Biol. Fish. 30 (1–2): 15-22

Slatkin, M., and R. R. Hudson. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129: 555-562

Slatkin, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47: 264-279

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139: 457-462

Smith, P. J., Jamieson, A., and A. J. Birley. 1990. Electrophoretic studies and the stock concept in marine teleosts. J. Cons. Int. Explor. Mer 47: 231-245

Smith, Wm. L., and W.C. Wheeler. 2004. Phylogeny of the mail-cheeked fishes (Teleostei: Scorpaeniformes): evidence from mitochondrial and nuclear sequence data. Mol. Phylogen. Evol. 32: 627-646

Sota, T. 2002. Radiation and reticulation: extensive introgressive hybridisation in the carabid beetles *Ohomopterus* inferred from mitochondrial gene genealogy. Popul. Ecol. 44: 145-156

StatSoft, Inc. 2003. STATISTICA (data analysis software system), version 6. www.statsoft.com

Stepien, C. A., and T. D. Kocher. 1997. Molecules and morphology in studies of fish evolution. Pp. 1-11 *in* Kocher, T. D., and C. A. Stepien, eds. Molecular systematics of fishes. Academic Press, San Diego, CA: 314 pp.

Storer, D. H. 1854. A new species of fish from Provincetown, found in the harbor at that place. Proc. Boston Soc. Nat. Hist. 31

Stransky, C. 2000. Migration of juvenile deep-sea redfish (*Sebastes mentella* Travin) from the East Greenland shelf into the central Irminger Sea. ICES C.M. 2000/N:28: 10 pp.

Stransky, C. 2005. Geographic variation of golden redfish (*Sebastes marinus*) and deep-sea redfish (*S. mentella*) in the North Atlantic based on otolith shape analysis. ICES J. Mar. Sci.: submitted

Stransky, C., Garbe-Schönberg, C.-D., and D. Günther. 2005. Microchemistry of Atlantic redfish otoliths: Temporal stability, geographic variation and migration. Mar. Freshw. Res.: submitted

Stransky, C., and S. E. MacLellan. 2005. Species separation and zoogeography of redfish and rockfish (genus *Sebastes*) by otolith shape analysis. Can. J. Fish. Aquat. Sci.: submitted

Strimmer, K., and A. von Haesseler. 1996. Quartet Puzzling: A quartet maximum-likelihood method for reconstructing tree topologies. Mol. Biol. Evol. 13: 964-969

Sundt, R. C., and T. Johansen. 1998. Low levels of interspecific DNA sequence variation of the mitochondrial 16S rRNA in North Atlantic redfish *Sebastes* (Pisces, Scorpaenidae). Sarsia 83: 449-452

Swofford, D. L. 1993. Phylogenetic Analysis Using Parsimony (PAUP), Version 3.1.1. Illinois Natural History Survey, Champaign, IL, USA

Swofford, D. L., Olsen, G. J., Waddell, P. J., and D. M. Hillis. 1996. Phylogenetic Inference. Pp. 407-514 *in* Hillis, D. M., and C. Moritz, eds. Molecular systematics, 2nd edition. Sinauer Associates, Sunderland, MA, USA: 655pp.

Swofford, D. L. 2002. PAUP*: PAUP*, Phylogenetic Analysis Using Parsimony (*and other Methods), version 4.0b10. Sinauer Associates, Sunderland, MA, USA

Takezaki, N., and M. Nei. 1996. Genetic distances and re-construction of phylogenetic trees from microsatellite DNA. Genetics 144: 389-399

Tamburrini, M., Verde, C., Olianas, A., Giardina, B., Corda, M., Sanna, M. T., Fais, A., Deiana, A. M., Di Prisco, G., and M. Pellegrini. 2001. The hemoglobin system of the brown moray *Gymnothorax unicolor* - structure/function relationships. Eur. J. Biochem. 268: 4104 -4111

Tautz, D. 1989. Hypervariability of simple sequences as a source for polymorphic DNA markers. Nucleic Acids Res. 17: 6463-6471

Tavolga, W. N. 1971. Sound production and detection. Pp. 135-205 *in* Hoar, W. S., and D. J. Randall, eds. Fish physiology Vol. V. Academic press, New York

Templeman, W., and E. J. Sandeman. 1959. Variations in caudal pigmentation in late-stage pre-extrusion larvae from *marinus*- and *mentella*-type female redfish from the Newfoundland area. J. Fish. Res. Board Can. 16: 763-789

Templeman, W. 1976. Biological and oceanographic background of Flemish Cap as an area for research on the reasons for year-class success and failure in cod and redfish. ICNAF Res. Bull. 12: 91-117

Templeman, W. 1980. Incidence subcaudal melanophores in preextrusion larvae of redfish species in the Newfoundland-Labrador area. J. Northw. Atl. Fish. Sci., 1: 7-19

Templeton, A. R., Crandall, K. A., and C. F. Sing. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132: 619-633

Thomas, W. K., and A. T. Beckenbach. 1989. Variation in salmonid mitochondrial DNA: Evolutionary constraints and mechanisms of substitution. J. Mol. Evol. 29: 233-245

Trautner, J. 2000. Genetische Untersuchungen an Wild- und Zuchtpopulationen der Regenbogenforelle, *Oncorhynchus mykiss* (Pisces). Doctoral Thesis, University of Hamburg, Germany: 95 pp.

Travin, V. I. 1951. *Sebastes mentella* Travin sp. nov. in Barentz Sea. C. R. (Doklady) Acad. Sci. URSS (N. S.): 741-744

Trottier, B. L., Rubec, P. J., and A. C. Ricard. 1989. Biochemical separation of Atlantic Canadian redfish: *Sebastes mentella* and *Sebastes norvegicus*. Can. J. Zool. 67: 1332-1335

Turner, T. F., Wares, J. P., and J. R. Gold. 2002. Genetic effective size is three orders of magnitude smaller than adult census size in an abundant, estuarine-dependent marine fish. Genetics 162: 1329-1339

van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., and P. Shipley. 2004. *micro-checker*: software for identifying and correcting genotyping errors in microsatellite data. Mol. Ecol. Notes 4: 535-538

Vekemans, X. 2002. AFLP-SURV version 1.0. Distributed by the author. Laboratoire de Génétique et Ecologie Végétale, Université Libre de Bruxelles, Belgium

Verspoor, E., and J. Hammar. 1991. Introgressive hybridisation in fishes: the biochemical evidence. J. Fish Biol. (Suppl. A) 39: 309-334

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23 (21): 4407-4414

Vos, P., and M. Kuiper. 1997. AFLP analysis. Pp. 115-131 *in* Caetano-Anollés, G., and P. M. Gresshoff, eds. DNA markers: protocols, applications, and overviews. Wiley-Liss Inc., New York: 364 pp.

Waples, R. S. 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. J. Hered. 89: 438-450

Ward, R. D., Woodward, M., and D. O. F. Skibinski. 1994. A comparison of genetic diversity levels in marine, freshwater and anadromous fishes. J. Fish Biol. 44: 213-232

Ward, R. D., and P. M. Grewe. 1995. Appraisal of molecular genetic techniques in fisheries. Pp. 29-54 *in* Carvalho, G. R., and T. J. Pitcher, eds. Molecular genetics in fisheries. Chapman & Hall, London: 141 pp.

Ward, R. D. 2000. Genetics in fisheries management. Hydrobiologia 420: 191-201

Ward, R. D. 2002. Genetics of fish populations. Pp. 200-225 *in* Hart, P. J. B., and J. D. Reynolds, eds. Handbook of fish biology and fisheries. Volume 1: Fish biology. Blackwell Science Ltd., Oxford, UK: 424 pp.

Watterson, G. A., and H. A. Guess. 1977. Is the most frequent allele the oldest? Theor. Popul. Biol. 11: 141-160

Wattier, R., Engel C. R., Saumitou-Laprade, P., and M. Valero. 1998. Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). Mol. Ecol. 7 (11): 1569-1573

Weber, J. L., and P. E. May. 1989. Abundant class of human DNA polymorphisms can be typed using the polymerase chain reaction. Am. J. Hum. Gen. 44: 388-396

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38 (6): 1358-1370

Weir, B. S. 1996. Genetic data analysis II. Sinauer Associates, Inc., Sunderland, MA: 445 pp.

Whitehead, P. J. P., Bauchot, M.-L., Hureau, J.-C., Nielsen, J., and E. Tortonese, eds. 1986. Fishes of the north-eastern Atlantic and the Mediterranean. Vol. III. UNESCO, Paris: Pp.1223-1227

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A., and S. V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl. Acids Res. 18: 6531-6535

Wilson, A. J., Gíslason, D., Skúlason, S., Snorrason, S. S., Adams, C. E., Alexander, G., Danzmann, R. G., and M. M. Ferguson. 2004. Population genetic structure of Arctic Charr, *Salvelinus alpinus* from northwest Europe on large and small spatial scales. Mol. Ecol. 13 (5): 1129-1142

Withler, R. E., Beacham, T. D., Schulze, A. D., Richards, L. J., and K. M. Miller. 2001. Co-existing populations of Pacific Ocean perch, *Sebastes alutus*, in Queen Charlotte Sound, British Columbia. Mar. Biol. 139: 1-12

Wourms, J. P. 1991. Reproduction and development of *Sebastes* in the context of the evolution of piscine viviparity. Environ. Biol. Fish. 30: 111-126

Wright, S. 1943. Isolation by distance. Genetics 28: 114-138

Wright, S. 1951. The genetical structure of populations. Ann. Eugen. 15: 323-354

Wright, S. 1965. The interpretation of population structure by F-statistics with special regards to systems of mating. Evolution 19: 395-420

Yan, G., Romero-Severson, J., Walton, M., Chadee, D. D., and D. W. Severson. 1999. Population genetics of the yellowfever mosquito in Trinidad: comparisons of amplified fragment length polymorphisms (AFLP) and restriction fragment length polymorphism (RFLP) markers. Mol. Ecol. 8 (6): 951–963

Yanulov, K. P. 1962a. Parasites as indicators of local rosefish stocks. Pp. 266-276 *in* Mart, Y. Y., ed. Soviet fisheries investigations in the Northwest Atlantic. English translation by Hershkowitz, B., Israel Program for Scientific translations, Jerusalem, 1963

Yanulov, K. P. 1962b. On the group of rosefish (*S. mentella* Travin) in the Labrador-Newfoundland area. Pp. 277-289 *in* Mart, Y. Y., ed. Soviet fisheries investigations in the Northwest Atlantic. English translation by Hershkowitz, B., Israel Program for Scientific translations, Jerusalem, 1963

Young, F. W. 1996. Vista: The Visual Statistics System. Research Memorandum 94-1(b). L. L. Thurstone Psychometric Laboratory, University of North Carolina, Chapel Hill, NC

Young, W. P., Ostberg, C. O., Keim, P., and G. H. Thorgaard. 2001. Genetic characterization of hybridisation and introgression between anadromous rainbow trout (*Oncorhynchus mykiss irideus*) and coastal cutthroat trout (*O. clarki clarki*). Mol. Ecol. 10 (4): 921-930

Zabeau, M., and P. Vos 1993. Selective restriction fragment amplification. A general method for DNA fingerprinting. European Patent Application 92402629.7 (publication number EP0534858A1)

Zhivotovsky, L. A. 1999. Estimating population structure in diploids with multilocus dominant DNA markers. Mol. Ecol. 8 (6): 907-913

ABBREVIATIONS

A	adenine
AFC	anal fin ray count
AFLP	amplified fragment length polymorphism
AMOVA	analysis of molecular variance
Arg	arginine
ATP	adenosine triphosphate
bp	base pair(s)
C	cytosine
°C	degree Celsius
CI	consistency index
conc.	concentration
D_{CE}	Cavalli-Sforza and Edwards (1967) chord distance
ddNTP	2',3'-dideoxynucleoside 5'-triphosphate
d.f.	degrees of freedom
dH_2O	deionised water
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	2'-deoxynucleoside 5'-triphosphate
DTT	1,4-dithiothreitol
E	east
E. coli	*Escherichia coli*
*Eco*RI	restriction enzyme from *Escherichia coli* strain RY 13
EDTA	ethylenediaminetetraacetic acid
e.g.	exempli gratia
EGM	extrinsic gas bladder muscle
etc.	et cetera
F'	forward
F81	model of Felsenstein (1981)
FCA	factorial correspondence analysis
F_{ST}	fixation index (Wright 1951; Weir and Cockerham 1984)
G	guanine
Gly	glycine
h	hour
$\hat{h}$	haplotype diversity
Hb	hemoglobin
H_e	expected heterozygosity, gene diversity

HI	homoplasy index
*Hind*III	restriction enzyme from *Haemophilus influenzae*
H_j	Nei's (1987) gene diversity (analogous to the average gene diversity H_e)
HKY85	Hasegawa-Kishino-Yano model (Hasegawa et al. 1985)
H_o	observed heterozygosity
H strand	heavy strand
HW	Hardy-Weinberg
HWE	Hardy-Weinberg equilibrium
HWP	Hardy-Weinberg principle
IAM	infinite allele model
IBD	isolation by distance
ICES	International Council for the Exploration of the Seas
IDHP	isocitrate dehydrogenase
IPTG	isopropyl ß-D-thiogalactopyranoside
IRD	infrared fluorescence dye
k	number of alleles
K	number of genetic clusters
KAc	potassium acetate
kb	kilo base pairs
K2P	Kimura 2-parameter model (Kimura 1980)
l	litre
*lac*Z	beta-galactosidase gene
LB	Luria-Bertani
LD	linkage disequilibrium
LE	linkage equilibrium
L strand	light strand
M13	bacteriophage M13
MCMC	Markov chain Monte Carlo
MCS	multiple cloning site
MDH	malate dehydrogenase
MDSA	multidimensional scaling analysis
MEP	malic enzyme
MgAc	magnesium acetate
min	minute
ML	maximum likelihood
MP	maximum parsimony
*Mse*I	restriction enzyme from *Micrococcus spp.*
mtDNA	mitochondrial DNA
n	number of individuals

NADH	nicotinamide adenine dinucleotide
NAFO	Northwest Atlantic Fisheries Organization
ND3	NADH dehydrogenase subunit 3
N_e	effective population size
NE	northeast
NJ	neighbour-joining
no.	number
O	upper strand
P	probability value
PCR	polymerase chain reaction
Q	membership coefficient
R^2	coefficient of determination
r	standardised Mantel test relationship coefficient
R'	reverse
r(50)	allelic richness
RAPD	randomly amplified polymorphic DNA
RC	rescaled consistency index
RFLP	restriction fragment length polymorphism
RI	retention index
RL	restriction-ligation
RNA	ribonucleic acid
Rnase	ribonuclease
rpm	rotations per minute
rRNA	ribosomal RNA
R_{ST}	microsatellite specific analogue of F_{ST} (Slatkin 1995)
S	Svedberg unit
SDS	sodium dodecyl sulphate
S.E.	standard error
SE	southeast
SMM	stepwise mutation model
sp.	species
SPN	statistical parsimony network
SW	southwest
T	thymine
T4	bacteriophage T4
Taq	*Thermus aquaticus*
TBE	Tris-borate-EDTA buffer
TBR	tree bisection and reconnection

TE	Tris-EDTA buffer
ti	transition
ti/tv	transition/transversion ratio
TPM	two phase model
Tris	tris(hydroxymethyl)aminoethane
Tris-HCl	HCl-buffered Tris
tRNA	transfer RNA
$tRNA^{Arg}$	transfer RNA for arginine
$tRNA^{Gly}$	transfer RNA for glycine
tv	transversion
u	Weiss unit
U	lower strand
UPGMA	unweighted pair group method with arithmetic means
UV	ultraviolet
V	volt
VNTR	variable number of tandem repeats
W	west
X	genotypes of the sampled individuals
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galactoside
Z	unknown clusters or populations of origin
$\hat{\pi}$	nucleotide diversity (Nei 1987)
Φ_{ST}	fixation index (Excoffier et al. 1992)

Other units of measurement not listed above were abbreviated according to the International Unit System SI (Système Internationale d`Unite).

ACKNOWLEDGEMENTS

I would like to express my deepest thanks to all friends and colleagues who supported me during the course of this study.
I want to thank Professor Dr. Axel Temming for being my supervisor and for supporting my work with his advice.
For being my co-supervisor, for many helpful comments and for his advice on phylogenetic analysis, I would like to thank Professor Dr. Reinhold Hanel.
Thanks to all the great colleagues from the Federal Research Centre for Fisheries (BFAFi) in Hamburg.
My sincere and warm thanks go to Dr. Jochen Trautner - for his friendship, advice, encouragement and support throughout all stages of my doctoral thesis.
Thanks also to Dr. Hans-Stephan Jenke for the opportunity to work at the laboratory of the Institute for Fisheries Ecology.
My warmest thanks and hugs go to Corena Wesche and Antje Bühringer for their skilful technical assistance and for their friendship - and for all the enjoyable coffee-breaks.
Special thanks and hugs also to Katrin Roth, not only for the great time we had during her stay at our lab and her constructive comments on my thesis, but also for her dear friendship.
For being a great project-coordinator and for always supporting my work and my decisions, I want to thank Dr. Hans-Joachim Rätz.
I also want to thank Dr. Einar Eg Nielsen (Danish Institute for Fisheries Research), for being always available when I needed advice on population genetics and microsatellite statistics.
Special thanks also to Dr. Sönke Jansen und Dr. Kristin Kaschner, who provided very helpful comments and corrections in the final stage of the thesis.
Thanks also to Jörg Appel for his advice on creating maps.

This work was partly funded by the European Commission within the research and development project REDFISH (Population structure, reproductive strategies and demography of redfish (Genus *Sebastes*) in the Irminger Sea and adjacent waters (ICES V, XII and XIV; NAFO 1), QLK5-CT1999-01222.
Thanks to all the nice people who participated in this project - for the great time, the fruitful redfish discussions and the fantastic and interesting project meetings.
For providing redfish samples, sample data and biological data, I would like to express my gratitude to Dr. Christoph Stransky (Federal Research Centre for Fisheries, Hamburg, Germany), Dolores (Lola) Garabana and Dr. Francisco (Fran) Saborido-Rey (Institute of Marine Research, Vigo, Spain), Dr. Anna-Kristín Daníelsdóttir, Þorsteinn Sigurðsson and their staff (Marine Research Institute, Reykjavík, Iceland), Dr. Torild Johansen (University of Bergen, Department of Fisheries and Marine Biology, Bergen, Norway), the people from the

Faroes Fisheries Laboratory (Tórshavn, Faroe Islands), Dr. James Glass (Tristan Natural Resources Department, Edinburgh, Tristan da Cunha) and Dr. Margaret Treble (Fisheries and Oceans Canada, Prince Rupert, British Columbia).
Thanks to Lola and Torild for providing me with information about the results of the morphometrical and protein analyses, respectively, of the specimens analysed in this thesis.
For their interest in my work and many interesting conversations, I would particularly like to thank Dr. Kjell Nedreaas, Prof. Dr. Gunnar Nævdal, Dr. Hartmut Rehbein, Torild, Lola, Fran, Chris and all the other participants of REDFISH.

There are no words to describe my gratitude to my love Markus and to my family, my parents and my grandmother, for their patience, continuous support and loving care they have given me during these years. This work is dedicated to them and to my late grandfather (who loved fish and fishing) and my paternal grandparents. I also want to thank my "family-in-law", for their encouragement, help and care in so many ways.