

Entwicklung molekularbiologischer Methoden für die Authentizitätsbestimmung von Haselnüssen

Dissertation

Zur Erlangung des akademischen Grades

Doctor rerum naturalium

Dr. rer. nat.

Aus dem Fachbereich Chemie der Universität Hamburg

Hamburg School of Food Science

vorgelegt von

Christina Lang

Hamburg

22.07.2023

1. Gutachter der Dissertation: Prof. Dr. Markus Fischer

2. Gutachter der Dissertation: Prof. Dr. Stephan Seifert

Prüfungskommission: Prof. Dr. Markus Fischer, Prof. Dr. Ralph Holl, Dr. Maria Riedner

Tag der Disputation: 10.11.2023

Tag der Druckfreigabe: 10.11.2023

Die vorliegende Arbeit wurde im Zeitraum von August 2016 bis Juli 2019 an der Hamburg School of Food Science des Fachbereichs Chemie der Fakultät für Mathematik, Informatik und Naturwissenschaften der Universität Hamburg unter der Anleitung von Herrn Prof. Dr. Markus Fischer angefertigt.

*»Um glücklich zu sein,
muss man seine Vorurteile abgelegt
und seine Illusionen behalten haben.«*

Émilie du Châtelet

Publikationsliste

Lang C., Felbinger C., Fischer M.: Sortendifferenzierung von Haselnüssen mittels molekularbiologischer Verfahren. *Deutsche Lebensmittel-Rundschau* (2019), 115, 9, S. 386-393.

Lang C., Weber N., Möller M., Schramm L., Schelm S., Kohlbacher O., Fischer M.: Genetic authentication: Differentiation of hazelnut cultivars using polymorphic sites of the chloroplast genome. *Food Control* (2021), 130, 108344.

Lang C., Scharf A., Fischer M.: Lebensmittelauthentifizierung – Anwendung der CRISPR-Cas9-Technologie zur Differenzierung von Edel- und Konsumkakao. *Deutsche Lebensmittel-Rundschau* (2021), 117, 5, S. 194-199.

Scharf A.*, **Lang C.***, Fischer M.: Genetic authentication: Differentiation of fine and bulk cocoa (*Theobroma cacao* L.) by a new CRISPR/Cas9 based *in vitro* method. *Food Control* (2020), 114, 107219.

* corresponding first authors

Schelm S., Siemt M., Pfeiffer J., **Lang C.**, Tichy H.-V., Fischer M.: Food Authentication: Identification and Quantitation of Different Tuber Species via Capillary Gel Electrophoresis and Real-Time PCR. *Foods* (2020), 9(4), 501.

Bachmann R., Jilani A., Ibrahim H., Bahmann D., **Lang C.**, Fischer M., Bisping B., Hackl T.: Signal pattern plot: a simple tool for time-dependent metabolomics studies by 1H NMR spectroscopy. *Analytical und Bioanalytical Chemistry* (2019), 411, 6857.

Inhaltsverzeichnis

Publikationsliste.....	IV
Inhaltsverzeichnis.....	V
Abkürzungsverzeichnis	IX
Maßeinheiten und Größen	XII
1 Zusammenfassung	1
2 Abstract.....	3
3 Einleitung	5
4 Grundlagen.....	6
4.1 Haselnuss	6
4.1.1 Botanische Einordnung	6
4.1.2 Bedeutung unterschiedlicher Haselnussorten und deren Anbau.....	7
4.1.3 Wirtschaftliche Bedeutung und Nutzung.....	9
4.1.4 <i>Food Fraud</i> im Bereich der Haselnusserzeugnisse	10
4.2 DNA	11
4.2.1 <i>Genomics</i> –basierte Ansätze in der Lebensmittelüberwachung	11
4.2.2 Nukleare DNA.....	12
4.2.2.1 ITS-Bereich.....	13
4.2.2.2 WRKY-Gene	15
4.2.3 Plastidäre DNA	16
4.3 Molekularbiologische Methoden.....	18
4.3.1 Polymerase chain reaction (PCR)	18
4.3.2 Detektion von PCR-Produkten	20
4.3.2.1 Agarosegelelektrophorese (AGE)	20
4.3.2.2 Kapillargelelektrophorese (CGE).....	21
4.3.3 Sequenzierverfahren	23
4.3.3.1 Sanger-Sequenzierung.....	23
4.3.3.2 <i>Next-Generation Sequencing</i>	24
4.3.4 Klonierung	28
4.3.5 Analyse von <i>Start Codon Targeted</i> (SCoT) Polymorphismen	30
4.3.6 SCAR-Marker	32
4.3.7 <i>PCR-Restriction fragment length polymorphism</i> (PCR-RFLP)	33
4.3.8 <i>High resolution melting analysis</i> (HRMA).....	35
5 Zielsetzung	37
6 Ergebnisse und Diskussion.....	38
6.1 Sequenzierung des cpGenoms mittels NGS	39
6.2 Identifizierung und Validierung von Sequenzunterschieden	40

6.3	Entwicklung und Anwendung von <i>targeted</i>-Methoden auf Basis polymorpher Bereiche des cpGenoms	42
6.3.1	Nachweis der Indels 1 und 3 in den Bereichen B und C des cpGenoms	42
6.3.1.1	Qualitativer Nachweis	42
6.3.1.2	Nachweis der Indels 1 und 3 in prozessierten Haselnüssen	48
6.3.1.3	Quantifizierung von Haselnussmischungen	48
6.3.2	PCR-RFLP für den Nachweis georgischer Haselnüsse	58
6.3.2.1	Qualitativer Nachweis	58
6.3.2.2	Nachweis in prozessierten Haselnüssen	61
6.3.2.3	Quantifizierung des Anteils georgischer Haselnüsse in Mischungen	62
6.3.3	HRMA für die Differenzierung von Haselnüssen unterschiedlichen geographischen Ursprungs	65
6.3.3.1	Ausgangssituation für die Methodenentwicklung	65
6.3.3.2	Berechnung von Schmelzkurven	66
6.3.3.3	Methodenentwicklung und -validierung	68
6.3.3.4	Überprüfung der Anwendbarkeit auf prozessierte Proben	73
6.3.3.5	Quantifizierung mittels HRMA	74
6.4	Analyse definierter Bereiche des Karyoms	79
6.4.1	Analyse der ITS-Bereiche	79
6.4.2	Analyse der WRKY-Gene	81
6.5	Analyse von SCoT Polymorphismen	87
6.5.1	Methodenentwicklung und -optimierung	87
6.5.2	Auswertung mittels <i>Neighbour-Joining</i> -Methode	88
6.5.3	Auswertung mittels Hauptkomponentenanalyse	91
6.5.4	Zusammenfassung und Diskussion der Analyse von SCoT-Polymorphismen	93
6.6	Entwicklung sortenspezifischer SCAR-Marker	95
6.7	Zusammenfassung der Analysenergebnisse	98
7	Material und Methoden	105
7.1	Probenmaterial	105
7.1.1	Haselnusskerne (Rohware)	105
7.1.2	Prozessierte Haselnussproben	105
7.1.3	Haselnussblätter	106
7.1.4	Kennzeichnung des Probenmaterials	106
7.2	Gesamt-DNA-Isolierung	106
7.2.1	Isolierung der Gesamt-DNA aus Haselnussblättern	107
7.2.2	Isolierung der Gesamt-DNA aus Haselnusskernen	107
7.2.2.1	CTAB-Methode	107
7.2.2.2	<i>Rapid Extraction</i>	108
7.3	Qualitätsbestimmung von DNA-Isolaten	108
7.3.1	Fluorimetrische Konzentrationsbestimmung	108
7.3.2	Photometrische Reinheitsbestimmung	109

7.4	Qualitative Polymerasekettenreaktion (PCR)	109
7.5	Detektion von PCR-Produkten mittels Elektrophorese	110
7.5.1	Agarosegelelektrophorese (AGE)	110
7.5.2	Kapillargelelektrophorese (CGE)	111
7.6	Aufreinigung von PCR-Produkten	112
7.6.1	<i>Monarch® PCR & DNA Cleanup Kit</i>	112
7.6.2	<i>peqGOLD Gel Extraction Kit</i>	112
7.7	Klonierung	112
7.7.1	Herstellung elektrokompetenter <i>E. coli</i> -Zellen	112
7.7.2	Herstellung und Aufreinigung des Ligationsansatzes.....	113
7.7.3	Elektroporation und Transformation	113
7.7.4	<i>Picking & Screening</i>	114
7.7.5	Plasmidextraktion.....	114
7.7.6	Finale Überprüfung und Sequenzierung	114
7.8	Erstellung von <i>Alignments</i>	115
7.9	Sequenzierung des cpGenoms mittels NGS	115
7.10	Resequenzierung vier ausgewählter Bereiche des cpGenoms	116
7.11	Entwicklung von <i>targeted</i>-Methoden auf Basis polymorpher Bereiche des cpGenoms ..	117
7.11.1	Nachweis der Indels 1 und 3 in den Bereichen B und C.....	117
7.11.1.1	Qualitativer Nachweis der Indels 1 und 3.....	117
7.11.1.2	Quantifizierung von Haselnussmischungen auf Basis der identifizierten Indels 1 und 3	117
7.11.2	PCR-RFLP für den Nachweis georgischer Haselnüsse auf Basis eines SNPs in Bereich D.....	119
7.11.2.1	Qualitativer Nachweis georgischer Haselnüsse mittels PCR-RFLP.....	119
7.11.2.2	PCR-RFLP für den semi-quantitativen Nachweis georgischer Haselnüsse.....	119
7.11.3	HRMA für die Differenzierung georgischer und aserbajdschanischer Haselnüsse	120
7.11.3.1	Berechnung von Schmelzkurven	120
7.11.3.2	Methodenentwicklung und -validierung	120
7.11.3.3	Quantitative Bestimmung von Haselnussmischungen	121
7.12	Analyse definierter Bereiche des Karyoms	121
7.12.1	Analyse der ITS-Bereiche	121
7.12.2	Analyse der WRKY-Gene	122
7.13	Analyse von SCoT Polymorphismen	123
7.13.1	Probenmaterial	123
7.13.2	Durchführung der PCR	123
7.13.3	Auswertung.....	123
7.14	Entwicklung sortenspezifischer SCAR-Marker	123
7.14.1	RAPD-PCR.....	124
7.14.2	Klonierung.....	124
7.14.3	Entwicklung der SCAR-Primer	124

8	Literaturverzeichnis	125
9	Anhang.....	141
9.1	Gefahrstoffverzeichnis.....	141
9.2	Verwendete Verbrauchsmaterialien.....	143
9.3	Verwendete Chemikalien, Puffer, Kits und Primer	144
9.4	Geräteparameter.....	150
9.5	Softwareparameter und Online-Tools	152
9.6	Probenmaterial	153
9.7	Kennzeichnung des Probenmaterials.....	163
9.8	Temperaturprogramme	165
9.9	Parameter für die statistische Bewertung von Kalibriergeraden	166
9.10	Ergänzungen zum Abschnitt 6 – Ergebnisse und Diskussion.....	169
9.10.1	Sequenzierung des cpGenoms mittels NGS.....	169
9.10.2	Identifizierung und Validierung von Sequenzunterschieden.....	170
9.10.3	Entwicklung und Anwendung von <i>targeted</i> -Methoden auf Basis polymorpher Bereiche des cpGenoms	179
9.10.4	Nachweis der Indels 1 und 3 in den Bereichen B und C des cpGenoms	179
9.10.4.1	Qualitativer Nachweis	179
9.10.4.2	Quantifizierung von Haselnussmischungen.....	181
9.10.5	PCR-RFLP für den Nachweis georgischer Haselnüsse	183
9.10.6	HRMA für die Differenzierung von Haselnüssen unterschiedlichen geographischen Ursprungs	185
9.10.6.1	Berechnung von Schmelzkurven	185
9.10.6.2	Methodenentwicklung und -validierung.....	186
9.10.6.3	Quantifizierung mittels HRMA.....	188
9.10.7	Analyse der ITS-Bereiche	190
9.10.8	Analyse der WRKY-Gene	192
9.10.9	Analyse von SCoT Polymorphismen.....	195
9.10.10	Entwicklung sortenspezifischer SCAR-Marker	208
10	Danksagung.....	210
11	Eidesstattliche Versicherung	211

Abkürzungsverzeichnis

Abkürzung	Erläuterung
&	und
<	kleiner
>	größer
®	<i>Registered mark</i>
A	Adenin
AGE	Agarosegelelektrophorese
Amp	Ampicillin
AZ	Aserbaidshon
bp	Basenpaare
BSA	Rinderserumalbumin (<i>Bovines Serumalbumin</i>)
bspw.	beispielsweise
bzgl.	bezüglich
C	Cytosin
CGE	Kapillargelelektrophorese (<i>capillary gel electrophoresis</i>)
cpDNA	Chloroplasten-DNA
CTAB	Cetyltrimethylammoniumbromid
DE	Deutschland
DNA	Desoxyribonukleinsäure (<i>Desoxyribonucleic acid</i>)
dNTP	Desoxyribonukleosidtriphosphat
dsDNA	doppelsträngige DNA (<i>double-stranded DNA</i>)
<i>E. coli</i>	<i>Escherichia coli</i>
EBI	Europäisches Institut für Bioinformatik (<i>European Bioinformatics Institute</i>)
EDTA	Ethylendiamintetraessigsäure
EMBL	Europäisches Laboratorium für Molekularbiologie (<i>European Molecular Biology Laboratory</i>)
ES	Spanien
EtOH	Ethanol
EU	Europäische Union / Europa
FAO	<i>Food and Agriculture Organization</i>
FAOSTAT	<i>Food and Agriculture Organization Corporate Statistical Database</i>
FR	Frankreich
FU	Fluoreszenz (<i>fluorescence units</i>)
FW	<i>forward</i>
G	Guanin
g.g.A.	Geschützte geographische Angabe
g.U.	Geschützte Ursprungsbezeichnung
GE	Georgien
GESTIS	Gefahrenstoffinformationssystem
GHS	<i>Globally Harmonized System</i>
Gr.	Gruppe
HRMA	Hochauflösende Schmelzpunktanalyse (<i>high resolution melting analysis</i>)
Inc.	<i>Incorporated</i>

Abkürzung	Erläuterung
Indel	Insertion oder Deletion
IT	Italien
ITS	<i>Internal transcribed spacer</i>
IUPAC	<i>International Union of Pure and Applied Chemistry</i>
LB	LB-Medium = Nährmedium für Bakterienkulturen (<i>lysogeny broth</i>)
LOD	Nachweisgrenze (<i>limit of detection</i>)
LOQ	Bestimmungsgrenze (<i>limit of quantitation</i>)
LSC-Region	<i>Large single copy region</i>
MgCl ₂	Magnesiumchlorid
mtDNA	Mitochondrien-DNA
NA	Unbekannt
NaCl	Natriumchlorid
NGS	<i>Next Generation Sequencing</i>
Nr.	Nummer
OD	Optische Dichte
PCA	Hauptkomponentenanalyse (<i>principle component analysis</i>)
PCR	Polymerasekettenreaktion (<i>polymerase chain reaction</i>)
<i>q.s.</i>	<i>Quantum satis</i>
qPCR	quantitative <i>real time</i> -PCR
r	Korrelationskoeffizient
R ²	Bestimmtheitsmaß
RAPD	<i>Randomly amplified polymorphic DNA</i>
rDNA	Ribosomale DNA
RFU	Relative Fluoreszenz (<i>relative fluorescence units</i>)
RNA	Ribonukleinsäure
rRNA	Ribosomale RNA
RT	Raumtemperatur
RV	<i>reverse</i>
SCAR	<i>Sequence characterized amplified region</i>
SCoT	<i>Start codon targeted</i>
SDS	Natriumdodecylsulfat
SNP	Einzelbasenaustausch (<i>single nucleotide polymorphism</i>)
SSC-Region	<i>Small single copy region</i>
ssDNA	einzelsträngige DNA (<i>single-stranded DNA</i>)
SSR	Mikrosatelliten (<i>simple sequence repeats</i>)
<i>s_{x0}</i>	Verfahrensstandardabweichung
<i>s_y</i>	Reststandardabweichung
T	Thymin
TAE	Tris-Acetat-EDTA
<i>Taq</i>	<i>Thermus aquaticus</i>
TE	Tris-EDTA
TM	<i>Trade mark</i>
T _m	Schmelzpunkt

Abkürzung	Erläuterung
TR	Türkei
TRIS	Trishydroxymethylaminomethan
TRIS-HCl	Trishydroxymethylaminomethanhydrochlorid
u.a.	unter anderem
USD	<i>United States Dollar</i>
v.a.	vor allem
Vgl.	Vergleich
v_{x0}	Verfahrensvariationskoeffizient
Wdh.	Wiederholung
x	Steigung der Kalibriergerade bei linearer Regression
y	Schnittpunkt der Kalibriergerade mit der y-Achse bei linearer Regression

Maßeinheiten und Größen

Abkürzung	Erläuterung
%	Prozent
°C	Grad Celsius
µg	Mikrogramm
µL	Mikroliter
µm	Mikrometer
µM	Mikromolar
c	Konzentration
cm	Zentimeter
g	Gramm
g; rcf	Relative Zentrifugalbeschleunigung
h	Stunde
I	Elektrische Stromstärke
kbp	Kilo-Basenpaare
kcal	Kilokalorie
L	Liter
mA	Milliampere
Mbp	Mega-Basenpaare
mg	Milligramm
min	Minute
mL	Milliliter
mm	Millimeter
mM	Millimolar
ng	Nanogramm
nm	Nanometer
nM	Nanomolar
P	Leistung
pH	<i>potentio Hydrogenii</i>
rpm	Umdrehung pro Minute (<i>revolutions per minute</i>)
s	Sekunde
S	Svedberg
t	Tonne
U	Elektrische Spannung
V	Volt
W	Watt

1 Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit der Entwicklung molekularbiologischer Methoden für die Authentizitätsbestimmung und Sortendifferenzierung von Haselnüssen (*Corylus avellana*). Ziel war zunächst die Identifizierung von Sequenzunterschieden. Anschließend sollten darauf basierend Analysenmethoden entwickelt werden, die sowohl den qualitativen Nachweis, als auch die Identifizierung und quantitative Abschätzung von Mischungen unterschiedlicher Proben ermöglichen.

Der Fokus lag dabei auf der Analyse des Chloroplastengenoms, da dieses im Gegensatz zum Kern- und Mitochondriengenom rein maternal vererbt wird. Es erfolgte die Sequenzierung von dem cpGenom zwölf unterschiedlicher Haselnussorten mittels NGS (*Next Generation Sequencing*). Die anschließende Assemblierung erfolgte gegen die in NCBI veröffentlichten cpGenome einer *C. chinensis* (KX814336.2) und *C. avellana* (KX822768.2) unbekannter Sorten. Auf Basis des cpGenoms war es nicht möglich sortenspezifische Sequenzunterschiede für den gezielten Nachweis einer Sorte zu identifizieren. Es wurden jedoch innerhalb der LSC-Region (*large single copy region*) zwei *Indels* mit Längen von 115 bzw. 118 bp identifiziert, die eine Gruppierung der untersuchten Sorten ermöglichen, denn es zeigte sich eine geografische Clusterung nach dem Anbaugebiet. Der qualitative Nachweis erfolgte mit einer einfach durchzuführenden PCR (*polymerase chain reaction*) und anschließender Trennung und Detektion mittels AGE (Agarosegelelektrophorese). Die Anwendbarkeit der Methode wurde auch anhand von verarbeiteten Haselnussprodukten überprüft. Des Weiteren wurde auf Basis der *Indels* ein quantitativer Ansatz entwickelt, um Mischungen von Haselnüssen der beiden Gruppen und damit auch wirtschaftlich relevante Verfälschungen wie z.B. die Beimischung kostengünstiger georgischer zu teuren italienischen Haselnüssen identifizieren und semi-quantitativ abschätzen zu können. Als Detektionsmethode wurde hierfür die Kapillargelelektrophorese (CGE, engl. *capillary gel electrophoresis*) herangezogen. Zudem wurde als kostengünstige Alternative zur CGE die Graustufenanalyse der Agarosegelelektrophorese evaluiert. Die Nachweisgrenzen (LOD, engl. *limit of detection*) lagen hierbei bei einem Anteil georgischer Haselnüsse von 5%. Es wurde zudem in der LSC-Region ein Einzelbasenaustausch (SNP, engl. *single nucleotide polymorphism*) identifiziert, der spezifisch für Haselnussproben aus Georgien ist. Dieser resultiert in den Proben georgischer Herkunft in einer Restriktionsschnittstelle des Enzyms *MfeI*, sodass basierend darauf eine PCR-RFLP (*polymerase chain reaction – restriction fragment length polymorphism*) für den spezifischen Nachweis georgischer Haselnüsse entwickelt werden konnte. Auch hier wurde sowohl ein qualitativer, als auch ein quantitativer Ansatz vergleichbar zu dem Nachweis der *Indels* verfolgt. Der LOD lag dabei, bezogen auf den Anteil georgischer Haselnüsse, bei 10%. Ebenso wurde die Einsatzmöglichkeit der Methode auf verarbeitete Produkte überprüft. Neben der PCR-RFLP wurde auf demselben DNA-Bereich eine hochauflösende Schmelzpunktanalyse (HRMA, engl. *high resolution melting analysis*) entwickelt. Mit dieser können die Haselnussorten in drei Gruppen unterteilt werden. Zum einen bilden sich hierbei dieselben Gruppierungen vergleichbar zu den identifizierten *Indels* aus und zum anderen führt der SNP bei georgischen Haselnussproben, für den auch die PCR-RFLP entwickelt wurde, zu einer weiteren Abtrennung dieser Proben. Die Unterteilung, die basierend

auf der Kombination aus den Indels und der PCR-RFLP möglich ist, kann mittels HRMA in einem *closed tube*-Ansatz nachgewiesen werden. Sie stellt damit eine alternative Analysemöglichkeit dar. Die Anwendbarkeit der Methode wurden ebenfalls hinsichtlich verarbeiteter Produkte überprüft. Ebenso wurde auch hier die Möglichkeit der quantitativen Abschätzungen von Mischungen evaluiert.

Neben dem cpGenom wurden auch spezifische Bereiche des Kerngenoms analysiert, die sich in verschiedensten Publikationen bereits als Zielregionen für die Identifizierung von Sequenzunterschieden und die Differenzierung biologischer Identitäten bewährt haben. Zum einen wurden die ITS-Regionen und zum anderen die WRKY-Gene auf Sequenzunterschiede hin untersucht. Bei beiden Ansätzen konnten heterozygote SNPs identifiziert werden, die allerdings nicht sortenspezifisch waren.

Neben den *targeted*-Ansätzen basierend auf identifizierten Sequenzunterschieden wurde auch der *non targeted*-Ansatz der SCoT-Marker (*start codon targeted*) auf 13 unterschiedliche Haselnussorten angewandt. Es handelt sich dabei um eine zu der RAPD-PCR vergleichbaren, jedoch weniger bekannten Methode, bei der die PCR ebenfalls mit nur einem Primer durchgeführt wird. Aus der Analyse resultieren Bandenmuster, die als spezifischer Fingerabdruck jeder Sorte angesehen werden können. Die anschließende Auswertung erfolgte sowohl mittels eines *Neighbour-Joining*-Diagramms, als auch mit einer Hauptkomponentenanalyse (PCA, engl. *principal component analysis*). Mit beiden Methoden konnte gezeigt werden, dass die Proben sortenspezifisch zusammen clustern. Zum Teil war auch ein geographischer Zusammenhang erkennbar, wobei der Sorteneinfluss überwiegt. Auf Basis einer zuvor von Felbinger et al. (2020) entwickelten RAPD-PCR, sollte zudem der Ansatz der sogenannten SCAR-Marker (*sequence characterized amplified region*) analysiert werden. Hierbei dienen die Ergebnisse der RAPD-PCR als Grundlage für die Entwicklung spezifischer Primer, mit dem schließlich der zuvor identifizierte Marker schnell und gezielt nachgewiesen werden kann. Es war jedoch nicht möglich basierend auf den Ergebnissen Sequenzunterschiede zu identifizieren oder spezifische Primer zu entwickeln.

Die Ergebnisse dieser Arbeit zeigen, dass die unterschiedlichen Haselnussorten sehr nah miteinander verwandt sind und es daher äußerst schwierig ist sortenspezifische Sequenzunterschiede zu identifizieren. Dennoch war es möglich Differenzen auf dem cpGenom zu identifizieren und darauf basierend Methoden zu entwickeln, mit denen die Authentizität von Haselnüssen in Bezug auf relevante Verfälschungen hin überprüft werden kann. Die entwickelten Analysen können daher sowohl im Rahmen der Lebensmittelüberwachung, als auch von der verarbeitenden Industrie herangezogen werden, um die biologische Identität und Herkunft von Haselnüssen neben der Prüfung der Dokumente und Frachtpapiere auch analytisch eingrenzen zu können.

2 Abstract

This work presents the development of molecular biological methods for determination of authentication and the differentiation of hazelnut (*corylus avellana*) cultivars. First of all, polymorphic variations and sequence differences should be identified. Subsequently, analytical methods were to be developed which allow both the qualitative detection and the identification and quantitative estimation of mixtures.

The focus of this study was on the analysis of the chloroplast genome, as this is inherited maternally in contrast to the nuclear and mitochondrial genomes. The cpGenome of twelve different hazelnut cultivars from seven growing countries was sequenced using NGS (Next Generation Sequencing). Subsequent assembly was performed against the cpGenomes published in NCBI of a *C. chinensis* (KX814336.2) and *C. avellana* (KX822768.2) unknown cultivars. Based on the cpGenome, it was not possible to identify specific sequence differences for targeted detection of cultivars. However, two indels with lengths of 115 and 118 bp were identified within the LSC (large single copy) region, allowing grouping of the investigated cultivars. Qualitative detection was performed using a simple PCR (polymerase chain reaction) followed by separation and detection by AGE (agarose gel electrophoresis). The results showed the geographical clustering into two groups. The suitability of the method was also verified using processed hazelnut products. Furthermore, based on the indels, a quantitative approach was developed to identify and semi-quantitatively estimate mixtures of hazelnuts from the two groups and thus also economically relevant adulterations such as the admixture of low-cost Georgian to expensive Italian hazelnuts. Capillary gel electrophoresis (CGE) was used as the detection method for this purpose. In addition, gray scale analysis of agarose gels was evaluated as a cost-effective alternative to CGE. The detection limits (LOD) were at a level of Georgian hazelnuts of 5%. In addition, a single nucleotide polymorphism (SNP) was identified in the LSC region, which is specific for hazelnut samples from Georgia. The SNP results in a restriction site of the enzyme *Mfel* in the samples of Georgian origin, so that based on this a PCR-RFLP (polymerase chain reaction - restriction fragment length polymorphism) for the specific detection of Georgian hazelnuts could be developed. Both a qualitative and a quantitative approach comparable to the detection of indels was pursued. The LOD was 10%, based on the proportion of Georgian hazelnuts. The applicability of the method to processed products was also verified. In addition to PCR-RFLP, a high resolution melting analysis (HRMA) was developed on the same DNA region. With this, the hazelnut cultivars can be divided into three groups. The same grouping, comparable to the identified indels, was shown and in addition, the SNP in Georgian hazelnut samples leads to a further separation of these samples. The subdivision, which is possible based on the combination of the indels and PCR-RFLP, can be detected using HRMA in one simple approach. It thus represents an alternative analytical option. The applicability of the method was also verified with respect to processed products. Also, the possibility of quantitative estimation of mixtures was evaluated. In addition to the cpGenome, specific regions of the nuclear genome were analyzed that have already been shown in various publications to be target regions for the identification of sequence differences and the differentiation of biological identities. First, the ITS region and second, the WRKY genes were

analyzed for sequence differences. In both approaches, heterozygous but not variety-specific SNPs could be identified.

In addition to the targeted approaches based on identified sequence differences, the non-targeted approach of SCoT (start codon targeted) markers was also applied to 13 different hazelnut cultivars. This is a method comparable to RAPD-PCR (randomly amplified polymorphic DNA), but less well known, in which the PCR is also performed with only one primer. The analysis results in banding patterns that can be considered as a specific fingerprint of each cultivar. Subsequent analysis was performed using both a neighbour-joining diagram and a principal component analysis (PCA). Both methods showed that the samples clustered together in a variety-specific manner. In part, a geographic correlation was also apparent, but the varietal influence seems to be dominating. Based on a RAPD-PCR previously developed by Felbinger et al. (2020), the approach of SCAR (sequence characterized amplified region) markers was also analyzed. The results of the RAPD-PCR serve as a basis for the development of specific primers, with which the previously identified marker can finally be detected quickly and specifically. However, it was not possible to identify sequence differences or develop specific primers based on the results.

The results of this work showed that the different hazelnut cultivars are very closely related to each other and it is therefore extremely difficult to identify variety-specific sequence differences. Nevertheless, it was possible to identify differences on the cpGenome and, based on this, to develop methods to verify the authentication of hazelnuts with respect to relevant adulterations. The developed analyses can therefore be used both in the context of food monitoring and by the processing industry to narrow down the biological identity and origin of hazelnuts analytically in addition to checking the documents.

3 Einleitung

Das Bewusstsein für Lebensmittel ist bei den Verbrauchern in den letzten Jahren enorm gestiegen. Fragen wie „Was wird genau verzehrt?“, „Wo kommen die Lebensmittel her?“, „Wie werden sie produziert?“ gewinnen zunehmend an Bedeutung. Das Vorhandensein entsprechender Informationen und damit auch die Überprüfung eben dieser werden immer wichtiger. Gleichzeitig führt die zunehmende Globalisierung zu längeren und zunehmend komplexeren Warenströmen, die schwieriger zu kontrollieren sind. In diesem Zuge rückt auch die Prüfung der Authentizität und die Identifizierung von ge- und verfälschten Lebensmitteln immer mehr in den Fokus der Lebensmittelüberwachung. Jährlich entstehen der globalen Lebensmittelindustrie aufgrund von *Food Fraud* ein geschätzter Schaden im Bereich von 10-15 Mrd. Dollar.^[1] *Food Fraud* bzw. *Food Fakery* bezeichnet dabei die Verfälschung von Lebensmitteln, bspw. durch Um- oder Falschetikettierung, aber auch die Streckung hochwertiger Produkte mit minderwertiger Ware oder die Beimengung nicht deklarierter Stoffe. Es geht dabei primär um die absichtliche Täuschung, i.d.R. wirtschaftlich und/oder finanziell motiviert, sowohl von Verbrauchern, als auch der verarbeitenden Industrie.

Food Fraud wie er bspw. 2011 im Rahmen des „Pferdefleischskandals“, bei dem nicht deklariertes Pferdefleisch in Lasagne nachgewiesen wurde, in den Medien weit verbreitet war, führt zu einem enormen Vertrauensverlust der Verbraucher in die verarbeitende Industrie und verursacht dieser gleichzeitig große Imageschäden und finanzielle Einbußen. Lebensmittelverfälschungen können jedoch auch gesundheitsschädliche Auswirkungen haben wie der Melaminskandal 2008 zeigte. Auch Fehldeklarationen oder Beimischungen von Lebensmitteln mit allergenem Potential können zu ernststen gesundheitlichen Komplikationen führen.

Auf EU-Ebene wurden daher 2011 die von Europol und Interpol koordinierten OPSON-Operationen ins Leben gerufen. Sie haben das Ziel Betrug im Bereich der Lebensmittel aufzudecken und zu bekämpfen. Im Rahmen der Operation OPSON VI in den Jahren 2016/2017 standen Haselnüsse und Haselnusserzeugnisse aus den Ländern Türkei, Italien und Georgien im Fokus der Ermittler. Hierbei wurden Beimischung anderer Nussarten wie Erdnüsse, Mandeln oder Cashews identifiziert. Solche Verfälschungen sind u.a. auf eine steigende Nachfrage bei gleichbleibenden Erntemengen und gleichzeitig höheren Gewinnen durch Streckung mit günstigerer Ware zurückzuführen.

Für den Nachweis vieler potentieller Verfälschungen gibt es bereits etablierte Methoden. Die Authentizität von Haselnüssen in Bezug auf Herkunft und Sorte erfolgt jedoch immer noch weitestgehend durch Prüfung der Frachtpapiere und Lieferdokumente. Für den einfachen und schnellen analytischen Nachweis der Herkunft oder Sorte von Haselnüssen und v.a. für die Detektion und Quantifizierung von Mischungen fehlen bislang jedoch geeignete Methoden. Da solche Streckungen sowohl für den Verbraucher, als auch die verarbeitende Industrie neben qualitativen Einbußen auch finanzielle Schäden bedeuten, ist die Entwicklung neuer Methoden für die Authentizitätsbestimmung in diesen Bereichen enorm wichtig.

4 Grundlagen

4.1 Haselnuss

4.1.1 Botanische Einordnung

Die Hasel (*Corylus*) gehört zu der Familie der Birkengewächse (Betulaceae) und der Gattung der Haselnussgewächse (Coryloideae).^[2] Ursprünglich beheimatet ist sie in Europa und Kleinasien und mittlerweile auch in den gemäßigten Regionen der nördlichen Hemisphäre verbreitet.^{[3],[4]} Die Hasel wächst monochasial als laubabwerfender Strauch oder Baum.^{[2],[5]} Es handelt sich um eine monözische Samenpflanze.^{[2],[5]} Das bedeutet, dass sowohl die männlichen, als auch die weiblichen Blüten an einer Pflanze zu finden sind. Die männlichen Blüten treiben als dichasiale Teilblütenstände aus den Achselknospen der vorjährigen Laubblätter aus.^[5] Sie bilden 8-10 cm lange Ähren (sogenannte Kätzchen), welche gruppiert über eine bewegliche Achse verbunden sind. Auch die weiblichen Blüten wachsen ebenfalls dichasial.^[5] Sie bestehen lediglich aus einem Fruchtknoten, welcher in zwei langen, roten Narbenästen endet.^[5] Sie werden von Knospenschuppen (sogenannten Tegmenten) umgeben und sind nur wenige Millimeter lang.^[5] Der Reifezeitpunkt der Blüten ist dabei zeitlich getrennt voneinander (dichogam).^[2] Während die männlichen Blüten bereits im November ausgebildet werden, reifen die weiblichen Blüten erst im Zeitraum von Januar bis März heran.^[5] Bei der Haselnuss handelt es sich um eine windblütige (anemophile) Pflanze.^{[2],[5]} Sie ist also auf eine Bestäubung durch den Wind angewiesen. Aufgrund ihrer Selbstinkompatibilität ist eine Selbstbefruchtung bzw. Fremdbestäubung durch eine Art mit demselben Inkompatibilitätsmuster ausgeschlossen.^[4] Nach der Bestäubung erfolgen nach einigen Wochen die Befruchtung und die Entwicklung der Früchte (einsamige Nüsse) bis zum Herbst.^[5] An einem Blütenstand bilden sich ein bis drei Früchte aus, welche becherförmig von einer blattartigen Hülle (Cupula) umgeben sind.^[5]

Auf der nördlichen Hemisphäre sind je nach Publikation zwischen zehn und 25 *Corylus*-Arten verbreitet.^{[2],[6]} Alle sind diploid und bilden essbare Nüsse aus.^[2] Bei der Haselnuss handelt es sich um die Frucht der Hasel und um eine echte Nuss im botanischen Sinne, da alle drei Schichten des Perikarps vollständig verholzt sind.^[7] Der Teil der verzehrt wird und umgangssprachlich ebenfalls als Nuss bezeichnet wird, ist lediglich der Samen mit einer dünnen ihm anliegenden Samenschale.^[5] Bei der am häufigsten kultivierten Art handelt es sich um die *C. avellana*, die auch als Europäische Hasel bezeichnet wird.^[2] Aufgrund ihres hohen Grads an Polymorphismen unterteilen manche Taxonomen diese Art weiter. So ordnen bspw. einige Publikationen türkische Haselnüsse den Arten *C. maxima* Mill. und *C. pontica* C. Koch zu. Dies ist allerdings sehr umstritten und wird ebenso oft in Veröffentlichungen abgelehnt.^{[4],[6],[8],[9]}

4.1.2 Bedeutung unterschiedlicher Haselnusssorten und deren Anbau

Die Anbauggebiete der Haselnuss liegen im Bereich des 45. Breitengrads und reichen von den USA über Europa, speziell Frankreich, Italien und Spanien, über die Türkei bis hin zum Kaukasus.^[10] Viele wichtige Haselnusssorten entstammen ursprünglich Wild-Sammlungen von Haselnüssen unbekannter Sorte in den entsprechenden Anbaugebieten.^{[2],[8]} Die ersten Bemühungen die Sorten genetisch hinsichtlich ihrer agronomischen Merkmale, wie z.B. Ölgehalt, Blanchierfähigkeit, Größe der Kerne sowie Resistenzen gegenüber Schädlingen, zu verbessern erfolgten in den 1960er Jahren.^{[8],[11]} Die Vermehrung kann sowohl mittels Stecklingen, als auch über Absenker oder das Verfahren der Pfropfung erfolgen. Stecklinge werden v.a. in der Türkei, Italien und Spanien eingesetzt. In Frankreich und den USA (v.a. Oregon) dominiert das Verfahren der Absenker.^[2] Die Haselnuss bevorzugt durchlässige, fruchtbare Böden. Sie wächst beispielsweise auf kalkhaltigen, vulkanischen oder lehmigen Böden, deren pH-Wert im Optimalfall bei etwa 6,5 liegt. Sie benötigt ein mildes Klima und eine hohe Luftfeuchtigkeit sowie milde Winter.^{[2],[11]} Bisher wurden weltweit rund 400 Sorten beschrieben. Davon sind jedoch nur etwa 20 für die Wirtschaft von Bedeutung.^[4] Sie unterscheiden sich u.a. in ihren ökologischen Ansprüchen, Krankheitsresistenzen, Verwendungsmöglichkeiten und Aromaprofilen.

Im Folgenden wird der Haselnussanbau des größten Haselnussproduzenten, der Türkei, sowie die wichtigsten Anbauländer in Europa detaillierter erläutert. In der Türkei erfolgt der Haselnussanbau v.a. in den Regionen Giresun, Ordu, Akçakoca und Trabzon.^{[2],[12]} Diese liegen alle an der Schwarzmeer-Küste, wo das Klima den Anbau begünstigt. Am weitesten verbreitet sind die Sorten Tombul, Çakildak, Mincane, Karafindik, Foşa und Sivri.^[2] 80% der in Giresun angebaute Haselnüsse sind der Sorte Tombul zuzuordnen.^{[2],[12]} Erfolgt der Anbau speziell in den Regionen Piraziz und Çarşıbaşı werden Nüsse dieser Sorte auch als Giresun-Qualität verkauft. Es handelt sich hierbei um sehr hochwertige und teure Nüsse, die sich durch einen hohen Fett-Gehalt, eine dünne Schale und ein leicht zu entfernendes Perisperm auszeichnen.^[13] Alle anderen Sorten werden als Levant-Qualität verkauft und preislich günstiger gehandelt.^[13] Die genaue Bezeichnung richtet sich dabei nach der Provinz des Anbaus. So werden z.B. Nüsse aus Ordu als Levant Ordu verkauft. Hierbei spielt die genaue Sorte keine Rolle.^[13] In der Region Ordu werden vorwiegend die Sorten Palaz, Çakildak und Tombul angebaut.^{[2],[12]} Akçakoca ist ein Gebiet, welches viele ältere Haselnussplantagen aufweist, auf denen häufig verschiedene Klone angebaut werden. Die Sorten Foşa und Mincane sind nicht speziell einem Anbaugebiet zuzuordnen.^[2]

Basierend auf ihrer Form werden türkische Haselnüsse in drei Formen unterteilt: rund, spitz und lang.^{[12],[13]} Zu den runden Haselnüssen zählen die häufigsten angebaute Sorten wie z.B. Çakildak, Foşa, Karafindik, Mincane, Palaz und Tombul. Sie werden aufgrund ihrer qualitativ hochwertigen Nüsse und der für die verarbeitende Industrie besser geeignete Form bevorzugt. Zu spitzen Haselnüssen, welchen meist eine geringere Qualität zugesprochen wird und die aufgrund ihrer Form auch in der Süßwarenindustrie nicht bevorzugt werden, zählen bspw. die Sorten Sivri, Kalinkara, Aci oder Incekara. Zu den langen Haselnüssen, welche in ihrer Form Mandeln ähneln, gehören z.B. die

Sorten Yassi Badem oder Yuvarlak Badem. Sie werden frisch verzehrt, eignen sich jedoch nicht für die weitere Verarbeitung.^{[12],[13]}

Tabelle 1: Synonyme einiger türkischer Haselnussorten.^[12]

Sorte	Synonyme
Tombul	Giresun Yağlısi, Yağlı Findik
Mincane	Sarifindik, Sariyağlı, Sirafindik
Foşa	Yomra
Çakıldak	Delisava, Gökindik
Karafindik	Karayağlı

Der Haselnussanbau in der Türkei ist traditionell und durch viele kleine familiengeführte Unternehmen geprägt. Für viele dort angebaute Sorten existieren unterschiedliche Synonyme, z.T. auch abhängig vom Anbaugebiet. Einige Synonyme wichtiger Sorten sind in Tabelle 1 zusammengefasst.^[12]

In Italien dominieren vier große Anbauregionen: Sizilien, Campanien, Latium und Piemont. Diese unterscheiden sich nicht nur geographisch, sondern auch an den jeweils dort überwiegend angebauten Sorten.^{[2],[14]} In Sizilien wird die Sorte Siciliana angebaut.^[2] In Campanien, wo alleine etwa 1/3 der Gesamternte herkommt, werden v.a. die Sorten Mortarella, San Giovanni und Tonda di Giffoni angebaut.^{[2],[14]} Weniger relevante Sorten, v.a. für den direkten Verzehr geeignet, sind Tonda bianca, Tonda rossa, Camponica und Riccia di Talanico.^[14] Die Sorte Tonda di Giffoni, welche neben Italien auch in Frankreich kultiviert wird, zeichnet sich durch eine gute Verarbeitungsqualität, konstante Erträge sowie eine hohe Resistenz gegenüber Krankheitserregern und Schädlingen aus.^[14] Für diese Sorte gibt es auch einen Eintrag bzgl. einer geschützten geografischen Angabe (g.g.A.) in der Datenbank eAmbrosia der Europäischen Union. Die Angabe „Nocciola di Giffoni“ dürfen nur Haselnüsse der Sorte Tonda di Giffoni aus speziellen Gebieten der Provinz Salerno tragen.^[15] Im Zentrum von Italien in Latium, speziell den Provinzen Rom und Viterbo wird die Sorte Tonda Gentile Romana angebaut. Diese wird v.a. aufgrund ihres Aromaprofils nach der Röstung, welches von grün, fruchtig, erdig und fettig bis hin zu Popcorn- und Kaffee-ähnlich reicht, gerne in der Süßwarenindustrie eingesetzt.^[16] Ein Nachteil der Sorte liegt jedoch in dem sich schlecht zu lösenden Perisperm, weshalb sie überwiegend ungeschält vermarktet wird.^{[2],[4]} Es gibt für diese Sorte auch eine in der Datenbank eAmbrosia der Europäischen Union eingetragene geschützte Ursprungsbezeichnung (g.U.) „Nocciola Romana“.^[15] Diese gilt jedoch nur für Haselnussplantagen, welche zu mind. 90% die Sorte T. G. Romana und max. 10% der Sorten Tonda di Giffoni oder Barrettona anbauen. Mögliche Anbauggebiete sind zudem beschränkt auf spezielle Regionen in den Provinzen Rom und Viterbo.^[15] Das vierte und sehr wichtige Anbaugebiet liegt im Piemont. Dort wird die Sorte Tonda Gentile Trilobata angebaut, früher auch Tonda Gentile della Langhe genannt.^[2] Die Sorte T. G. Trilobata eignet sich sehr gut für die großtechnische Röstung, da hierbei überwiegend vorteilhafte Aromakomponenten wie süß, fruchtig und nach gebackenem Brot entstehen.^{[17],[18]} Zudem lässt sich das Perisperm nach der Röstung sehr leicht entfernen, was sie ebenfalls zu einer in der Süßwarenindustrie bevorzugten Sorte macht. Für spezielle Anbauggebiete dieser Sorte in der Region Piemont in Italien gibt es ebenfalls einen Eintrag in der Datenbank eAmbrosia.^[15] Die g.g.A. „Nocciola Piemonte“ bzw. „Nocciola del Piemonte“ ist Nüssen dieser Sorte aus der Region Piemont vorbehalten.

Die Spanien findet der Haselnussanbau überwiegend in der Region Tarragona in Katalonien statt.^[2] Dort werden knapp 90% der landesweiten Ernte produziert.^[19] Überwiegend wird die Sorte Negret

kultiviert.^[2] Auf neuen Plantagen wird vermehrt auch die Sorte Pauetet angebaut.^[2] Weniger verbreitet sind Sorten wie Gironell, Griffol, Morell oder Culpla.^{[2],[19]} Für Haselnüsse der Sorten Negret, Pauetet, Gironell, Morell und Culpla gibt es ebenfalls in der Datenbank eAmbrosia eine eingetragene geschützte Ursprungsbezeichnung. Voraussetzung ist jedoch ihr Anbau in speziellen Regionen in Tarragona. Werden alle Voraussetzungen erfüllt, dürfen die Haselnüsse unter der g.U. „Avellana de Reus“ verkauft werden.^[15] Spanische und italienische Haselnussorten zeigten bei bisherigen Untersuchungen auf Basis von Mikrosatelliten immer einen hohen Verwandtschaftsgrad auf.^{[6],[8]}

Die großen Haselnussplantagen Frankreichs liegen überwiegend im Südwesten des Landes in der Nähe der Stadt Cancon in der Region Nouvelle-Aquitaine.^[2] Früher dominierte dort v.a. die Sorte Fertile de Coutard (Synonym: Barcelona), welche überwiegend für den *in shell*-Markt angebaut wird.^[2] Zunehmend kamen und kommen die Sorten Ennis und Corabel hinzu. Auf neuere Plantagen wird vermehrt auch die Sorte Pauetet kultiviert.^[2]

4.1.3 Wirtschaftliche Bedeutung und Nutzung

Haselnüsse gelten schon lange aufgrund ihrer Inhaltsstoffe als ernährungsphysiologisch wertvoller Bestandteil unserer Nahrung. Mit 6,29 kcal/g besitzen Haselnüsse eine hohe Energiedichte (> 4 kcal/g).^[20] Den größten Bestandteil stellen hierbei die Fette mit 61,6% dar.^{[20],[21]} Der Lipidanteil setzt sich überwiegend aus ungesättigten Fettsäuren (v.a. Öl- und Linolsäure) zusammen.^{[21]–[23]} Zudem besitzen sie einen hohen Gehalt an Vitamin E.^{[4],[23]} Diese Inhaltsstoffe können sich positiv auf die Gesundheit auswirken und u.a. kardiovaskulären Erkrankungen oder Arteriosklerose entgegenwirken.^{[24]–[26]} Mit 12% Proteinen stellen Haselnüsse zudem eine gute Quelle für pflanzliche Proteine dar.^{[20],[21]} Es wurde bereits mehrfach beobachtet, dass sowohl die Sorte, als auch die Herkunft, Umwelteinflüsse und Anbaubedingungen Auswirkungen auf die Fettsäure-, Sterol- und Tocopherolzusammensetzung sowie den Gehalt an Mineralstoffen (wie z.B. Magnesium, Kupfer oder Calcium) haben können.^{[10],[22],[23],[27]}

Nur etwa 10% der geernteten Haselnüsse werden als *in shell*-Ware verkauft. Die restlichen 90% gelangen weiter zu der verarbeitenden Industrie, wo sie für Schokoladenprodukte, Pralinen, Feingebäck oder Eiscreme eingesetzt werden.^{[4],[28]} Der größte Teil (ca. 80%) wird in der Schokoladenindustrie verarbeitet.^[12] Nur aus einem sehr kleinen Anteil wird Haselnussöl hergestellt, welches neben dem Genuss als Speiseöl auch Anwendung in der Ölmalerei oder in Kosmetikprodukten findet.^[5]

2020 wurden weltweit etwa 1 Mrd. Tonnen Haselnüsse angebaut. Die größten Anbauländer sind hierbei die Türkei, Italien, USA, Aserbaidschan, Georgien und Chile. Führender Produzent ist hierbei mit 665.000 t die Türkei. Das entspricht einem Anteil an der Gesamternte von etwa 64%. Die Preise für Haselnüsse auf dem Weltmarkt sind sehr stark abhängig von dem jeweiligen Anbauland wie Abbildung 1 verdeutlicht. Das obere Preissegment bilden hierbei Haselnüsse aus der Türkei und

Italien mit 3.000 bzw. 3.600 USD/t im Jahr 2020. Nüsse aus Georgien und Aserbaidtschan hingegen werden für lediglich etwa 1.200 USD/t gehandelt.^[29]

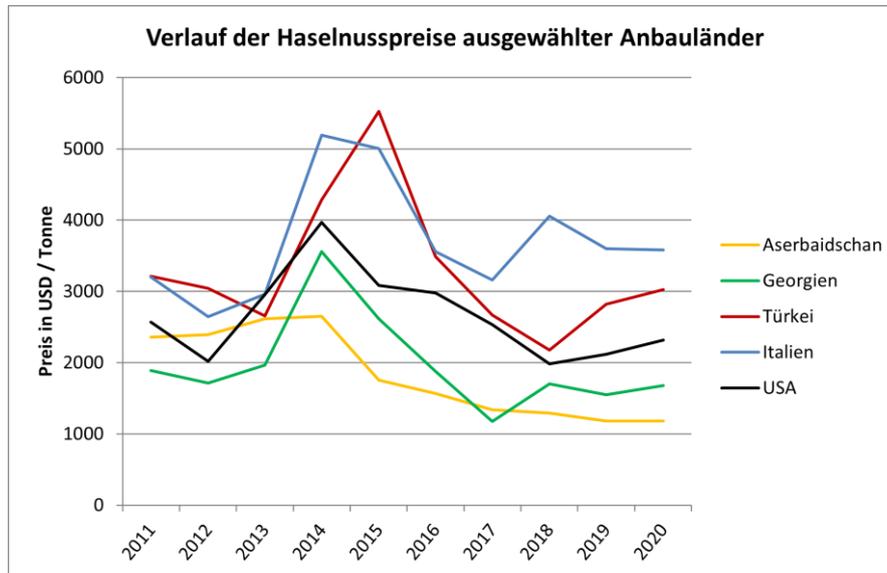


Abbildung 1: Verlauf der Preise für Haselnüsse (mit Schale) der fünf wichtigsten Anbauländer in den Jahren 2011-2020 in USD/Tonne.^[29]

Im Jahr 2014 ist ein deutlicher Preisanstieg zu erkennen. Dieser ist auf einen späten Kälteeinbruch in der Türkei zurückzuführen, welcher im Vergleich zu dem Vorjahr zu einem Preisanstieg von 100-240% auf dem Weltmarkt führte. Um dieser Abhängigkeit von nur einem Anbauland entgegenzuwirken, werden zunehmend auch neue Anbaugelände, wie z.B. Chile, erschlossen. Dort lag die Produktion 2020 bereits bei 33.939 t.^[29]

Die Türkei ist bisher nicht nur führender Produzent, sondern auch Exporteur. Im Zeitraum von 2000-2016 exportierte die Türkei jährlich zwischen 120.000 t und 160.000 t Haselnüsse. Ein Großteil davon wird in Deutschland verarbeitet. Die deutsche Wirtschaft ist einer der größten Importeure von Haselnüssen. Allein im Jahr 2020 wurden fast 70.000 t importiert.^[30] Das entspricht ca. 6,4% der Welternte.^[29]

4.1.4 Food Fraud im Bereich der Haselnusserzeugnisse

Food Fraud bezeichnet allgemein Verstöße gegen das Lebensmittelgesetz im Bereich der Verbrauchertäuschung. Lebensmittelverfälschungen stellen hierbei ein Problem sowohl für die Verbraucher als auch die Wirtschaft dar. Das Vertrauen der Verbraucher in die lebensmittelverarbeitende Industrie kann nachhaltig geschädigt werden und die Betriebe können dadurch Imageschäden sowie wirtschaftliche Folgen erleiden.

Im Bereich der Haselnüsse bieten bereits die geschilderten Preisunterschiede zwischen den verschiedenen Anbauländern einen Anreiz zur Streckung hochwertiger bzw. teurer mit minderwertiger/günstiger Ware. Des Weiteren gibt es auch bei Haselnüssen Produkte mit der

Auslobung einer geschützten Herkunftsbezeichnung. Solche Angaben führen häufig zu erhöhten Preisen und stellen damit ebenfalls einen Anreiz für die fälschliche Nutzung dar. Bei den geschützten Herkunftsbezeichnungen wird primär zwischen einer geschützten Ursprungsbezeichnung (g.U.) und einer geschützten geografischen Angabe (g.g.A.) unterschieden. Die entsprechenden Siegel sind in Abbildung 2 dargestellt.



Abbildung 2: Siegel für Produkte mit einer geschützten Herkunftsangabe. Links: geschützte Ursprungsbezeichnung (g.U.), Rechts: geschützte geografische Angabe (g.g.A.).

Aktuell sind in der Datenbank eAmbrosia (EU-Register der geografischen Angaben) der Europäischen Kommission zwei geschützte Ursprungsbezeichnungen und drei geschützte geografische Angaben aus Italien, Frankreich und Spanien eingetragen.^[31] Ausschlaggebend sind hierbei, wie unter 4.1.2 bereits anhand einiger Beispiele vorgestellt, sowohl spezielle Haselnussorten sowie entsprechende Anbauggebiete. So ist beispielsweise die Bezeichnung „Nocciola del Piemonte“ bzw. „Nocciola Piemonte“ Haselnüssen der Sorte Tonda Gentile

Trilobata aus der Region Piemont in Italien vorbehalten.^[31] Nüsse mit solchen Auslobungen zeichnen sich in der Regel durch besondere Eigenschaften aus. Dies spiegelt sich dementsprechend auch im Preisniveau wider. Aus diesem Grund ist es notwendig, Produkte mit den genannten Siegeln auf ihre Richtigkeit zu überprüfen.

Um Lebensmittelverfälschungen aufzudecken gibt es seit 2011 die von Europol und Interpol ins Leben gerufenen OPSON-Operationen. Es handelt sich hierbei um ein länderübergreifend koordiniertes Vorgehen gegen Lebensmittelbetrug.^[32] In den Jahren 2016/2017 waren Haselnusserzeugnisse speziell aus den Ländern Türkei, Georgien und Italien im Rahmen der Operation OPSON VI im Visier der Ermittler.^[33] Ausgangspunkt war hierbei der Verdacht auf Beimischungen von Erdnüssen, Mandeln und Cashewkernen. Diese konnten auch bestätigt werden. Es wurden Streckungen durch die genannten Nüsse im Bereich von 8-45% nachgewiesen.^[33] Begründet wurde der Verdacht auf die enorm gestiegenen Preise von Haselnüsse aufgrund der kältebedingten Ernteausfälle in der Türkei im Jahr 2014. Diese führten dazu, dass andere Nüsse wie Mandeln oder Cashewkerne zu geringeren Preisen gehandelt wurden.

4.2 DNA

4.2.1 Genomics –basierte Ansätze in der Lebensmittelüberwachung

Genomics ist eine der vier *Omics*-Disziplinen und bildet zusammen mit *Transcriptomics*, *Proteomics* und *Metabolomics* die sogenannte *Omics*-Kaskade. Die vier Teilgebiete bilden ein aufeinander aufbauendes System.^[34] Den Anfang bildet das Genom. Es definiert den Genotyp eines Organismus und ist stets individuell und eindeutig.^[35] Das Genom wird über den Prozess der Transkription in einzelsträngige RNA, das sogenannte Transkriptom übersetzt. Dieses wiederum bildet die Vorlage für

die Proteine, die im Zuge der Translation gebildet werden.^[36] Die Summe aller Proteine bildet das Proteom. Insgesamt können bei diesem Prozess aus einem Gen funktionell unterschiedliche Proteine gebildet werden.^[35] Dies gründet u.a. auf verschiedenen Regulationsmechanismen. Zudem können unterschiedliche posttranslationale Modifikationen bei den Proteinen auftreten.^[36] Das Proteom beeinflusst schließlich die von dem Organismus gebildeten Metabolite, also das Metabolom, das damit auch den Endpunkt der *Omics*-Kaskade markiert und den größten Einfluss auf den Phänotyp eines Organismus besitzt.^{[34],[35]} Grundsätzlich nimmt die Varianz vom Genom hin zum Metabolom zu. Gleichzeitig steigt auch der Einfluss exogener und anthropogener Faktoren.^{[35],[37]}

Unter Berücksichtigung aller *Omics*-Disziplinen kann im Bereich der Lebensmittel ein genaues *Food Profile* erstellt werden, das sowohl Rückschlüsse auf die Art des Organismus, als auch seine Herkunft/Ursprung und die Art des Anbaus bzw. der Aufzucht zulässt.^{[35],[37]} Aufgrund der begrenzten Stabilität des Transkriptoms wird häufig anstelle dessen das Isotopenprofil (*Isotopolomics*) mit einbezogen.^[37] Das Metabolom und Isotopenprofil eignen sich sehr gut für den Herkunftsnachweis, da hier äußere Faktoren einen stärkeren Einfluss aufweisen.^[38] Bei dem Genom hingegen fallen diese Einflüsse so gering aus, dass dieses in den in der Lebensmittelanalytik betrachteten Zeiträumen als unveränderlich angesehen werden kann. Daher wird es im Bereich der Authentizitätsbestimmung bevorzugt für den Nachweis und die Prüfung der biologischen Identität herangezogen.^[37] So werden molekularbiologische Methoden basierend auf der Detektion spezifischer DNA z.B. für den Nachweis von Pferdefleisch in verarbeiteten Fleischprodukten eingesetzt.^{[39],[40]} Auch im Rahmen der unter 4.1.4 erwähnten OPSON-Operation wurden neben immunologischen Methoden auch molekularbiologische Ansätze für die Bestätigung auffälliger Befunde herangezogen.^[41] Insgesamt gibt es eine Vielzahl wissenschaftlicher Publikationen, die sich mit der Methodenentwicklung zur Überprüfung der biologischen Identität von Lebensmitteln beschäftigen, wie z.B. der Unterscheidung zwischen Marzipan und Persipan^[42], der Differenzierung von Edel- und Konsumkakao^[43] oder der Nachweis unterschiedlicher Fischarten^{[44],[45]}.

Im Rahmen des *Genomics* besteht entweder die Möglichkeit der Differenzierung anhand des genetischen Fingerabdrucks oder basierend auf dem spezifischen Nachweis von Sequenzunterschieden mit Hilfe zielgerichteter, sogenannter *targeted*-Methoden. Die Gesamtheit des Genoms, welches für die Methodenentwicklung herangezogen werden kann, setzt sich dabei aus dem Kern- und Mitochondriengenom zusammen. Bei Pflanzen kommt zusätzlich noch das Genom der Chloroplasten hinzu.

4.2.2 Nukleare DNA

Bei Eukaryonten ist das Kerngenom im Zellkern lokalisiert und in mehreren linearen Chromosomen organisiert. Jedes Chromosom setzt sich aus zwei Chromatiden zusammen, die am Centromer miteinander verbunden sind.^[46] Ein Chromatid wiederum ist aus Chromatin aufgebaut, also der doppelsträngigen DNA und chromosomalen Proteinen sowie Histonen. 146 bp der Doppelhelix winden sich dabei jeweils um ein Histonoktamer und bilden damit ein Nucleosom.^[47] Diese werden

durch Interaktion mit anderen chromosomalen Proteinen enger gepackt und bilden schließlich das Chromatin. Ohne diese enge Packung wäre eine Unterbringung der DNA im Zellkern aufgrund ihrer Größe nicht möglich.^[47] Homologe Chromosomen eines diploiden Chromosomensatzes sind bei obligaten Selbstbefruchtern homozygot, bei Fremdbestäubern hingegen heterozygot, da zwar der Grundaufbau und Genbestand gleich bleiben, aber i.d.R. Unterschiede in der Basensequenz vorliegen.^[7]

Das Genom beinhaltet neben Genen auch nicht-kodierende Bereiche und hochrepetitive Sequenzen. Diese können sowohl zwischen als auch innerhalb der Gene liegen. Bei den innerhalb eines Gens liegenden nicht-kodierenden Bereichen handelt es sich um sogenannte Introns.^[47] Die kodierenden Abschnitte eines Gens werden als Exons bezeichnet. Zudem besitzen sie Abschnitte mit regulatorischen Funktionen wie Promotor und Terminator. Im Prozess der Genexpression wird im ersten Schritt durch die Transkription die DNA abgelesen. Das entstehende primäre Transkript, die sogenannte Prä-mRNA, weist auch die Introns noch auf. Sie werden bei dem Vorgang des sogenannten Spleißens entfernt, sodass die reife mRNA lediglich aus den kodierenden Exons besteht.^[47] Diese wird vom Zellkern ins Cytoplasma transportiert, wo anschließend die Translation stattfindet.^[47] Hierbei erfolgt die Übersetzung der Geninformation in eine Aminosäuresequenz, welche schließlich das codierte Protein bildet.

Die Haselnuss besitzt einen diploiden Chromosomensatz mit 11 Chromosomen.^[11] Die Größe eines Genoms liegt abhängig vom Organismus im Bereich von 10^6 - 10^{11} bp.^[36] Sie kann mittels Durchflusszytometrie abgeschätzt werden.^[48] Das Genom der *C. avellana* wurde über diese Methode 2012 am USDA-ARS-NCGR (United States Department of Agriculture, Agricultural Research Service, National Clonal Germplasm Repository) auf eine Länge von 378 Mbp geschätzt.^[11] Im Vergleich dazu besitzt bspw. das Humangenom eine Länge von 2,91 Gbp, das Genom der Walnuss (*Juglans regia*) 667 Mbp und das Genom des Pfirsichs (*Prunus persica*) 265 Mbp.^{[49]–[51]} Der Längenunterschied ist zum größten Teil auf die nicht-kodierenden Bereiche zurückzuführen. Er sagt nichts über die Komplexität eines Organismus aus. Ein längeres Genom steht nicht automatisch für eine höhere Anzahl an Genen. Rowley et al. (2018) gelang eine *de novo* Sequenzierung des Kerngenoms der Haselnuss. Es konnten 345 Mbp (entspricht ca. 91 % der geschätzten Länge) assembliert werden.^[11] Es wurden knapp 35.000 Protein-kodierende Bereiche vorhergesagt. Hiervon konnten etwa 22.500 erfolgreich annotiert werden. 82,5 % dieser Proteine sind ebenfalls in den Sequenzen der nahverwandten Gattungen *Vitis*, *Prunus*, *Populus* und *Ricinus* vorhanden.^[11]

4.2.2.1 ITS-Bereich

Der ITS-Bereich (*internal transcribed spacer region*) ist ein Bestandteil der ribosomalen DNA (rDNA), welche die rRNA (ribosomale RNA) kodiert.^[52] Die Abschnitte der rDNA waren die ersten eukaryotischen Gene, die molekularbiologisch analysiert wurden.^[53] Neben dem Karyom weisen auch die Genome der Chloroplasten und Mitochondrien rDNA-Gene auf.^[36] Größere Bedeutung in wissenschaftlichen Untersuchungen erlangte jedoch die rDNA des Karyoms.

Grundsätzlich gibt es vier rDNA-Abschnitte: 18S, 5,8S, 28S und 5S. Die Bezeichnung beruht dabei auf der Größe der entsprechenden rRNA, welche in der Svedberg-Einheit (S) angegeben wird.^[52] Die ersten drei Genabschnitte liegen als Gen-Kassette vor, die mehrfach hintereinander wiederholt wird (*rDNA-repeat*).^{[52],[54]} Die Anzahl der *rDNA-repeats* schwankt sehr stark sowohl intra-, als auch interspezifisch. Es wird hierbei ein Zusammenhang zwischen der Genomgröße und der Anzahl an Wiederholungen vermutet.^[55] Pflanzen weisen i.d.R. deutlich höhere Kopienzahlen (250-20.000 (haploid)) im Vergleich zu Tieren mit 100-1.000 auf.^[56] So wurden bei Untersuchungen der rDNA der Ackerbohne (*Vicia faba*) zwischen 250 und 22.000 Wiederholungen nachgewiesen. Es zeigten sich hierbei neben den intraspezifischen Differenzen auch Unterschiede in den verschiedenen Geweben einer Pflanze.^[57] Die einzelnen rDNA-Abschnitte treten in der genannten Reihenfolge hintereinander auf. Sie werden durch zwei nicht-kodierende Bereiche (ITS-1 und ITS-2) voneinander getrennt. Zwischen der 18S rDNA und 5,8S rDNA befindet sich der ITS-1 und zwischen der 5,8S rDNA und 28S rDNA der ITS-2. Im Folgenden wird der Abschnitt bestehend aus ITS-1, 5,8S rDNA und ITS-2 als ITS-Bereich bezeichnet.^[58] Den Anfang einer Transkriptionseinheit stellt der ETS-Bereich (*external transcribed spacer region*) dar. Am Ende folgt der NTS-Bereich (*non-transcribed spacer region*), welcher zwei *rDNA-repeats* voneinander trennt.^{[52],[54]} Der Aufbau ist schematisch in Abbildung 3 dargestellt.

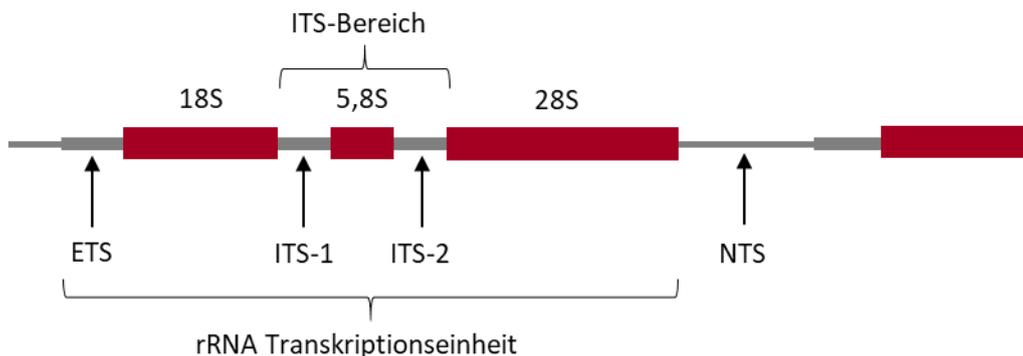


Abbildung 3: Schematische Darstellung einer rRNA-Transkriptionseinheit (*rDNA-repeat*) bestehend aus der 18S rDNA, 5,8S rDNA und 28S rDNA sowie den beiden *internal transcribed spacer regions* ITS-1 und ITS-2, dem ETS-Bereich (*external transcribed spacer region*) und dem NTS-Bereich (*non-transcribed spacer region*) (modifiziert nach Baldwin et al. (1995) und Hillis und Dixon (1991)^{[58],[59]}).

Die rDNA wird durch die RNA-Polymerase I transkribiert. Während der Transkription werden die ITS- und ETS-Bereiche entfernt und anschließend die rRNA-Einheiten voneinander getrennt. Die 5S rRNA wird auf einem anderen Chromosomenabschnitt kodiert und durch die RNA-Polymerase III transkribiert.^[36] Die rRNA-Einheiten bilden zusammen mit den ribosomalen Proteinen die beiden Untereinheiten der Ribosomen. Sie 5S, 5,8S und 28S rRNA bilden hierbei die große 60S Untereinheit und die 18S rRNA die kleine 40S Untereinheit.^[52]

Aufgrund der Variabilität hat sich der ITS-Bereich als Grundlage für Verwandtschaftsanalysen etabliert. Etwa ein Drittel aller veröffentlichter phylogenetischer Studien bezieht sich ausschließlich auf die Abschnitte des Karyoms.^[60] Die auftretenden Sequenzunterschiede machen diese Bereiche zu einem häufig eingesetzten Marker für den Nachweis bzw. die Differenzierung verschiedener Organismen.^{[61]–[65]}

4.2.2.2 WRKY-Gene

Die WRKY-Gene codieren die WRKY-Proteine. Hierbei handelt es sich um eine Gruppe pflanzenspezifischer Transkriptionsfaktoren, welche die physiologische Reaktion einer Pflanze auf abiotischen und biotischen Stress regulieren.^{[66],[67]} Bei abiotischen Stressfaktoren handelt es sich um Umwelteinflüsse wie Trockenheit, Hitze, Frost oder Nährstoffmangel, also Stressfaktoren, die nicht durch Lebewesen verursacht oder beeinflusst werden. Im Gegensatz dazu werden biotische Stressfaktoren durch Lebewesen verursacht. Es handelt sich dabei z.B. um den Befall durch Phytopathogene (wie Viren oder Pilze), Parasiten oder Symbionten.^[68]

Die WRKY-Proteine sind an Prozessen wie der Seneszenz, Trichomentwicklung sowie der Abwehr von Pathogenen und der Synthese sekundärer Metabolite beteiligt.^{[67],[69]} Es handelt sich dabei um Mechanismen, mit denen die Pflanze versucht sich vor dem Einfluss unterschiedlicher Stressfaktoren zu schützen. So weist bspw. das Trichom, die sogenannten Pflanzenhaare, unterschiedlichste Funktionen auf. Es kann als Schutz vor Sonnenstrahlung wirken, aber auch durch seine Optik Insekten anziehen oder im Gegensatz dazu die Pflanze vor Fraßfeinden schützen.^[7]

Die ersten WRKY-Proteine wurden bei der Süßkartoffel, Wildhafer, Petersilie und der Acker-Schmalwand nachgewiesen.^{[70]-[73]} Ihren Namen haben die WRKY-Proteine von der gleichnamigen WRKY-Domäne.^[67] Es handelt sich hierbei um eine 60 Aminosäuren lange Sequenz, mit einem konservierten WRKYGQK-Motiv am N-terminalen Ende und einem atypischen Zink-Finger-Motiv am C-Terminus.^{[67],[73],[74]} Die WRKY-Proteine werden basierend auf der Anzahl vorhandener Domänen sowie der Struktur des Zink-Finger-Motivs in drei Hauptgruppen (I-III) unterteilt. Gruppe I weist dabei zwei Domänen auf und Gruppe II und III hingegen nur eine.^{[67],[74]} Gruppe II kann zudem basierend auf weiteren konservierten Bereichen in fünf Untergruppen (a-e) unterteilt werden.^{[67],[74]}

2005 wurde von Yamasaki et al. mit Hilfe der NMR-Spektroskopie der erste Strukturvorschlag für eine WRKY-Domäne erarbeitet und 2007 von Duan et al. mittels Kristallographie bestätigt.^{[75],[76]} Abhängig von der Länge der Domäne bildet sie ein antiparalleles β -Faltblatt bestehend aus vier bis fünf β -Strängen sowie einem Zink-bindenden Bereich an einem Ende des β -Faltblatts, gebildet aus einem Cystidin- und Histidin-Rest.^{[75],[76]} Die WRKYGQK-Reste bilden dabei den N-terminalen β -Strang, der an der Oberfläche des Proteins liegt und darüber an die große Furche eines DNA-Strangs binden kann. Die WRKY-Proteine binden spezifisch an die W-Box.^[75] Es handelt sich dabei um eine DNA-Sequenz, die als regulatorisches cis-Element fungiert. Sie besitzt die Sequenz TTGAC(C/T), wobei die Basen TGAC ein invariables Kernmotiv bilden, das für die Bindung und die Funktion der WRKY-Proteine enorm wichtig ist.^{[67],[74]} Die Sequenz ist Bestandteil vieler Promotor in pflanzlichen Genen, die im Zusammenhang mit Abwehrreaktionen stehen.^{[77],[78]} Die WRKY-Proteine spielen daher eine entscheidende Rolle bei der Regulierung der Expression dieser Gene.^[67]

Innerhalb des Genoms der *Arabidopsis thaliana* wurden bereits über 70 WRKY-Gene identifiziert. Es ist jedoch davon auszugehen, dass es insgesamt wahrscheinlich bis zu 1000 unterschiedlicher Gene gibt.^{[67],[77]}

Alle WRKY-Gene haben ein Intron gemeinsam, welches die kodierende Region der C-terminalen Domäne der Gruppe I und der Einzel-Domänen der Gruppe II und III unterteilt. Die Position dieses Introns ist hochkonserviert.^[67] Im Zuge der Untersuchungen der WRKY-Proteine des *Theobroma cacao* von Borrone et al. (2004) wurden innerhalb dieses Introns, SNPs identifiziert.^[69] Daher wurden die WRKY-Gene im Rahmen dieser Arbeit als Ziel-Gen für eine Untersuchung bzgl. potentieller Sequenzunterschiede herangezogen und analysiert.

4.2.3 Plastidäre DNA

Neben dem Kerngenom beinhalten pflanzliche Zellen weitere semiautonome Zellorganellen, die ebenfalls eine eigene DNA aufweisen – Mitochondrien und Plastiden. Sie weisen nicht nur eine eigene DNA, sondern auch ein eigenes Replikations- und Transkriptionssystem auf und besitzen die Fähigkeit der selbstständigen Proteinbiosynthese.^[47] Sie besitzen einen prokaryotischen Ursprung. Plastiden gehen dabei auf einen Vorläufer der heutigen Cyanobakterien zurück und sind charakteristisch für photoautotrophe Eukaryoten.^[47] Zu der Gruppe der Plastiden zählen neben den Chloroplasten auch deren degenerierte Form, die sogenannten Gerontoplasten, sowie Chromoplasten, Leukoplasten, Etioplasten oder Amyloplasten.^{[47],[79]} Sie liegen also in unterschiedlichen Differenzierungszuständen mit sowohl funktionellen als auch strukturellen Unterschieden und eigenen, spezifischen Aufgaben vor. Sie gehen jedoch alle auf dieselbe Vorstufe, den Proplastiden, zurück und besitzen die Fähigkeit sich ineinander umzuwandeln.^[47] Lediglich die Bildung der Gerontoplasten erfolgt irreversibel.^[7]

Die Gesamtheit der genetischen Information der Plastiden wird auch Plastom (ptDNA) genannt.^[36] Die genetische Information der unterschiedlichen Plastid-Formen ist dabei identisch.^[79] ptDNA ist zirkulär aufgebaut und setzt sich aus zwei *inverted repeat* (IR) Regionen zusammen, die jeweils durch zwei Einzelkopieabschnitte (*large single copy* (LSC) und *small single copy* (SSC) Region) voneinander getrennt werden (siehe Abbildung 4).^{[36],[80]} Die Größe der ptDNA schwankt in Abhängigkeit des betrachteten Organismus, liegt jedoch meist im Bereich von 120-160 kbp.^{[79],[80]} Auf dem Plastom sind etwa 60-200 Proteine kodiert.^[54] Der Anteil kodierender Sequenzen kann dabei 50-90% der ptDNA ausmachen.^[81] Die kodierten Gene können grob in zwei Kategorien unterteilt werden. Zum einen Gene, die der Erhaltung des genetischen Systems der Plastiden selbst dienen, also bspw. Gene für tRNA, rRNA sowie ein Drittel der ribosomalen Proteine ebenso wie Untereinheiten von RNA-Polymerasen, und zum anderen um Proteinkomponenten für die Photosynthese.^{[36],[54]} Die rRNA-Gene sind dabei

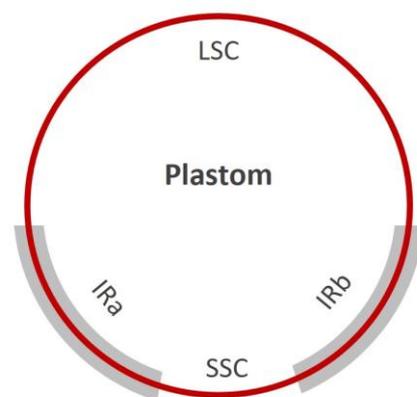


Abbildung 4: Schematische Darstellung des Aufbaus des Plastoms. LSC = *large single copy*, SSC = *small single copy*, IR = *inverted repeat*.^[36]

alle in den IR-Bereichen angesiedelt.^[36] Aufgrund der beschränkten Codierungskapazität sind jedoch letztendlich 90% der plastidenspezifischen Proteine auf dem Kerngenom lokalisiert.^[7]

Ein Chloroplast weist mehrere Kopien der ptDNA auf. Die DNA-Moleküle sind mit RNA und Proteinen in Strukturen, vergleichbar zu den Nucleoiden von Bakterien, organisiert, wobei die Nucleoide der Plastide im Gegensatz zu Bakterien mehrere Kopien der ptDNA enthalten.^{[47],[81]} Bei Untersuchungen der *Beta vulgaris* lagen 4-7 Kopien pro Nucleoid vor.^[82] Die Summe an Nucleoiden und damit auch an Kopienzahlen der ptDNA pro Plastid ist dabei abhängig vom Entwicklungsstand der Pflanze und der Chloroplasten.^{[83]-[86]} Im Falle der *Beta vulgaris* lagen letztendlich 12-330 Kopien der ptDNA pro Plastid vor.^[82] Ob der DNA-Gehalt im Laufe des Alterungsprozesses der Chloroplasten abnimmt und bis zu welchem Zeitpunkt ptDNA nachweisbar ist, ist sehr umstritten.^[81] Die veröffentlichten Informationen verhalten sich diesbezüglich sehr kontrovers.^{[79],[83],[87],[88]}

Die Vererbung des Plastoms erfolgt in den meisten Fällen uniparental.^[89] Es gibt jedoch auch Ausnahmen beispielsweise innerhalb der Gattung *Pelargonium*, welche eine biparentale Vererbung des mt- und cpGenoms aufweisen.^[90] Im Fall der Koniferen verläuft die Vererbung paternal.^[91] Bei Angiospermen hingegen erfolgt sie zu 80% maternal.^[89] Speziell für die Gattung *Corylus* gibt es bisher keine Veröffentlichungen bzgl. der Vererbung des cpGenoms. Da jedoch andere Vertreter der Ordnung Fagales, der auch die Haselnuss zuzuordnen ist, wie z.B. die Stieleiche (*Quercus robur*) eine maternale Vererbung aufweisen^[92], ist davon auszugehen, dass dies auch bei der Haselnuss der Fall ist. Bezüglich seiner Mutationsrate liegt das cpGenom zwischen dem mtGenom und dem Karyom. Es wird im Durchschnitt von $1-3 \times 10^{-9}$ Basen pro Jahr ausgegangen.^[93] Damit liegt die Mutationsrate in Abhängigkeit der verwendeten Berechnungsmethode um das drei- bis vierfache unterhalb der des Karyoms und um das dreifache über der des Mitochondriengenoms.^{[92],[93]}

Im Falle der Haselnuss besitzt das Plastom bzw. cpGenom eine Länge von ca. 160 kbp.^{[94],[95]} Jeweils 26-27 kbp sind dabei den IR-Bereichen zuzuordnen. Unterteilt werden diese durch den etwa 18 kbp langen SSR-Bereich. Die verbliebenen 88-90 kbp bilden die LSC-Region.^{[94],[95]} Insgesamt sind etwa 130 Gene auf dem cpGenom der Haselnuss kodiert.^{[94],[95]} Der G/C-Gehalt liegt über das gesamte cpGenom betrachtet bei ca. 36,5%. Im Bereich der IRs ist er jedoch im Vergleich zu der LSC- und SSC-Region mit 42,4% erhöht.^[95] Dies folgt aus den überwiegend in den IR-Bereichen kodierten G/C-reichen rRNA- und tRNA-Genen.^[95]

4.3 Molekularbiologische Methoden

4.3.1 Polymerase chain reaction (PCR)

Die Polymerase-Kettenreaktion (engl. *polymerase chain reaction*, PCR) wurde 1983 von Kary Mullis entwickelt. Die erste Publikation zu dieser Technik folgte 1985 von Saiki et al., die mit Hilfe der PCR einen spezifischen Sequenzabschnitt des β -Globin im Rahmen der Diagnose der Sichelzellenanämie amplifizierten.^[96] 1993 wurde Kary Mullis schließlich mit dem Nobelpreis für Chemie für die Entwicklung der PCR ausgezeichnet.^[97] Er schuf damit den Grundstein vieler moderner molekularbiologischer Methoden. Es handelt sich bei der PCR um ein Verfahren zur Vervielfältigung eines definierten DNA-Abschnitts.^{[98],[99]}

Ein Reaktionsansatz für die PCR beinhaltet Wasser, einen enzymespezifischen Puffer, das Template (= die zu amplifizierende DNA), Primer, eine DNA Polymerase und dNTPs (Desoxynucleosidtriphosphate).^{[36],[100]} Der Puffer beinhaltet alle für die DNA Polymerase notwendigen Substanzen und sorgt damit für optimale Reaktionsbedingungen.^[100] Die Zusammensetzung kann abhängig vom Hersteller variieren, da sie jeweils an die Enzyme spezifisch angepasst werden. Ein unabdingbarer Bestandteil stellt das $MgCl_2$ dar. Dieses dient der Polymerase als Cofaktor.^{[101],[102]} Abhängig von der eingesetzten DNA Polymerase wird manchmal auch $MgSO_4$ anstelle des $MgCl_2$ eingesetzt.^[100] Die Mg^{2+} -Ionen beeinflussen die Enzymaktivität und bilden mit den dNTPs einen Komplex, der letztendlich die Form darstellt, in der die dNTPs von der DNA Polymerase in den DNA-Strang eingebaut werden können.^{[99],[100],[102]} Die Konzentration liegt im Bereich von 1-3 mM pro Reaktionsansatz.^[99] Bei Primern handelt es sich um 20-35 bp lange Oligonucleotide. Bei der Standard-PCR wird jeweils ein Primerpaar eingesetzt, das sich aus dem Vorwärts- und Rückwärts-Primer zusammensetzt. Jeweils ein Primer lagert sich pro DNA-Einzelstrang am jeweiligen 3'-Ende des Ziel-Fragments, das amplifiziert werden soll, an.^{[100],[103]} Für die Amplifikation wurde ursprünglich das Klenow-Fragment der *E.coli* DNA Polymerase I eingesetzt.^[104] Diese wies jedoch den großen Nachteil auf, dass sie thermolabil ist und deshalb nach jedem Denaturierungsschritt erneut zugegeben werden musste.^[104] Eine deutliche Erleichterung brachte daher die Etablierung der sogenannten *Taq* DNA Polymerase, die aus dem Bakterienstamm *Thermus aquaticus* isoliert wurde, welches natürlich in Geysiren und heißen Quellen zu finden ist und daher thermostabile Enzyme aufweist.^{[103]-[105]} Es handelt sich bei dieser DNA Polymerase auch heute noch um die am weitesten verbreitete und am häufigsten eingesetzte Polymerase im Bereich der PCR.^[100] Die meisten DNA Polymerasen besitzen eine 5' \rightarrow 3'-Syntheserichtung und benötigen dsDNA inklusiv eines intakten 3'-OH-Endes.^{[100],[102]} Es gibt heute neben der *Taq* DNA Polymerase eine Vielzahl weiterer DNA Polymerasen, die sich u.a. in ihren Synthesegeschwindigkeiten und Genauigkeiten unterscheiden.^[100]

Grundsätzlich basiert die Methode der PCR auf der zyklischen Wiederholung dreier Teilschritte: Denaturierung, Annealing und Elongation.^{[36],[46],[98],[100]} In Abbildung 5 ist ein PCR-Zyklus schematisch dargestellt. Im Folgenden werden die Teilschritte näher erläutert.

Denaturierung:

Bei der Denaturierung wird der Reaktionsansatz auf $> 90\text{ }^{\circ}\text{C}$ erhitzt.^[99] Dabei zerlegt sich der DNA-Doppelstrang in seine Einzelstränge. Initial wird ein verlängerter Denaturierungsschritt durchgeführt, da dies unerlässlich für das Gelingen der PCR ist. Die wiederkehrenden Denaturierungsschritte im Zuge der Wiederholungen sind i.d.R. deutlich kürzer.^[100]

Primer-Annealing:

Nach der Denaturierung erfolgt der Schritt des Annealings. Hierbei lagern sich zwei sequenzspezifische Oligonucleotide, sogenannte Primer, jeweils am 3'-Ende des zu amplifizierenden Abschnitts an.^[102] Die Temperatur des Reaktionsansatzes wird dabei auf die primerspezifische Annealingtemperatur reduziert. Diese wird durch die Länge und die Sequenz des Primers bestimmt.^[46]

Elongation:

Im dritten Schritt, der Elongation, vervielfältigt eine DNA Polymerase unter Verbrauch von Desoxynucleosidtriphosphaten (dNTPs) den Doppelstrang, indem sie am freien 3'-OH-Ende des Primers ansetzt und einen zum Einzelstrang komplementären Strang aufbaut. Die Temperatur des Reaktionsansatzes beträgt hierbei $72\text{ }^{\circ}\text{C}$.^[102] Damit wird in einem Zyklus jeder Doppelstrang verdoppelt.^[103]

Durch die mehrfache Wiederholung dieser drei Schritte bzw. eines PCR-Zyklus (i.d.R. 20-35 mal), erfolgt eine exponentielle Vervielfältigung des gewünschten spezifischen DNA-Abschnitts.^{[99],[104]} Durch den bereits beschriebenen Einsatz thermostabiler Enzyme kombiniert mit programmierbaren Thermocyclern kann die PCR in einem Reaktionsgefäß vollautomatisch durchgeführt werden.^[102]

Es gibt heutzutage eine Vielzahl unterschiedlicher Techniken und Methoden, die auf dem Grundprinzip der PCR aufbauen. Einige dieser Methoden können nicht nur für qualitative, sondern auch quantitative Analysen genutzt werden. Auch im Bereich der Klonierung und der Sequenzierung ist die PCR ein sehr wichtiger Bestandteil. In vielen Fällen fungiert jedoch weiterhin die „klassische“ PCR als Vorstufe zur Vervielfältigung eines spezifischen DNA-Abschnitts für weiterführende Analysen.^[100]

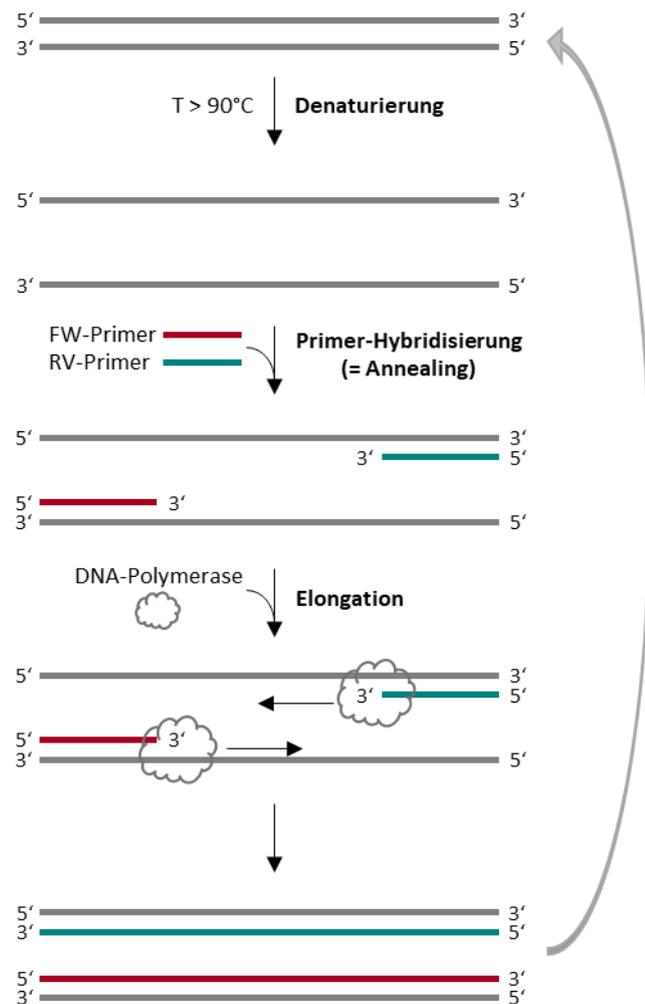


Abbildung 5: Schematischer Ablauf eines PCR-Zyklus (modifiziert nach ^[102]).

4.3.2 Detektion von PCR-Produkten

4.3.2.1 Agarosegelelektrophorese (AGE)

Bei der AGE handelt es sich um ein elektrophoretisches Trennverfahren, um Moleküle unterschiedlicher Größe voneinander zu trennen. Die Elektrophorese wurde bereits 1937 von Arne Tiselius entwickelt.^{[106],[107]} Sie beschreibt grundlegend die Wanderung geladener Teilchen in einem elektrischen Feld abhängig von ihrer Ladung und dem Molekulargewicht.^[36]

Nukleinsäure besitzt aufgrund ihres geladenen Zucker-Phosphat-Rückgrates eine negative Ladung. Da sich diese proportional zu der Länge des DNA-Stranges verhält, erfolgt eine Trennung nach der molekularen Masse und damit der Länge von DNA-Fragmenten.^{[102],[108]} Für die Analyse von Nukleinsäure wird i.d.R. Agarose als Trennmaterial eingesetzt. Die Verwendung von Agarose in diesem Kontext beschrieb Stellan Hjertén erst 1961.^[109] Es handelt sich hierbei um ein Polysaccharid, welches aus β -D-Galactopyranose und 3,6-Anhydro- α -L-galactopyranose aufgebaut ist.^[110] Wird Agarose mit einem Puffer aufgekocht, geliert es anschließend beim Abkühlen und bildet aufgrund von Wasserstoffbrückenbindungen ein feines Netzwerk aus. Abhängig von der gewählten Agarose-Konzentration (i.d.R. 0,5-3%) kann der Porendurchmesser variiert werden. Mit zunehmender Konzentration nehmen Porendurchmesser und der Trennbereich ab.^[102]

Die Elektrophorese selbst erfolgt in einer Elektrophoresekammer, welche mit einem Puffer befüllt ist in den das ausgehärtete Gel eingelegt wird. Bevor die DNA in die Probenaschen pipettiert wird, wird sie mit *Loading Dye* versetzt. Hierbei handelt es sich um eine Mischung aus Puffer, Glycerin und Farbstoff. Aufgrund des hohen Anteils an Glycerin (ca. 50%) wird die Probe beschwert, wodurch sie nicht aus den Probenaschen herausfließen kann. Die Farbstoffe (Xylencyanol und/oder Bromphenolblau) ermöglichen eine Visualisierung während des Pipettierens und eine Nachverfolgbarkeit des Fortschritts der Elektrophorese. Durch Anlegen einer Gleichspannung an der Elektrophoresekammer wandert die negativ geladene DNA in Richtung der positiv geladenen Anode. Nach beendeter Auftrennung wird die DNA durch Verwendung eines Fluoreszenzfarbstoffes unter UV-Licht detektiert.^[102] Diese Farbstoffe interkalieren in die doppelsträngige DNA oder binden an den kleinen Furchen der DNA und werden unter UV-Licht zur Fluoreszenz angeregt wodurch die DNA sichtbar wird.^[111] Durch Verwendung eines Molekulargewichtsmarkers kann die Länge der erhaltenen Fragmente abgeschätzt werden.^[102] Der Ablauf ist schematische in Abbildung 6 dargestellt.

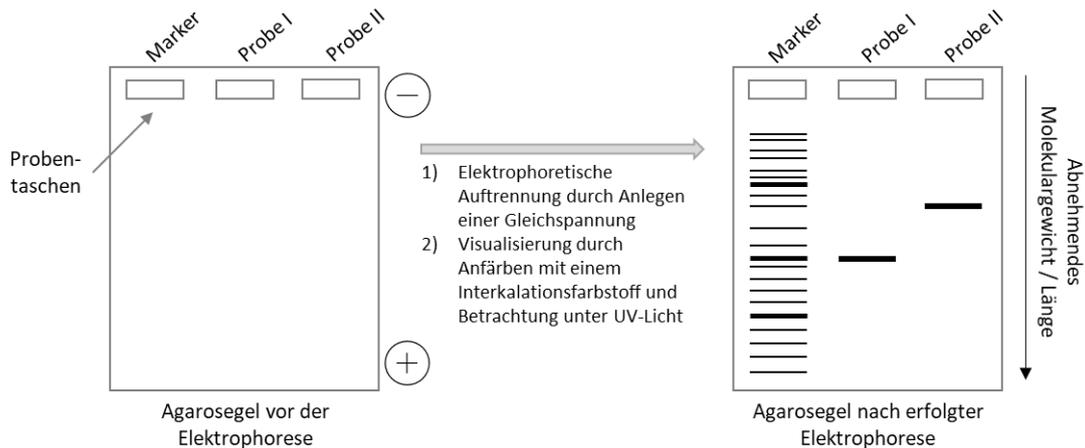


Abbildung 6: Grafische Darstellung einer elektrophoretischen Auftrennung von DNA-Fragmenten mittels AGE.

Ethidiumbromid ist einer der am weitesten verbreiteten Farbstoffe zum Anfärben von DNA. Seine Verwendung wurde bereits in den 70er Jahren beschrieben.^{[112],[113]} Da bei diesem Stoff jedoch von einem genverändernden Potenzial ausgegangen wird^[114], werden heutzutage oft alternative Farbstoffe wie GelRed® eingesetzt. Diese bestehen häufig aus zwei Monomeren bekannter Fluoreszenzfarbstoffen wie bspw. Ethidiumbromid, welche über sogenannte *bridges* oder *linker*, bestehend aus 8-150 Atomen miteinander verbunden sind.^[111] Dadurch bilden sich bei Abwesenheit von DNA *hairpin*-Strukturen, welche dazu führen, dass die Farbstoffe im Vergleich zu Ethidiumbromid eine deutlich verringerte Hintergrund-Fluoreszenz, interkaliert bzw. gebunden an einen DNA-Strang jedoch eine erhöhte Sensitivität aufweisen.^{[111],[115],[116]} Des Weiteren führen die modifizierten Strukturen nachweislich zu einer verringerten Membranpermeabilität und damit einem geringeren toxischen Potenzial.^{[115],[116]}

4.3.2.2 Kapillargelelektrophorese (CGE)

Bei der Kapillargelelektrophorese (CGE, engl. *capillary gel electrophoresis*) handelt es sich ebenfalls um ein Verfahren zur Trennung und Detektion von geladenen Teilchen basierend auf dem Verfahren der Elektrophorese. Die angewendete CGE gehört zu der Form der Mikrochip-CGE. Bereits 1994 demonstrierten Woolley und Mathies erstmals die Trennung von DNA-Molekülen mit Hilfe einer Mikrochip-Elektrophorese.^[117] Ihr Vorgehen damals unterscheidet sich nicht wesentlich von der Anwendung heute. Die Trennung erfolgte in feinen in einen Glaschip geätzten Kapillaren, welche mit einem Polymer befüllt wurden. Die zu trennende DNA wurde mit einem Fluoreszenzfarbstoff, der in dem eingesetzten Puffer enthalten war, markiert, wodurch eine Detektion nach erfolgter Trennung durch Anlegen einer Spannung mittels eines Lasers möglich war.^[117]

Für die in dieser Arbeit vorgestellten Anwendungen der CGE wurde der 2100 Bioanalyzer der Firma Agilent Technologies, Inc. (Santa Clara, USA) verwendet. Es handelt sich hierbei um eines der kommerziell erhältlichen Systeme für eine automatisierte Trennung und Detektion von Biomolekülen (DNA, RNA und Proteine) basierend auf der Mikrochip-Technologie. Abbildung 7 zeigt beispielhaft

einen Chip zur Trennung von DNA-Fragmenten. Links (A) ist der gesamte Chip abgebildet, welcher in den 2100 Bioanalyzer eingesetzt wird, und rechts (B) die Unterseite des Mikrochips aus Glas mit den eingezätzten Kapillaren in denen die Trennung stattfindet.

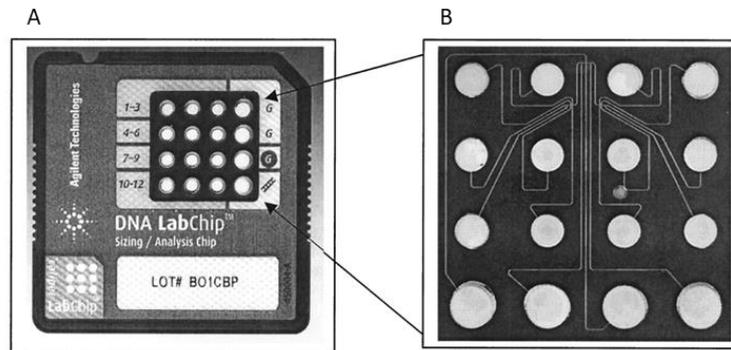


Abbildung 7: DNA LabChip™ für die Analyse von DNA mit Hilfe des 2100 Bioanalyzers.^[118]

Die Gelmatrix befindet sich hierbei in dünnen, in den Mikrochip geätzten Kanälen. Dadurch ist die Oberfläche im Vergleich zu dem Gelvolumen sehr groß, weshalb die bei der Elektrophorese entstehende Wärme schneller und leichter abgeführt werden kann. Dies ermöglicht eine Auftrennung bei einer deutlich höheren Spannung und verkürzt damit die benötigte Analysenzeit auf ca. drei Minuten pro Probe.^[119] Der Plastikchip, in welchen der Mikrochip eingefasst ist, weist auf der Vorderseite die sogenannten *wells* auf, in welche die Gelmatrix, bzw. die Proben oder die zugehörige DNA-Leiter hinein pipettiert werden. Durch Druck wird mit Hilfe einer sogenannten *Priming station* die mit einem Fluoreszenzfarbstoff versetzte Gelmatrix vor der Benutzung in die Kapillaren hineingepresst. Die Detektion erfolgt mittels eines Lasers und des Interkalationsfarbstoffes, der gebunden an die DNA zur Fluoreszenz angeregt wird und diese dadurch detektierbar macht.^[118] Insgesamt weist dieses System im Vergleich zu der herkömmlichen AGE einige Vorteile auf. Es stellt eine Möglichkeit einer schnellen, automatisierten *high-throughput* Analyse dar, welche zudem nur eine sehr geringe Probenmenge benötigt.^[119] So kann mit dem vorgestellten System von Agilent ein Chip mit 12 Proben in nur ca. 30 Minuten analysiert werden und das benötigte Probenvolumen liegt bei 1 μL mit Konzentrationen von lediglich 5 $\text{pg}/\mu\text{L}$ bis 50 $\text{ng}/\mu\text{L}$ abhängig vom verwendeten Kit.^{[118],[120],[121]} Zudem besteht die Möglichkeit einer deutlich präziseren Bestimmung der Fragmentlänge. Damit stellt sie eine gute und zum Teil unabdingbare Alternative zur AGE dar, da sie zusätzlich auch die Möglichkeit einer Quantifizierung bietet.^[122] Des Weiteren wurde der 2100 Bioanalyzer in der vorliegenden Arbeit auch für die Überprüfung der *library* im Zuge der Probenvorbereitung für die Sequenzierung mittels NGS herangezogen.^[123]

4.3.3 Sequenzierverfahren

4.3.3.1 Sanger-Sequenzierung

Die Sanger-Sequenzierung, auch bekannt unter der Kettenabbruchmethode oder der Didesoxymethode wurde in den 1970er Jahren von Sanger und Coulson entwickelt.^{[46],[124]} Die erste vollständige Sequenzierung eines Genoms (Bakteriophage ϕ X174) basierend auf dieser Methode wurde 1977 vorgestellt.^{[125],[126]}

Die Methode beruht dabei auf den Grundlagen der PCR. Der zu sequenzierende DNA-Abschnitt wird zunächst durch Erhitzen denaturiert. Anschließend lagert sich ein Primer (Sequenzierprimer) an den Einzelstrang an und eine Polymerase verlängert den Primer, beginnend am 3'-Ende, und synthetisiert den komplementären DNA-Strang. Im Gegensatz zur herkömmlichen PCR enthält der Reaktionsansatz neben den Desoxynucleotiden (dNTPs) zusätzlich Didesoxynucleotide (ddNTPs).^{[36],[46],[126]} Diesen fehlt die Hydroxylgruppe am 3'-Ende des Zuckerbausteins, weshalb nach Einbau eines ddNTPs in den DNA-Strang die Synthese abgebrochen wird. Dies führt dazu, dass Teilfragmente der Sequenz gebildet werden.^[127] Ursprünglich wurde der Reaktionsansatz in vier Aliquote unterteilt und jeweils eines der ddNTPs (ddATP, ddTTP, ddCTP, ddGTP) hinzugegeben.^[126] Jeder Reaktionsansatz enthielt nach abgeschlossener PCR alle möglichen Fragmente, die mit der jeweiligen Base des entsprechenden ddNTPs enden.^[46] Anschließend wurden die Fragmente mittels hochauflösender Polyacrylamidgelelektrophorese (PAGE) aufgetrennt. Ein Anfärben der DNA-Banden mit einem Interkalationsfarbstoff wie Ethidiumbromid ist in diesem Fall nicht möglich, da die DNA-Menge pro Bande zu gering ist.^[46] Zu Beginn erfolgte die Detektion mittels Autoradiographie.^{[46],[128]} Hierbei wurden radioaktiv markierte Primern, dNTPs bzw. ddNTPs eingesetzt.^{[46],[126]} Bei den eingesetzten Isotopen handelte es sich zunächst um ^{32}P und später auch um ^{33}P oder ^{35}S .^[128] Nach erfolgter Trennung der DNA-Fragmente wurde über Nacht ein Röntgenfilm auf das Polyacrylamidgel gelegt. Aufgrund der radioaktiven Strahlung wurden die Positionen der Fragmente auf dem Film geschwärzt, sodass sich nach Entwicklung ein Bandenmuster abbildete.^[46] Der große Nachteil dieses Verfahrens ist die radioaktive Strahlung der Reagenzien.^[129] Neben der radioaktiven Markierung wurde im Laufe der Zeit eine alternative Möglichkeit der Markierung entwickelt. Diese basiert auf dem Einsatz von Fluoreszenzfarbstoffen.^[128] Hierbei können ebenfalls die Primer oder ddNTPs markiert werden. Von einer Markierung der dNTPs wird in der Regel abgesehen, da die Fluoreszenzfarbstoffe aufgrund ihrer Größe das Laufverhalten der DNA-Fragmente bei der anschließenden elektrophoretischen Trennung beeinflussen würden.^[46] Heutzutage werden i.d.R. vier unterschiedliche Fluoreszenzfarbstoffe eingesetzt, mit denen die vier verschiedenen ddNTPs markiert werden.^{[128],[130]} Damit ist es möglich, die Reaktion in nur einem Ansatz durchzuführen. Die anschließende Trennung der DNA-Fragmente erfolgt heutzutage über eine hochauflösende Kapillargelelektrophorese (CGE).^[128] Am Ende der Kapillare befindet sich ein Laser, der die Fluoreszenzfarbstoffe anregt.^{[36],[46],[128]} Die anschließend emittierte farbstoff- und damit basenspezifische Strahlung wird von einem Detektor erfasst und grafisch als Elektropherogramm für die weitere Auswertung ausgegeben. Die Funktionsweise der modernen Sanger-Sequenzierung

mittels fluoreszenzmarkierten ddNTPs und Auftrennung mittels CGE ist in Abbildung 8 schematisch dargestellt.

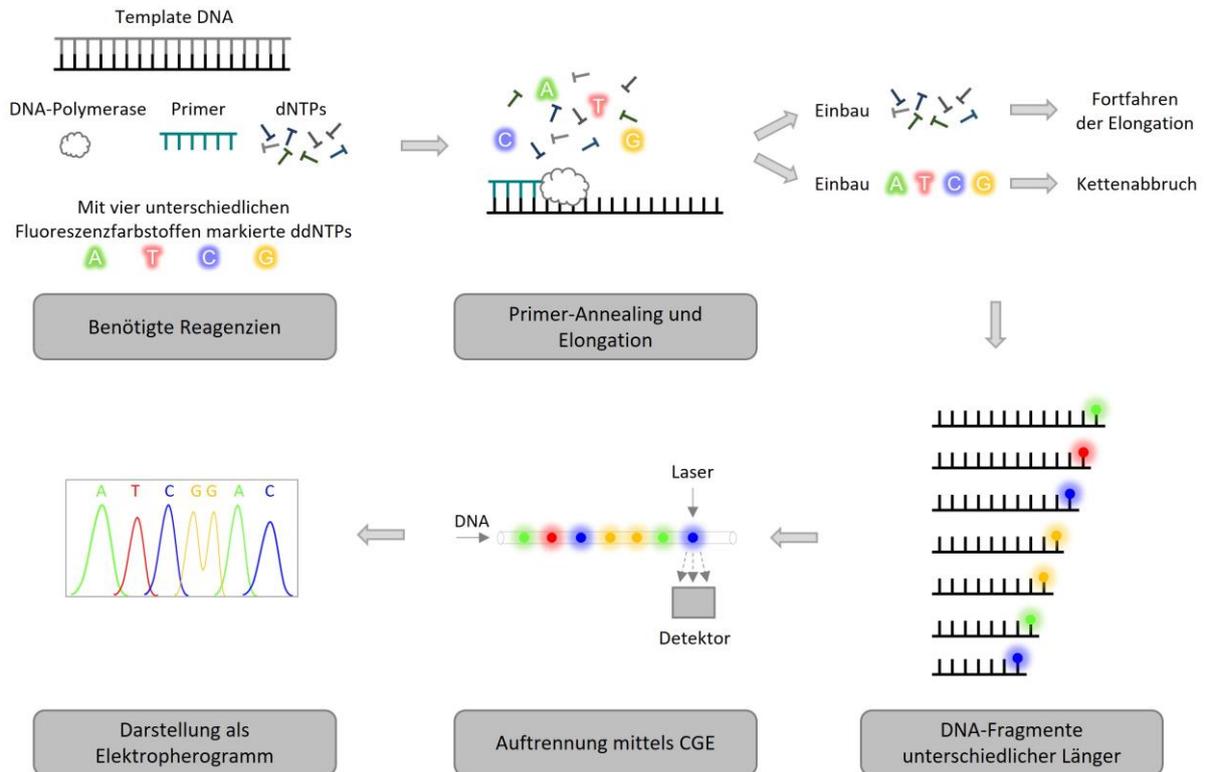


Abbildung 8: Funktionsweise der Sanger-Sequenzierung.

Die Sanger-Sequenzierung wurde in ihrer optimierten Form mit einer Leselänge von 750-1100 bp^{[131],[132]} zum Goldstandard der *high-throughput* Sequenzbestimmung.^[130] Sie gilt als erste Generation der automatisierten DNA-Sequenzierung.^[128]

4.3.3.2 Next-Generation Sequencing

Die Sanger-Sequenzierung prägte über zwei Jahrzehnte hinweg die Genomanalyse und trug zum Gelingen großer Projekte, wie dem Humangenom-Projekt bei. Trotz technischer Fortschritte kam diese Technologie auch aufgrund sich verändernder Anforderungen an ihre Grenzen.^[133] Zum einen ist die Leselänge, aber auch der generierte Datenoutput pro Lauf begrenzt. Zudem ist das Verfahren unter Berücksichtigung der generierten Datenmenge sehr kostenintensiv.^{[131],[133]} So folgte Anfang der 2000er die Entwicklung eine Reihe neuer Technologien der DNA-Sequenzierung. Diese können in die 2., 3. und 4. Generation unterteilt werden. Häufig werden sie jedoch einfach unter *Next-Generation Sequencing*, kurz NGS, zusammengefasst. Eine Gemeinsamkeit all dieser neuen Verfahren ist der deutlich höhere Datenoutput sowie die Tatsache, dass verschiedenste Abschnitte, aber auch ein Fragment in mehrfacher Kopie parallel sequenziert werden können, weshalb häufig auch von „*massively parallel sequencing*“ gesprochen wird.^[134]

Die Fragmentierung langer DNA-Abschnitte wird im Zuge der Probenvorbereitung entweder enzymatisch oder mechanisch durchgeführt. Sie erfolgt ungerichtet, um eine statistische Verteilung sich überlappender Fragmente zu erhalten, die im Anschluss an die Sequenzierung zu einer consensus-Sequenz zusammengefügt werden können. Dieser Ansatz wird auch Schrottschuss-Sequenzierung (engl. *shotgun sequencing*) genannt.^{[102],[135]} Die weitere Probenvorbereitung kann im Wesentlichen nach zwei verschiedenen Vorgehensweisen erfolgen. Zum einen besteht die Möglichkeit der *clonally amplified templates* und zum anderen die des *single DNA-molecule templates*.^[133] Im Fall der klonal amplifizierten Templates werden die DNA-Fragmente vor der Sequenzierung mittels PCR vervielfältigt. Dies kann entweder mittels Emulsions-PCR oder einer Festphasen-PCR erfolgen.^[133] Bei dem Einsatz eines *single-DNA molecule templates* hingegen findet im Vorfeld der eigentlichen Sequenzierung keine Vervielfältigung statt. Dies hat den Vorteil, dass die PCR als potentielle Fehlerquelle für den Einbau einer falschen Base ausgeschlossen wird. Des Weiteren wird dadurch auch die Diskriminierung von AT- bzw. GC-reichen Fragmenten während der PCR und damit einhergehenden verringerten Ausbeuten in Bezug auf den Datenoutput ausgeschlossen.^[133] Im Zuge der Probenvorbereitung werden die DNA-Fragmente lediglich mit Adaptern an den Enden versehen. Vor der Sequenzierung werden die Fragmente i.d.R. immobilisiert. Hierbei besteht die Möglichkeit entweder die zu sequenzierenden Fragmente zu immobilisieren oder die für die DNA-Sequenzierung benötigte DNA Polymerase.^[133]

Die Verfahren der 2. und 3. Generation arbeiten grundsätzlich bei der Sequenzierung nach zwei Prinzipien: *sequencing by synthesis* (SBS) oder *sequencing by ligation* (SBL).^[136] Der Bereich SBS kann noch in folgende drei Kategorien unterteilt werden: *cyclic reversible termination* (CRT), *single-nucleotide addition* (SNA) und *real-time sequencing*.^[133]

Ein Verfahren der 2. Generation, das die Probenvorbereitung der *clonally amplified templates* nutzt und mit dem Prinzip der *sequencing by synthesis* arbeitet ist die Plattform von Illumina Inc. (San Diego, USA). Es erfolgt dabei eine Vervielfältigung der DNA-Fragmente mittels isothermaler Brückenamplifikation, die als Festphasen-PCR auf der sogenannten *flow cell* durchgeführt wird.^{[133],[137]} Dadurch entstehen Cluster bestehenden aus jeweils mehreren hundert Kopien eines jeden Fragments.^{[138],[139]} Die anschließende Sequenzierung erfolgt nach der CRT-Methode (*cyclic reversible termination*). Wie der Name bereits andeutet, wird zyklisch vorgegangen, wobei jeder Sequenzierungszyklus aus drei Schritten besteht: (1) Einbau eines dNTPs, (2) Detektion/Visualisierung des Fluoreszenzsignals und damit Identifizierung der eingebauten Base, (3) Abspaltung des Terminators und des Fluorophors. Zum Einsatz kommen modifizierte dNTPs. Sie weisen einen reversiblen Terminator an der 3'-OH-Gruppe des Zuckerbausteins auf, der verhindert, dass pro Zyklus mehr als eine Base eingebaut werden kann. Bei der Schutzgruppe handelt es sich um eine Azidomethyl-Gruppe.^[137] Zur Unterscheidung der vier Basen werden vier unterschiedliche Fluoreszenzfarbstoffe eingesetzt, die über Verbindungsstrukturen an die Basen gebunden sind.^{[133],[137]} Zu Beginn eines Zyklus werden alle vier unterschiedlichen dNTPs zur Verfügung gestellt und die jeweils komplementären Basen durch die Polymerase eingebaut. Durch den Terminator kann pro Zyklus jeweils nur eine einzige Base eingebaut werden. Anschließend werden die verbliebenen dNTPs entfernt und die eingebauten dNTPs über die Fluoreszenz der Farbstoffe mittels TIRF (*total*

internal reflection fluorescence) Mikroskopie optisch detektiert. Im letzten Schritt werden die Fluoreszenzfarbstoffe und Terminatoren entfernt, sodass im darauffolgenden Zyklus wieder eine neue Base eingebaut werden kann.^[140] Die Readlängen liegen bei diesem Verfahren je nach Ausführung bei max. 300 bp.^[139]

Neben Illumina Inc. (San Diego, USA) wurden auch von Roche (Basel, Schweiz), Helicos (Cambridge, USA) und Life Technologies™ (Carlsbad, USA) Verfahren entwickelt, die der zweiten Generation zuzuordnen sind und ebenfalls mit einer optischen Detektion arbeiten.^[136] Der Antrieb für die Entwicklung der Geräte der dritten Generation war schließlich v.a. die Reduzierung der Kosten pro Sequenzierung. Da u.a. die optischen Detektionseinheiten für die Kosten verantwortlich waren, wurden Verfahren entwickelt, die auf alternativen Detektionsmethoden beruhen.^[136] Ein Beispiel ist die Ion Torrent NGS der Firma Thermo Fisher Scientific™ (Waltham, USA). Diese NGS-Plattform arbeitet ebenfalls mit dem SBS-Prinzip kombiniert mit einer Form der Halbleitersequenzierung.^[136] Die dNTPs weisen in diesem Fall keine Blockierung am 3'-Ende auf, da hier nach dem SNA-Verfahren (*single-nucleotide added*) gearbeitet wird und deshalb pro Zyklus nur eine Art von dNTP zugegeben wird.^{[140],[141]} Die Detektion erfolgt in diesem Fall mittels ISFET (*Ion-Sensitive Field Effect Transistor*) über die pH-Wert-Änderung, die durch das freiwerdende H⁺-Ion beim Einbau eines dNTPs durch eine Polymerase auftritt. Im Zuge der Probenvorbereitung wird die DNA ebenfalls fragmentiert und Adapter ligiert. Anschließend erfolgt eine Vervielfältigung der zu sequenzierenden DNA-Fragmente mittels Emulsions-PCR.^{[100],[141]} Dadurch werden Beads erhalten, die jeweils eine tausendfache Kopie von nur einem DNA-Fragment aufweisen. Die Beads werden im nächsten Schritt auf den Sequenzierchip gegeben, wobei pro *well* jeweils nur ein Bead, also ein DNA-Fragment vorliegt.^[141] Die Sequenzierung der einzelnen Fragmente findet also räumlich getrennt voneinander statt.^[100] Während des Sequenziervorgangs werden die unterschiedlichen dNTPs zeitlich getrennt voneinander zugegeben. Erfolgt der Einbau einer komplementären Base, wird ein Proton freigesetzt. Dadurch sinkt der pH-Wert um ca. 0,02 Units ab.^[141] Diese Reduzierung des pH-Werts wird durch einen Sensor am Boden des *wells* detektiert. Werden aufgrund eines homopolymeren Bereichs mehrere dNTPs eingebaut, so wird ein entsprechend proportional höherer Abfall des pH-Werts detektiert.^[141] Die Leselänge bei diesem Verfahren liegt im Bereich von 100-200 bp.^{[136],[141]} Ein Vorteil dieser Methode liegt darin, dass im Vergleich zur optischen Detektion höhere Sequenziergeschwindigkeiten erreicht werden können, da das Abspalten von Terminator und Fluoreszenzfarbstoff je Zyklus bei den Methoden von Illumina Inc. (San Diego, USA) viel Zeit kostet.^[100] Zudem sind keine modifizierten dNTPs notwendig.^[136] Nachteilig ist jedoch die kurze Leselänge sowie die vielen wiederkehrenden Waschschriffe, die zu einer Akkumulation von Fehlern führen könne. Des Weiteren weist diese Methode Schwierigkeiten bei der Sequenzierung von homopolymeren oder hoch repetitiven Bereichen auf.^[136]

Ein weiteres Verfahren der 3. Generation, aber aus dem Bereich der *Long-read* Sequenzierung, das v.a. für die Erstellung von Scaffolds herangezogen wird, ist das System von Pacific Biosciences, meist kurz als PacBio bezeichnet. Diese Technologie verfolgt den *single-molecule real-time* (SMRT) Ansatz.^{[136],[140]} Zum Einsatz kommt bei diesem Verfahren eine spezielle *flow cell* mit vielen kleinen *wells* im Picoliter-Bereich.^[140] Der Boden ist transparent und weist *zero-mode waveguide* (ZMW)

Nanostrukturen auf, die eine optische Detektion einzelner Fluoreszenz-Moleküle ermöglichen.^{[142],[143]} Genau umgekehrt im Vergleich zu *short-read*-Verfahren wird in diesem Fall jedoch nicht die DNA gebunden, sondern die DNA Polymerase ist am Boden der *wells* immobilisiert.^[140] Bei der Sequenzierung kommen fluoreszenzmarkierte dNTPs zum Einsatz. Wird eine Base am Boden eines *wells* durch die DNA Polymerase eingebaut, so wird das emittierte Licht des entsprechenden Fluoreszenzfarbstoffs durch ein Laser- und Kamera-System detektiert. Anschließend wird die Bindung zu dem Fluorophor durch die Polymerase gespalten, wodurch es sich räumlich trennt. Es handelt sich also ebenfalls um einen SBS-Ansatz mit optischer Detektion.^{[136],[140]} Eine weitere Besonderheit dieses Systems ist das Template, das in zirkulärer Form eingesetzt wird. Das sogenannte „SMRTbell template“ wird dadurch generiert, dass nach der Fragmentierung universelle *hairpin loops* an beide Enden des DNA-Fragments ligiert werden. Dadurch entsteht ein zirkuläres Template bei dem beide Stränge des DNA-Doppelstrangs sequenziert werden können. Eine weitere Amplifikation der DNA vor dem Sequenziervorgang ist nicht notwendig. Daher zählt dieses Verfahren auch zu dem Bereich des *single-molecule sequencing*. Dies bringt zwei Vorteile mit sich; zum einen wird die Zeit der Probenvorbereitung verkürzt und zum anderen die Genauigkeit der Sequenzierung erhöht.^[136] Aufgrund der zirkulären Form des Templates ist auch eine mehrfache Sequenzierung möglich. Wie oft das Template sequenziert werden kann, ist jedoch von dessen Länge abhängig. So gestaltet sich eine mehrfache Sequenzierung bei einer Länge von 3 kbp bereits sehr schwer, je kürzer desto öfter kann das Template sequenziert werden.^[140] Der wahrscheinlich größte Vorteil der PacBio-Technologie liegt in der hohen Leselänge, die bei modernen Geräten im Bereich von mehreren kbp liegt. Einen Nachteil stellt jedoch die vergleichsweise geringe Kapazität dar.^{[136],[140]}

Eine weitere Technologie, die ebenfalls auf dem Prinzip des *long-read single-molecule sequencing* beruht, jedoch aufgrund ihres Fortschritts bereits als Technologie der 4. Generation bezeichnet wird, ist die Nanopore-Sequenzierung.^{[136],[144],[145]} Bei diesem Verfahren werden für die Detektion Nanoporen eingesetzt, die in eine (synthetische) Membran eingebettet sind. Es handelt sich dabei um Proteine wie bspw. α -Hämolyisin. Dieses wurde bereits als eines der ersten Modell-Proteine für die Entwicklung dieser Technologie herangezogen. Der Durchmesser einer Pore entspricht etwa der Größe eines einzelsträngigen DNA-Moleküls.^[144] Die Detektion beruht auf einem konstanten Ionenstrom, der verändert wird, sobald ein Molekül die Pore passiert. Diese Veränderung wird detektiert und in die entsprechende Basenabfolge übersetzt.^{[144],[146]} Einer der bekanntesten Hersteller für Geräte basierend auf dieser Technologie ist Oxford Nanopore Technologies (Oxford, UK). Hierbei bildet α -Hämolyisin die Pore gekoppelt mit einem zusätzlichen Motorprotein. Im Zuge der Probenvorbereitung wird die DNA fragmentiert und anschließend zwei unterschiedliche Adapter (*Leader* und *Hairpin*) ligiert, wobei jedoch nicht gezielt gesteuert wird, welcher Adapter an welches Ende des DNA-Strangs gebunden wird. Im Optimalfall weist ein Ende den *Leader*-Adapter und das andere den *Hairpin*-Adapter auf.^[140] Der *Leader*-Adapter hat dabei die Funktion das DNA-Fragment Richtung Membranoberfläche und in die Pore zu leiten. Der *Hairpin*-Adapter ermöglicht die Sequenzierung des komplementären Strangs direkt im Anschluss an die Sequenzierung des ersten Einzelstrangs.^[140] Mit dieser Methode können Leselängen von bis zu 4 Mbp und ein Datenoutput von über 400 Basen pro Sekunde erreicht werden.^[147] Ebenso ist im Gegensatz zu anderen NGS-Verfahren

auch die Detektion von Modifikationen der Basen möglich.^[146] Die Methode weist zudem ebenfalls den Vorteil anderer *single-molecule* Sequenzierungen auf, dass vor der Sequenzierung keine Amplifikation notwendig ist. Des Weiteren gibt es sehr kleine transportable Geräte mit einer Größe von nur ca. 3x10 cm wie bspw. den MinION von Oxford Nanopore Technologies (Oxford, UK), die an jeden herkömmlichen Computer angeschlossen werden können. Zusammen mit der einfachen und wenig aufwendigen Probenvorbereitung ist diese Technologie auch für die Analytik vor Ort einsetzbar. Zudem kann in Echtzeit auf die generierten Daten zugegriffen und diese ausgewertet werden.^{[138],[140],[148]}

Bei den hier vorgestellten Methoden handelt es sich nur um eine Auswahl der wichtigsten Verfahren. Alle auf dem Markt verfügbaren NGS-Plattformen unterscheiden sich in Bezug auf ihre Fehlerraten, Leselängen, Read-Typen, Kapazitäten sowie Anschaffungs- und Unterhaltungskosten und den Kosten pro Experiment.^{[102],[133],[149]} Im Anschluss an oder während jeder Sequenzierung müssen die erhaltenen Daten ausgewertet werden. Hierfür werden zunächst die Adaptersequenzen entfernt. Anschließend werden die Reads basierend auf überlappenden Sequenzen zu Contigs (*contiguous sequence*) zusammengesetzt. Bei einer ausreichend hohen Datenmenge und dem Vorhandensein eines Referenzgenoms kann im Optimalfall bereits aus den Contigs eine consensus-Sequenz gebildet werden. Der Vorgang des Zusammensetzens einzelner Reads unter zur Hilfenahme einer Vorlage, also eines Referenzgenoms, wird auch Assemblierung genannt. Ist die Erstellung einer consensus-Sequenz nicht möglich, werden die Contigs zunächst zu einem Gerüst, auch Scaffold genannt, zusammengesetzt. Anschließend können die Lücken z.B. durch andere Sequenziermethoden geschlossen werden, sodass am Ende eine vollständige consensus-Sequenz vorliegt.^[102] Für komplexe *de novo* Sequenzierungen werden daher häufig unterschiedliche Methoden miteinander kombiniert. So wird bspw. mittels PacBio oder Nanopore Sequenzierung ein Scaffold erstellt und die verbliebenen Lücken anschließend mittels *short-read* Sequenzierung wie der Illumina-Technologie geschlossen.

4.3.4 Klonierung

Bei dem Verfahren der Klonierung wird ein DNA-Fragment oder Gen (Insert) in einen Klonierungsvektor eingebracht und anschließend mit Hilfe einer geeigneten Wirtszelle vermehrt.^[46] Es handelt sich also um eine Neukombination und Vermehrung von DNA, die *in vitro* durchgeführt wird. Da DNA unterschiedlichen Ursprungs miteinander verbunden werden, wird auch von Rekombination und entsprechend rekombinanter DNA gesprochen.^[46] Als Vektoren werden allgemein DNA-Moleküle bezeichnet, die Fremd-DNA aufnehmen und in die Wirtszellen übertragen können. Es gibt verschiedene Arten von Vektoren: Plasmide, Phagen, Cosmide oder YACs. Sie unterscheiden sich hauptsächlich in ihrer Aufnahmekapazität, also nach der Größe des Inserts, das in die Vektoren eingebaut werden kann.^{[36],[46]} Im Folgenden wird auf Plasmide als Vektor, welche heute auch die am häufigsten verwendeten Klonierungsvektoren sind, näher eingegangen. Sie besitzen einen Replikationsursprung (*ori*) für die Vermehrung in der Wirtszelle und Selektionsmarker, auch

Selektionsgene oder Marker-Gene genannt, bei denen es sich meist um Resistenzgene z.B. gegen Ampicillin oder Kanamycin handelt. Diese dienen der Selektion rekombinanter Bakterienzellen.^{[46][36]} Zudem besitzen Plasmide i.d.R. Schnittstellen für Restriktionsenzyme, über welche die ringförmige DNA geöffnet und das Insert eingebaut werden kann. Häufig liegen mehrere verschiedene Schnittstellen für unterschiedliche Restriktionsendonukleasen vor, die sich direkt hintereinander auf dem Plasmid befinden. Solche Bereiche werden auch als multiple Klonierungsstelle (MCS, engl. *multiple cloning sites*) oder *Polylinker* bezeichnet.^{[36],[46]}

Für die Herstellung rekombinanter DNA wird zwischen der *Sticky-end*- und *Blunt-end*-Ligation unterschieden.^[100] Bei der *Blunt-end*-Ligation weisen sowohl das Insert, als auch der Vektor glatte Enden auf, die mittels einer Ligase kovalent verknüpft werden.^[100] Mögliche Überhänge werden hierbei mittels Restriktionsenzymen entfernt. Für eine erfolgreiche Ligation sollte hierfür jedoch die Viskosität des Reaktionsansatzes höher gewählt werden. Zudem muss das Insert ein phosphoryliertes 5'-Ende besitzen.^[100] Bei der *Sticky-end*-Ligation werden das Plasmid und das Insert mit demselben Restriktionsenzym geschnitten. Dieses führt aufgrund seiner Erkennungssequenz zu sogenannten *sticky ends*, die beiden Stränge des DNA-Doppelstrangs werden also versetzt geschnitten.^[150] Dadurch liegen bei dem Insert und dem Plasmid komplementäre Enden vor, über die sich das Insert und das Plasmid aneinander anlagern und anschließend mittels einer Ligase verbunden werden können.^{[46],[100]} In diesen Bereich fällt auch die sogenannte T/A-Klonierung. Der linearisierte Vektor besitzt hierbei einen Thymidin-Überhang. Solch ein T/A-Klonierungsvektor kann synthetisiert werden, indem ein Vektor mittels eines *Blunt-end*-Restriktionsenzym geschnitten und anschließend durch Zugabe von Didesoxythymidintriphosphat (ddTTP) und einer terminalen Transferase ein Thymidin-Überhang angebaut wird.^[151] Alternativ kann das Plasmid auch mit einem geeigneten *Sticky-end*-Restriktionsenzym geschnitten werden, welches ebenfalls den gewünschten Überhang generiert. Voraussetzung für eine erfolgreiche Ligation mit einem T/A-Klonierungsvektor ist die Generierung des Inserts mit einer *Taq* Polymerase, da diese eine terminale Transferase-Aktivität besitzen und am 3'-Ende des Inserts einen Adenosin-Überhang bilden.^[152] Über die jeweiligen A/T-Überhänge können Insert und Vektor miteinander hybridisieren und anschließend mittels einer Ligase kovalent verknüpft werden.^[151] Dieses Verfahren ist deutlich effizienter im Vergleich zu einer *Blunt-end*-Ligation und bietet zudem den Vorteil, dass keine zusätzlichen enzymatischen Modifikationen des Inserts notwendig sind.^[153] Daher ist die Methode auch für die Klonierung von Inserts mit unbekannter Sequenz geeignet. Eine Variation der T/A-Klonierung ist die sogenannte TOPO-TA-Klonierung, bei der die Ligation anstelle einer Ligase mit einer Topoisomerase I erfolgt. Das Enzym erkennt die Schnittstelle 5'-YCCTT-3' des Vektors und schneidet *upstream* der Erkennungssequenz. Anschließend bindet die Topoisomerase I kovalent über eine 3'-Phosphotyrosyl-Bindung an den 3'-Enden des Vektors. Damit liegt der Vektor linearisiert vor und durch die kovalente Bindung zwischen Enzym und DNA wird eine erneute Selbstligation des Vektors verhindert. Steht nun ein Insert mit freiem 5'-Ende zur Verfügung, wird die kovalente Bindung übertragen und das Insert in den Vektor eingebaut.^[154]

Im Anschluss an die Ligation erfolgt die Transformation, also die Aufnahme des Vektors in eine Zelle. Als Wirtszellen dienen i.d.R. gram-negative Bakterien wie *Escherichia coli*. Bei diesen muss die

Kompetenz, also die Fähigkeit DNA aus der Umgebung aufnehmen zu können, künstlich induziert werden, da sie von Natur aus nicht kompetent sind.^[155] Die Transformation kann mit chemisch oder elektrokompenten Zellen durchgeführt werden. Bei der chemischen Transformation, die erstmals in den 1970er Jahren beschrieben wurde, werden die Bakterienzellen mit einer Salzlösung, wie bspw. CaCl_2 oder MgCl_2 und der rekombinanten DNA versetzt und auf Eis gekühlt. Dabei lagert sich die DNA außen an die Zellen an. Durch einen kurzzeitigen Hitzeschock auf 37-42 °C wird die Aufnahme der DNA induziert.^{[46],[156]–[159]} Bei der Elektroporation, erstmals in den 1980er Jahren von Neumann et al. beschrieben, werden die elektrokompenten Bakterienzellen einem elektrischen Schock (ca. 2000 V für einige Millisekunden) ausgesetzt.^{[160],[161]} Dadurch erfolgt eine kurzzeitige Änderung der Lipidanordnung der Membran und der Vektor kann in die Zelle aufgenommen werden.^[161] Für eine erfolgreiche Elektroporation muss darauf geachtet werden, dass die rekombinante DNA salzfrei ist, damit kein Kurzschluss verursacht wird und die Zellen absterben. Die Transformationseffizienz beträgt je nach Verfahren bis zu 10^9 Kolonien pro μg Plasmid-DNA.^[159] Nach einer kurzen Regeneration von etwa einer Stunde bei dem Temperaturoptimum für das Wachstum der Bakterien von 37 °C werden diese auf Nähragar ausplattiert. Dieser enthält, abhängig davon, welche Resistenzgene auf dem Plasmid vorhanden sind, ein entsprechendes Antibiotikum. Damit wachsen selektiv nur die Zellen, die erfolgreich einen Vektor aufgenommen haben und deshalb eine entsprechende Resistenz aufweisen.^[46] Eine Auswahl der auf dem Nähragar gewachsenen Kolonien werden einer *Colony*-PCR unterzogen. Mit Hilfe dieser wird überprüft, ob die Zellen ein Plasmid enthalten, welches auch das gewünschte Insert trägt.^[162] Nach erfolgreicher PCR werden die geeigneten Kolonien in Flüssigmedium überimpft, um ausreichend Zellmaterial zu gewinnen.^[46] Zuletzt erfolgt die Plasmidisolierung und anschließende Sequenzierung. Damit können Mischungen verschiedener DNA-Fragmente der gleichen Längen erfolgreich selektiert und sequenziert werden.

4.3.5 Analyse von *Start Codon Targeted* (SCoT) Polymorphismen

Collard und Mackill (2009) beschrieben erstmals die Untersuchung von *Start Codon Targeted* Polymorphismen als Grundlage phylogenetischer Untersuchungen.^[163] Es handelt sich hierbei um eine *non targeted*-Methode, da keine detaillierten Sequenzinformationen vorliegen müssen. Der Fokus liegt hierbei auf den Bereichen, die das Startcodon ATG flankieren.^[163] Dieses Basentriplett kodiert für das entsprechende Startcodon AUG auf der mRNA, welches den Beginn der Translation markiert und Teil der Kozak-Sequenz ist.^[164] Bei dieser handelt es sich um eine Konsensus-Sequenz des Bereichs, der das Startcodon flankiert. Er wurde nach der US-amerikanischen Biochemikerin Marilyn Kozak benannt.^[165] Verschiedene Untersuchungen zeigten, dass zwischen den Sequenzen unterschiedlicher Gruppen von Eukaryoten deutliche Unterschiede auftreten können.^{[165],[166]} Die Analyse der SCoT Polymorphismen richtet sich gezielt auf Mutationen in diesem Abschnitt.

Die Methode von Collard und Mackill (2009) beruht auf einer normalen Endpunkt-PCR, die mit nur einem Primer durchgeführt wird. Die Entwicklung der Primer erfolgte auf der Basis vorheriger Untersuchungen der Bereiche um das Startcodon herum von Joshi et al. (1997) und Sawant et al.

(1999).^{[166],[167]} Sie wurden so gewählt, dass die Sequenz des Startcodons jeweils ein Teil der Primer ist. Auf Grundlage der vorhandenen Sequenzinformationen wurden manche Nukleotide in der Basenabfolge fest definiert (siehe Tabelle 2).^[163]

Tabelle 2: Darstellung fest definierter Nukleotide bei der Entwicklung der SCoT-Primer.^[163] Das Startcodon ist grau hinterlegt.

Position	+1	+2	+3	+4	+5	+6	+7	+8	+9
Base	A	T	G	G	C	N	A	C	C

Anschließend wurden die Primer sowohl in die 5′-, als auch 3′-Richtung mit variablen Basen verlängert.^[163] Insgesamt weisen sie eine Länge von 18 Nukleotiden auf, da Primer mit einer Länge von 18-24 Basen eine bevorzugte Länge für die Entwicklung reproduzierbarer Marker darstellen.^[168] Zudem waren 18 Nukleotide das Maximum, welches aufgrund der Vielzahl konservierter Basen möglich war. Der GC-Gehalt liegt zwischen 50% und 72%.^[163] Die verschiedenen Primer unterscheiden sich um mindestens eine Base. Bevorzugt liegt der Basenaustausch innerhalb der letzten vier Basen am 3′-Ende, da sich dieser Teil der Primer an den nicht konservierten Bereich anlagert.^[163] Voraussetzung für die Entstehung eines Amplifikats, welches anschließend mittels AGE detektiert werden kann, ist die räumliche Nähe zweier Gene und daraus resultierender Primer-Bindungsstellen zueinander. Damit werden mit jedem Primer für die verschiedenen Sorten spezifische Bandenmuster erhalten, die mit einem universellen Fingerabdruck verglichen werden können. Die unterschiedlichen Banden werden auch als sogenannte SCoT-Marker bezeichnet. Abbildung 9 stellt schematisch die Funktionsweise der Methode dar.

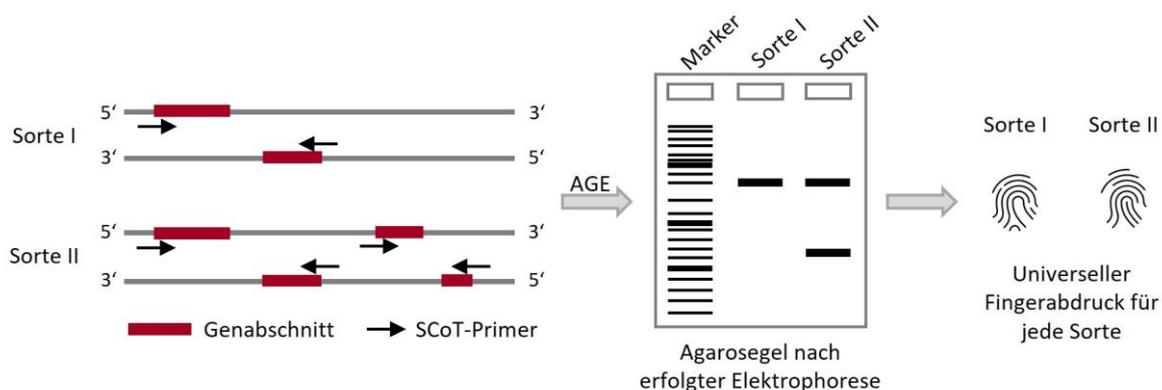


Abbildung 9: Schematische Darstellung der Funktionsweise der Analyse der SCoT Polymorphismen. Modifiziert nach Collard und Mackill (2009)^[163]

Der Ansatz der Analyse von SCoT-Polymorphismen wurde bereits erfolgreich zur Analyse beispielsweise von Mango, Kartoffeln oder Arten der Gattung *Citrus* L. sowie zur Identifizierung genetischer Variationen innerhalb der Art *Citrus sinensis* eingesetzt.^{[169]–[172]}

Insgesamt ist der Aufbau und die Funktionsweise dieser Methode vergleichbar zu anderen *non targeted*-Methoden wie die Analyse der ISSR (*inter-simple sequence repeats*) oder RAPD (*randomly amplified polymorphic DNA*). Im Vergleich zu der RAPD-PCR binden die Primer gezielt auf den

Bereichen der Startcodons und damit codierender, konservierter Bereiche. RAPD-Primer hingegen binden nicht gerichtet überall auf dem Genom, auch innerhalb nicht kodierender Bereiche. Vorteil der SCoT Marker ist, dass häufig eine höhere Anzahl polymorpher Banden resultiert.^{[170],[171]} Dies führt zu einer besseren Unterscheidung verschiedener Arten bzw. Sorten.

4.3.6 SCAR-Marker

SCAR-Marker (SCAR, engl. *sequence characterized polymorphic DNA*) wurden erstmals Anfang der 90er Jahre von I. Paran und R. W. Michelmore (1993) im Rahmen der Analyse von Resistenzgenen gegen falschen Mehltau bei Kopfsalat entwickelt.^[173] Es handelt sich um spezifische Marker, die bspw. im Rahmen von *non targeted*-Analysen identifiziert und anschließend charakterisiert werden. Mittels sequenzspezifischer Primer können diese schließlich gezielt amplifiziert werden.^[174] In der Regel werden sie basierend auf *fingerprinting*-Methoden wie bspw. der RAPD-PCR entwickelt.^[175]

Die RAPD-PCR wurde 1990 erstmals von Williams et al. vorgestellt.^[176] Es handelt sich dabei um eine *non targeted*-Methode, bei welcher kurze, ca. 10 bp-lange Primer für die PCR eingesetzt werden. Die Annealingtemperatur wird hierbei mit 35-45 °C vergleichsweise niedrig gewählt. Dadurch binden die Primer an unterschiedlichen Stellen der DNA und es entsteht ein Bandenmuster.^{[176],[177]} Da die Ergebnisse der RAPD-PCR häufig schwierig zu reproduzieren sind, wurden die SCAR-Marker entwickelt, welche im Anschluss den gezielten Nachweis reproduzierbar ermöglichen.^[175]

Hierfür werden identifizierte Marker kloniert und anschließend sequenziert. Der Schritt der Klonierung ist notwendig, da eine Bande häufig aus unterschiedlichen Fragmenten der gleichen Länge besteht und diese vor der Sequenzierung getrennt werden müssen. Anhand der Sequenzinformationen können schließlich sequenzspezifische Primer entwickelt werden, welche für die gezielte Amplifikation des Markers eingesetzt werden können.^[175] Für das Primerdesign werden in der Literatur zwei unterschiedliche Möglichkeiten vorgestellt. Paran und Michelmore (1993) entwickelten ursprünglich den Ansatz der Verlängerung der RAPD-Primer um die daran anschließenden 14 Basen.^{[173],[178]} Marieschi et al. (2012) hingegen schoben die spezifischen Primer ein, sodass die Sequenz des RAPD-Primers nicht Teil der neuen Primersequenzen war. Ziel war es damit die Spezifität zu erhöhen.^[179] Als Alternative zu der RAPD-PCR wird sehr häufig auch die AFLP (engl. *amplified fragment length polymorphism*) als Grundlage für die Entwicklung von SCAR-Markern herangezogen.^{[180],[181]}

Das Einsatzgebiet der SCAR-Marker ist sehr vielfältig. Häufig werden sie für den spezifischen Nachweis spezieller Gene herangezogen. So wurde diese Analysenmethode zum Beispiel für die Detektionen eines Gens, welches mit der Kernlosigkeit von Weintrauben in Verbindung steht, herangezogen.^[182] Des Weiteren werden SCAR-Marker auch für den Nachweis spezieller Resistenzgene genutzt, wie es bereits I. Paran und R. W. Michelmore (1993) im Zuge der Etablierung dieser Vorgehensweise taten.^{[173],[183]} Ein weiterer Einsatzbereich ist die Differenzierung unterschiedlicher Spezies. So wurden SCAR-Marker bspw. für die Authentizitätsbestimmung von

Safran (*Crocus sativus* L.), Granatapfel (*Punica granatum*) oder Thailand Jasminreis und dem Nachweis möglicher Verfälschungen dieser Lebensmittel entwickelt.^{[179],[184]–[186]}

4.3.7 PCR-Restriction fragment length polymorphism (PCR-RFLP)

Bei der PCR-RFLP handelt es sich um eine sehr schnelle Methode, die zur Unterscheidung von Organismen, Arten oder Sorten herangezogen werden kann. Kern dieser Methode ist das spezifische Schneiden eines DNA-Strangs durch ein Restriktionsenzym. Diese sind ein essentieller Bestandteil bei Klonierungsarbeiten und Genanalysen. Bei Restriktionsenzymen handelt es sich um Restriktionsendonukleasen. Ursprünglich kommen sie in Prokaryoten vor und dienen dort dem Schutz der DNA vor eingedrungener Bakteriophagen-DNA.^{[187],[188]} Traditionell werden sie nach dem entsprechenden Ursprungsorganismus mit einem Buchstaben-Triplet benannt. Zunächst wird der erste Buchstabe der Gattung und anschließenden die ersten zwei Buchstaben der Art angegeben. Weitere Angaben bzgl. Modifikationen sind möglich. In der Regel folgt noch eine fortlaufende Nummer für den Fall, dass mehrere Enzyme aus einem Organismus gewonnen werden.^[189] Die verschiedenen Enzyme besitzen unterschiedliche spezifische Erkennungssequenzen, die den Ort der Schnittstelle markieren. Die Mehrheit der bekannten Restriktionsenzyme gehört zu den Typ II Enzymen.^{[150],[188]} Diese zeichnen sich dadurch aus, dass sie i.d.R. als Homodimer mit einer molekularen Masse von ca. 2×30 kDa vorliegen.^[150] Ihre Erkennungssequenzen sind vier bis acht Basen lang und palindromisch aufgebaut.^{[150],[188]} Sie schneiden den DNA-Doppelstrang in unmittelbarer Nähe oder innerhalb der Erkennungssequenz in Anwesenheit von Mg^{2+} -Ionen.^[150]

Wird einer DNA-haltigen Lösung solch ein Restriktionsenzym zugegeben, so bindet dieses zunächst unter Freisetzung von Wasser unspezifisch durch Wechselwirkungen mit dem Phosphat-Rückgrat an die DNA.^{[150],[188]} Anschließend wandert das Enzym durch lineare Diffusion am DNA-Strang entlang.^[188] Für das Restriktionsenzym EcoRV wurde unter optimalen Bedingungen eine maximale Diffusionsgeschwindigkeit von $1,7 \times 10^6$ bp/s analysiert, wobei pro Bindung bis zu 2×10^6 bp gescreent werden können.^[190] Erreicht das Enzym seine Erkennungssequenz werden zwischen dem Enzym (Dimer) und den Basen der Restriktionsschnittstelle ca. 15-20 Wasserstoffbrückenbindungen und van der Waals-Wechselwirkungen ausgebildet. Zusätzlich liegen weitere Wasserstoffbrückenbindungen zu dem Phosphat-Rückgrat vor.^[150] Dieser Erkennungsprozess führt zu einer Konformationsänderung bei dem Enzym und der DNA und damit zu einer Aktivierung des katalytischen Zentrums. Nachdem die Phosphodiester-Bindungen beider DNA-Stränge geschnitten wurden, werden die DNA-Fragmente entweder durch direkte Dissoziation des Enzym-DNA-Komplexes oder Übertragung des Enzyms auf einen unspezifischen Abschnitt desselben DNA-Moleküls freigesetzt.^[150] Enzyme können die DNA abhängig von ihrer Restriktionsschnittstelle auf zwei unterschiedliche Arten schneiden. So können entweder glatte Enden (*blunt ends*) oder überstehende Enden (*sticky ends*) mit einem ein bis vier Basen langen 5'- oder 3'-Überhang entstehen.^{[150],[188]} Bei Typ II-Enzymen resultiert ein 5'-Phosphat- und 3'-OH-Ende.^[150]

Das in dieser Arbeit verwendete Restriktionsenzym ist *MfeI* und das entsprechende Isoschizomer zu dem Enzym *MunI*. Es besitzt die palindromische Erkennungssequenz 5'-C/AATTG-3' und wurde

ursprünglich aus dem Bakterium *Mycoplasma fermentas* isoliert.^[191] Heute wird es kommerziell über *E. coli* Stämme, welche das *MfeI*-Gen tragen, hergestellt. Sein Temperaturoptimum für eine hohe Aktivität liegt bei 37 °C. Aufgrund seiner Erkennungssequenz und dem Schnittmuster resultieren *sticky ends* mit einem vier Basen langen 5'-Überhang.^[191] Untersuchungen des Isoschizomers *MunI* zeigten, dass während der spezifischen Bindung zwischen Enzym und DNA 16 Wasserstoffbrückenbindungen und sechs van der Waals-Wechselwirkungen pro Erkennungssequenz ausgebildet werden.^[150]

Voraussetzung für die Unterscheidung zweier Organismen bzw. Arten oder Sorten mittels PCR-RFLP ist das Vorliegen einer Mutation (SNP oder Indel) innerhalb der Erkennungssequenz eines Restriktionsenzym. Für die Verwendung der PCR-RFLP als *targeted*-Methode wird zunächst mittels PCR ein Amplifikat, welches die Erkennungssequenz enthält, amplifiziert und dieses anschließend unter Zugabe eines geeigneten Puffers enzymatisch verdaut. Liegt eine intakte Schnittstelle vor, so wird das DNA-Fragment durch das Enzym geschnitten. Befindet sich jedoch eine Mutation innerhalb der Erkennungssequenz, erfolgt kein Verdau. Die Amplifikate können anschließend mittels AGE oder CGE auf das unterschiedliche Fragmentierungsmuster hin analysiert werden. Der Ablauf ist schematisch in Abbildung 10 dargestellt.

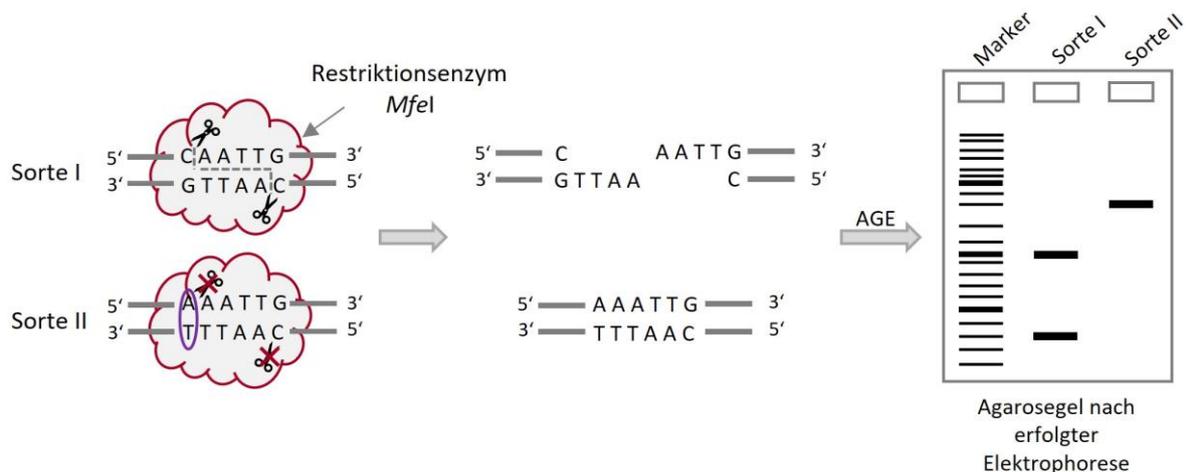


Abbildung 10: Schematischer Ablauf einer PCR-RFLP mit dem Restriktionsenzym *MfeI* und anschließender gelelektrophoretischer Auftrennung.

Die PCR-RFLP wurde auch im Bereich der Authentizitätsbestimmung bereits für verschiedenste Lebensmittel etabliert, z.B. für die Unterscheidung von Edel- und Konsumkakao, verschiedener Trüffel- oder Kaffee-Arten sowie zur Differenzierung von Fleischsorten um nur einige wenige Beispiele zu nennen.^{[43],[192]–[194]}

4.3.8 High resolution melting analysis (HRMA)

Die Schmelzpunktanalyse wurde in den 1990er Jahren zur Detektion von Unterschieden in der DNA-Sequenz entwickelt.^{[195],[196]} Durch die Weiterentwicklung der Interkalationsfarbstoffe sowie der zunehmenden Sensitivität der Analysengeräte wurde daraus die Hochauflösende Schmelzpunktanalyse (HRMA, engl. *high resolution melting analysis*).^[197]

Bei der HRMA werden Amplifikate im Anschluss an die PCR einem Temperaturgradienten ausgesetzt, wobei die Temperatur schrittweise von 65 °C auf 75 bzw. 80 °C erhöht wird. Die Start- und Endtemperatur können je nach zu erwartendem Schmelzpunkt angepasst werden.^{[198]–[200]} Die Temperaturerhöhung erfolgt i.d.R. in Schritten von 0,01 bis 0,2 °C.^[197] Ist der Schmelzpunkt eines DNA-Fragments erreicht, dissoziiert der Doppelstrang in seine Einzelstränge. Dies kann mit Hilfe von Interkalationsfarbstoffen wie *SYBR® Green* oder *SYTO™ 9* visualisiert und detektiert werden.^{[201],[202]} Zerfällt der Doppelstrang mit zunehmender Temperatur in seine Einzelstränge wird der Interkalationsfarbstoff freigesetzt und die Fluoreszenz nimmt ab (siehe Abbildung 11).

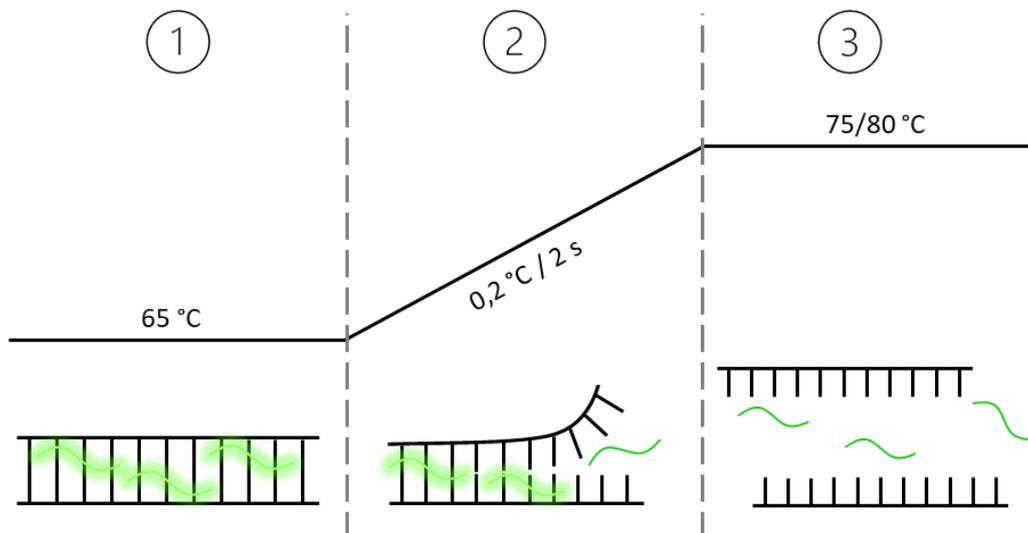


Abbildung 11: Drei Phasen eines beispielhaften Temperaturgradienten für eine HRMA. (1) Starttemperatur, bei der der DNA-Doppelstrang intakt ist und der Interkalationsfarbstoff eingelagert ist und fluoresziert. (2) Schrittweise Erhöhung der Temperatur, zunehmende Dissoziation des DNA-Doppelstrangs und Freisetzung des Interkalationsfarbstoffs und daraus resultierende Abnahme der Fluoreszenz. (3) Endtemperatur, vollständige, reversible Denaturierung der DNA, keine Fluoreszenz mehr detektierbar.

Aus der Betrachtung der Fluoreszenz in Abhängigkeit von der Temperatur ergibt sich die Schmelzkurve, der die Schmelztemperatur bzw. der Schmelzpunkt entnommen werden kann (siehe Abbildung 12). Der Schmelzpunkt ist definiert als die Temperatur, bei der die größte Abnahme der Fluoreszenz zu erkennen ist. Deutlicher wird dies, wenn die negative Ableitung der Fluoreszenz über die Temperatur ($-d(F)/d(T)$) gegen die Temperatur aufgetragen wird. Das Maximum der daraus resultierenden Kurve stellt den Schmelzpunkt dar (siehe Abbildung 12).^[200] Es handelt sich dabei also um die Temperatur, bei der die Hälfte der DNA noch doppelsträngig vorliegt und die andere Hälfte bereits als Einzelstrang.^[101]

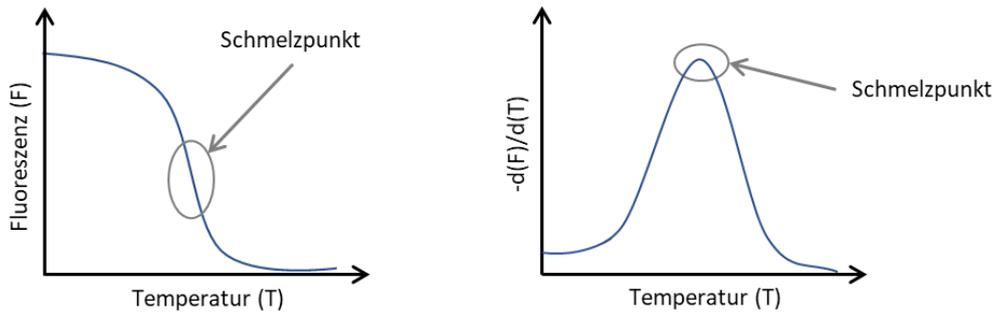


Abbildung 12: links: Schmelzkurve, erhalten durch Auftragen der Fluoreszenz gegen die Temperatur, rechts: Schmelzkurve, erhalten durch Auftragen der negativen Ableitung der Fluoreszenz über die Temperatur ($-d(F)/d(T)$) gegen die Temperatur.

Da das Fluoreszenzsignal auch bei gleichbleibenden PCR-Bedingungen und eingesetzten DNA-Mengen variieren kann und deshalb ein Vergleich verschiedener Schmelzkurven erschwert wird, ist für die Auswertung eine Normalisierung der Schmelzkurven notwendig. Hierbei werden die Fluoreszenzlevels zu Beginn und zum Ende der Schmelzkurve auf 100% bzw. 0% normalisiert. Um anschließend auch kleine Unterschiede zwischen den Schmelzkurven hervorzuheben, können sogenannte *Difference Plots* berechnet werden. Dazu wird eine Probe als Referenz definiert und die Differenzen der anderen Schmelzkurven in Bezug auf die Referenzprobe als Kurve angegeben.^[200]

Die Schmelztemperatur eines DNA-Fragments und die Form der Schmelzkurve werden von verschiedenen Faktoren beeinflusst: der Länge des Amplifikats, dessen GC-Gehalt, die Komplementarität und die Thermodynamik der Probenlösung.^{[195],[197],[198],[203]} Unterschiede innerhalb einer Sequenz in Form von SNPs, Insertionen oder Deletionen führen zu unterschiedlichen Schmelzpunkten sowie voneinander abweichenden Verläufen. Sie können damit sowohl über den Wert der Schmelzpunkte selbst, als auch die Form der Schmelzkurve detektiert werden. Es handelt sich bei der HRMA um eine schnelle und sensitive Methode zur Charakterisierung von PCR-Amplifikaten.^{[199],[201]} Sie findet einen breiten Anwendungsbereich in der Genotypisierung und dem Nachweis von Mutationen.^{[197],[202]} Die HRMA wurde bereits für die Analyse verschiedenster Pflanzen etabliert. So wurden auf der Basis von Mikrosatelliten, SNPs oder Indels beispielweise Wein- und Olivensorten, Basmatireis sowie Proben der Art *Citrus* und *Prunus* untersucht.^{[198],[204]–[206]} Dies sind nur einige wenige Beispiele. Die HRMA stellt damit ein neues, innovatives Tool für die Authentizitätsbestimmung von Lebensmitteln dar.^[200] Auch in der klinischen Forschung und Diagnostik ist sie eine inzwischen etablierte Methode.^{[207]–[209]}

Zusammenfassend handelt es sich bei der HRMA um eine einfach durchzuführende Analyse, welche auch für ein Hochdurchsatz-Screening geeignet ist.^{[197],[207]} Zudem wird durch den *closed-tube*-Ansatz das Risiko einer Kontamination verringert und die Amplifikate können im Anschluss für weitere Analysen herangezogen werden, da die Denaturierung der DNA-Doppelstränge reversibel verläuft.^{[199],[207],[208]}

5 Zielsetzung

Lebensmittelsicherheit und die Überwachung der Authentizität von Lebensmitteln gewinnt zunehmend an Bedeutung, sowohl für Verbraucher als auch die verarbeitende Industrie. Bisher erfolgt die Prüfung von Handelsware lediglich anhand von Lieferdokumenten. Es fehlt jedoch an analytischen Methoden mit Hilfe derer entsprechende Angaben verifiziert werden können. Bei Haselnüssen spielen sowohl die Herkunft, als auch die Sorte eine große Rolle. Haselnüsse unterschiedlicher Sorten weisen morphologische und organoleptische Unterschiede auf, die sich auf die Verarbeitung und die Sensorik beim Verzehr auswirken. Diese Parameter haben dadurch auch einen Einfluss auf den Preis.

Ziel der Arbeit war daher die Identifizierung von Sequenzunterschieden zwischen den Genomen unterschiedlicher Haselnussorten. Der Fokus lag hierbei auf der Untersuchung des Chloroplastengenoms, da dieses aufgrund seiner Größe von ca. 160.000 bp im Vergleich zum Kerngenom mit über 500 Mio. bp leichter zu sequenzieren ist.

Im ersten Schritt sollte also das Chloroplastengenom von mind. 12 unterschiedlichen Haselnussorten mittels *Next-Generation-Sequencing* (NGS) und einem MiSeq™ der Firma Illumina Inc. (San Diego, USA) sequenziert werden. Anhand der erhaltenen consensus-Sequenzen sollen polymorphe Bereiche definiert und mittels Sanger-Sequenzierung validiert werden. Neben dem cpGenom wurden parallel auch andere polymorphe Regionen auf dem Karyom, wie der ITS-Bereich oder die WRKY-Gene auf Sequenzunterschiede hin untersucht. Im nächsten Schritt sollten basierend auf identifizierten Differenzen *targeted*-Methoden für den spezifischen Nachweis von Sorten entwickelt werden. Die Wahl der dafür zugrundeliegenden Methode ist sowohl von der polymorphen Region selbst, als auch den dazu benachbarten Sequenzen abhängig. Mit Hinblick auf den Einsatz der entwickelten Methoden in der Routineanalytik der Lebensmittelüberwachung, von Dienstleistungslaboratorien sowie in der Wareneingangskontrolle der haselnussverarbeitenden Industrie sollen diese auch speziell für den Nachweis von Mischungen unterschiedlicher Haselnussorten optimiert und validiert werden. Zuletzt sollen die Methoden auf verarbeitete Produkte wie geröstete Haselnüsse übertragen werden, um sicherzustellen, dass diese in den genannten Bereichen eine Anwendung finden und bei jedem Prozessierungsschritt eingesetzt werden können.

Parallel sollte neben der Entwicklung von *targeted*-Methoden auch der Bereich der *non targeted*-Methoden erweitert werden. Da bereits viele Publikationen veröffentlicht wurden, die sich mit der Analyse der Mikrosatelliten sowie der RAPD-PCR befassen, sollte der Fokus in dieser Arbeit auf einer weiteren *non targeted*-Methode, der Analyse der SCoT-Marker, liegen.

6 Ergebnisse und Diskussion

Ziel der Arbeit war die Entwicklung von Methoden für den spezifischen Nachweis von Haselnussorten. Hierfür mussten zunächst Sequenzunterschiede identifiziert werden. Der Fokus lag dabei auf dem Chloroplastengenom, da dieses den Vorteil der rein maternalen Vererbung bietet. Parallel wurden auch Bereiche des Karyoms herangezogen, um die Anzahl an Ansatzpunkten zu erhöhen. Basierend auf identifizierten Sequenzunterschieden sollten anschließend molekularbiologische Methoden entwickelt werden mit Hilfe derer die Authentizität von Haselnussorten verifiziert werden kann. Unabhängig von den *targeted* Ansätzen wurde auch ein *non targeted*-Ansatz, die Analyse von SCoT-Polymorphismen herangezogen. Der Vorteil eines solchen *non targeted*-Ansatzes liegt darin, dass für die Durchführung der Analyse keine Sequenzinformationen bekannt sein müssen und er daher vollkommen unabhängig von anderen Analysen erfolgen kann. Die vorliegende Arbeit gliedert sich daher in die folgenden in Abbildung 13 grafisch zusammengefassten Abschnitte.

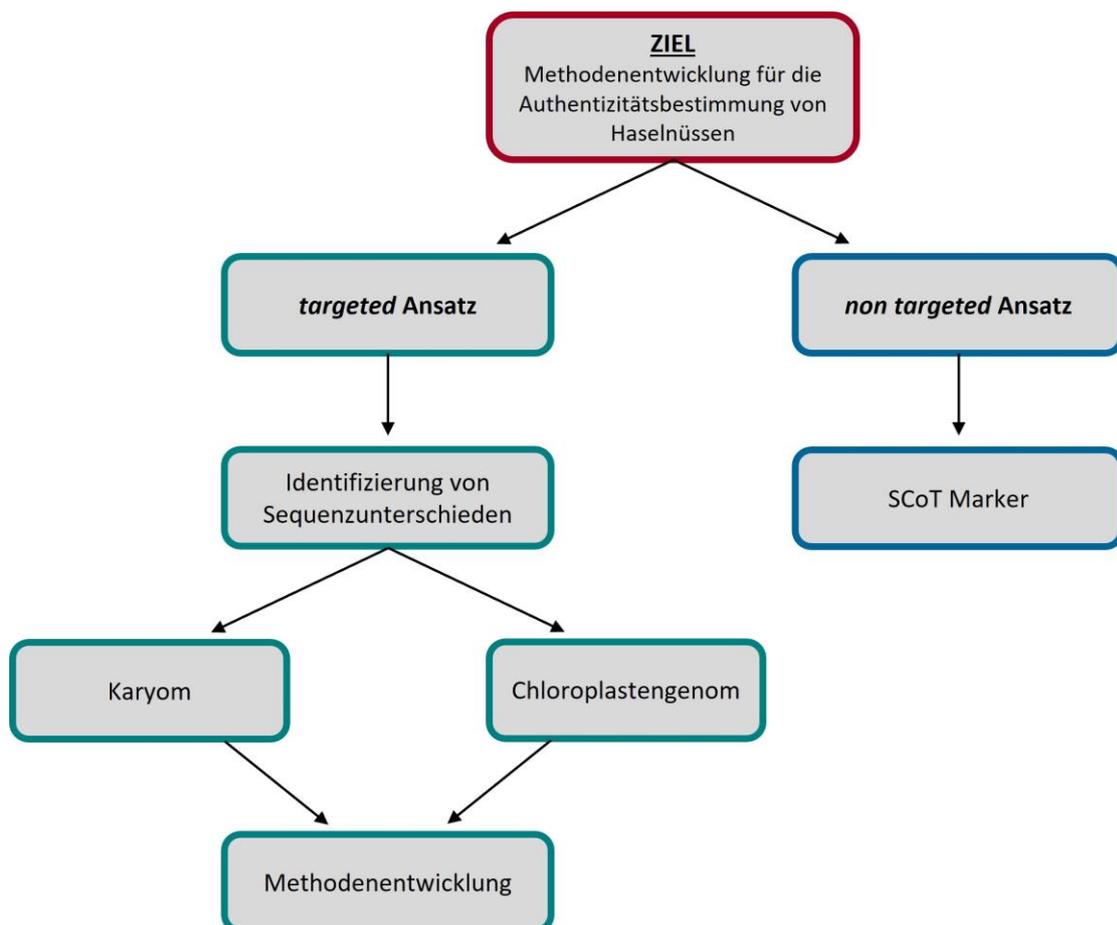


Abbildung 13: Grafische Zusammenfassung der analytischen Ansatzpunkte für die Methodenentwicklung zur Sortendifferenzierung von Haselnüssen.

6.1 Sequenzierung des cpGenoms mittels NGS

Im Verlauf dieser Arbeit wurde das cpGenom von 12 unterschiedlichen Haselnussorten (Ata Baba, Corabel, Barcelona, Butler, Hallsche Riesen, Negret, Palaz, Pautet, Segorbe, Tonda di Giffoni, T. G. Romana, T. G. Trilobata) wie unter 7.9 beschrieben mittels NGS sequenziert. Die Sequenzierungen beruhten dabei auf Gesamt-DNA-Isolaten von Haselnussblättern der genannten Sorten und wurden intern am hauseigenen Sequenziergerät (MiSeq™ von Illumina Inc., San Diego, USA) bzw. über einen externen Dienstleister (GATC Biotech; heute: Eurofins Genomics) durchgeführt. Der Tabelle 18 (Abschnitt 7.9) ist zu entnehmen, welche Proben intern und welche extern sequenziert wurden. Die generierten Rohdaten wurden in der *National Center for Biotechnology Information* (NCBI) Datenbank veröffentlicht. Die entsprechenden *Accession numbers* sind unter 9.10.1 in Tabelle 37 zusammengefasst. Die bioinformatische Auswertung der Sequenzierläufe erfolgte am Lehrstuhl für Angewandte Bioinformatik der Universität Tübingen wie in der zugehörigen Publikation (Lang et al. 2021) beschrieben.^[210] Pro Sequenzierung wurden zwischen 0,5 und 25 Mio. Reads mit einer Länge von 150 bzw. 250 bp generiert. Die Readlänge ist dabei abhängig von dem für die Probenvorbereitung verwendeten Kit der Firma Illumina Inc. (San Diego, USA). In Tabelle 18 im Abschnitt 7.9 ist aufgeschlüsselt, für welche Probe welches Kit eingesetzt wurde. Alle Sequenzierungen, die für die weitere Auswertung herangezogen wurden, zeigten bei der Qualitätsüberprüfung mittels FastQC keine abweichenden Parameter. Zudem wurden $\leq 0,5\%$ der Reads bei der Überprüfung auf Kontaminationen gemapped. Hierbei wurde bspw. auf die Verunreinigung mit humaner DNA im Zuge der Probenvorbereitung oder der Verunreinigung mit *Pseudomonas*, die sich häufig auf der Oberfläche von Blättern aufhalten, getestet. Für die Erstellung einer consensus-Sequenz wurden die Reads gegen die veröffentlichten cpGenome der *C. chinensis* (KX814336.2) und der *C. avellana* (KX822768.2) gemapped. Um welche Sorten es sich genau handelt, ist bei den veröffentlichten Genomen nicht angegeben. Es konnten jeweils zwischen 2,5% und 14% der erhaltenen Reads zu den Referenzgenomen gemapped werden. Da die Kontrollen auf Kontaminationen negativ ausfielen, ist davon auszugehen, dass die verbliebenen Reads dem Mitochondrien- und Kerngenom zuzuordnen sind. Damit wurde abhängig vom jeweiligen Sequenzierlauf eine *coverage* (= mittlere Abdeckung) von 34x bis 3600x erreicht. Illumina Inc. (San Diego, USA) empfiehlt für eine Gesamtgenom-Sequenzierung eine *coverage* von 30x bis 50x.^[211] Rowley et al. (2018) wählte für eine Resequenzierung des Kerngenoms sieben europäischer Haselnussorten lediglich eine *coverage* von etwa 20x.^[11] Damit liegen die durchgeführten Sequenzierungen im empfohlenen Bereich. Insgesamt war der Datenoutput der extern durchgeführten Sequenzierungen im Vergleich zu den internen deutlich höher, woraus auch eine höhere *coverage* resultierte. Dies zeigte ein vorhandenes Optimierungspotenzial in Bezug auf die Sequenzierung inkl. der Probenvorbereitung. Gleichzeitig stellt eine ausführliche Optimierung aufgrund der Vielzahl an zu optimierenden Parametern und hohen Kosten eines Sequenzierlaufs mittels NGS auch einen deutlichen Kostenfaktor dar.

In den meisten Publikationen werden inzwischen Gesamt-DNA-Isolate für die Sequenzierung eingesetzt.^{[212]–[214]} Hierbei gestaltet sich die Assemblierung deutlich komplexer, da die erhaltenen

Reads den unterschiedlichen Genomen zugeordnet werden müssen. Aufgrund der in den letzten Jahren stark gestiegenen Anzahl an veröffentlichten Sequenzen, kann jedoch in den meisten Fällen eine veröffentlichte Sequenz einer möglichst nah verwandten Art oder Sorte als Vorlage für die Assemblierung herangezogen werden, so wie es auch in der vorliegenden Arbeit durchgeführt wurde. Damit ist eine vorherige Isolierung der Plastide in möglichst hochreiner Form nicht mehr notwendig.^[213] Eine weitere Alternative stellt die Sequenzierung von *long range* PCR-Produkten dar.^[215] Dabei werden unter entsprechenden PCR-Bedingungen Amplifikate mit einer Länge von mehreren 10 kbp generiert.^{[100],[215]} Dieses Verfahren wurde bspw. von Li et al. (2015) für die Sequenzierung des cpGenoms der *Ostrya rehderiana* eingesetzt. Es handelte sich dabei um das erste veröffentlichte cpGenom innerhalb der Familie der Betulaceae.^[216] Dieses wiederum ermöglichte anschließend eine einfachere Assemblierung anderer Arten innerhalb dieser Familie (z.B. der *Corylus fargesii* oder der *Corylus chinensis*) auf Basis einer Sequenzierung des Gesamtgenoms.^{[94],[217]}

Da allgemein für das cpGenom eine geringe Mutationsrate von nur $1-3 \times 10^{-9}$ Basen pro Jahr publiziert ist, wurde analysiert, inwieweit sich die verwendeten Referenzgenome (*C. avellana* vs. *C. chinensis*) ähneln.^[93] Ein Vergleich der beiden Genome zeigte, dass eine 99,96 %ige Übereinstimmung vorliegt. Die 0,04% Unterschied entsprechen bei einer Länge von ca. 160.000 bp immerhin noch etwa 6400 interspezifischen SNPs. Es ist jedoch zu beachten, dass hier zwei unterschiedliche botanische Arten verglichen wurden. Bei der vorliegenden Arbeit wurden jedoch Proben unterschiedlicher Sorten auf intraspezifische Unterschiede hin untersucht. Daher ist davon auszugehen, dass auf Sortenebene eine noch deutlich größere Übereinstimmung vorliegt.

Im Rahmen der bioinformatischen Auswertung wurden pro Sorte 1-65 *high quality* SNPs identifiziert. Daraus resultierten schließlich 92 SNPs, die letztendlich individuell für eine Sorte sein sollten. Zudem wurden 251 Indels identifiziert. Die erfassten Sequenzunterschiede müssen jedoch sorgfältig validiert werden, da diese v.a. in Bereichen niedriger *coverage* auftraten. Daher muss ausgeschlossen werden, dass es sich um Sequenzier- oder Auswertefehler handelt.

6.2 Identifizierung und Validierung von Sequenzunterschieden

Auf Basis der unter 6.1 dargestellten Ergebnisse wurden vier Bereiche des cpGenoms (siehe Tabelle 3) ausgewählt, die möglichst viele der identifizierten SNPs enthielten, und mittels Sanger-Sequenzierung resequenziert, um die potentielle Sequenzunterschiede zu validieren. Die Sequenzen der vier Abschnitte sowie die zugehörigen Primerbindungsstellen sind im Anhang unter 9.10.2 in Abbildung 55 bis Abbildung 58 dargestellt.

Tabelle 3: Resequenzierte Bereiche des cpGenoms. Die angegebenen Sequenzpositionen beziehen sich auf das veröffentlichte Referenzgenom einer *C. avellana* unbekannter Sorte mit der *Accession No.* KX822768.2.

Bereich	Sequenzabschnitt	
A	76 – 773 bp	Überwiegend nicht-kodierender Bereich, bis Pos. 95 trnH-GTG, ab Pos. 535 psbA-Gen
B	4455 – 5046 bp	Nicht-kodierender Bereich zwischen matK und rps16
C	35831 – 36324 bp	Nicht-kodierender Bereich zwischen trnT-GGT und psbD
D	60770 – 61439 bp	Nicht-kodierender Bereich zwischen rbcl und psal

Von neun ausgewählten Sorten (Ata Baba, Palaz, Segorbe, Pauetet, Hallsche Riesen, Negret, Tonda di Giffoni, T. G. Romana und T. G. Trilobata), erfolgte die Resequenzierung ebenfalls auf Basis von Blatt-Material. Die DNA-Isolate wurden jeweils neu hergestellt und beruhten dabei stets auf nur einem Blatt, um mögliche Unterschiede zwischen verschiedenen Pflanzen auszuschließen. Denn eine rein maternale Vererbung des cpGenoms bei Haselnüssen ist, wie unter 4.2.3 näher erläutert, anzunehmen, es gibt jedoch bisher keine Publikation, die das abschließend untersucht und bestätigt. Von jeder Sorte wurden mindestens zwei unterschiedliche Isolate analysiert.

Das Ergebnis war eine eindeutige Unterteilung in zwei Gruppen. Gruppe 1 bilden die Sorten Ata Baba aus Aserbaidschan und Palaz aus der Türkei. Die Gruppe 2 bilden alle europäischen Proben. Aufgrund der deutlichen geografischen Clusterung wurde nachträglich noch eine Probe georgischer Haselnüsse ebenfalls von allen vier Bereichen sequenziert. Die DNA-Isolate beruhten dabei auf Nussmaterial, da keine Blätter aus diesem Land vorhanden waren. Eine tabellarische Übersicht der identifizierten Sequenzunterschiede ist Tabelle 38 im Anhang unter 9.10.2 zu entnehmen. Des Weiteren sind diesem Abschnitt Alignments aller vier Bereiche zu entnehmen, die basierend auf je einer Probe pro Sorte erstellt wurden. Insgesamt konnten sieben SNPs (SNP 1-7) sowie fünf Indels (Indel 1-5) identifiziert und validiert werden. Die im Folgenden verwendete Nummerierung wird für den weiteren Verlauf der Arbeit beibehalten. Bei drei der Indels (Indel 2, 4 und 5) handelt es sich lediglich um Einschübe einer Base (A bzw. T) in einer Poly-A- bzw. Poly-T-Region. Indel 2 und 4 sind dabei spezifisch für die Sorten Ata Baba und Palaz und Indel 5 für georgische Haselnüsse. Die anderen beiden Indels (Indel 1 und 3) befinden sich in den Bereichen B und C. Es handelt sich um deutlich größere Insertionen mit einer Länge von 118 bzw. 115 bp. Beide Insertionen sind nur bei den georgischen Haselnüssen sowie den Sorten Palaz und Ata Baba nachzuweisen. Neben diesen beiden Inserts ermöglichen die SNPs 1, 5 und 7 dieselbe Trennung. Die SNPs 2 und 4 befinden sich jeweils innerhalb eines der Inserts und ermöglichen eine Abtrennung der georgischen Nüsse von den Sorten Ata Baba und Palaz. Zudem sind auch die SNPs 3 und 6 spezifisch für die georgischen Proben. Der Basenaustausch von SNP 6 (A/C-Austausch) resultiert bei den Haselnüssen aus Georgien in der Erkennungssequenz des Restriktionsenzym *MfeI* (5'-CAATTG-3').

Die erhaltenen Ergebnisse weisen auf eine deutliche geografische Gruppierung hin. Es können insgesamt drei Gruppen definiert werden (siehe 6.7 Abbildung 52). Die Inserts bieten hierbei den Vorteil, dass sie aufgrund ihrer Länge bereits mittels einer einfachen AGE nach erfolgter PCR nachweisbar sind. Zusätzlich stellt SNP 6 den für den Nachweis georgischer Proben bevorzugten

Sequenzunterschied dar, da er mittels PCR-RFLP nachgewiesen werden kann. Im Folgenden wurde daher auf diese Sequenzunterschiede der Fokus gelegt. Ziel war die Entwicklung von *targeted*-Methoden für deren spezifischen Nachweis. Die etablierten Methoden sollten schließlich auf einen vergrößerter Probenpool, sowohl in Bezug auf Sorte als auch Anbaugebiet, angewendet werden. Die Resultate werden in den folgenden Kapiteln dargestellt.

6.3 Entwicklung und Anwendung von *targeted*-Methoden auf Basis polymorpher Bereiche des cpGenoms

6.3.1 Nachweis der Indels 1 und 3 in den Bereichen B und C des cpGenoms

6.3.1.1 Qualitativer Nachweis

Der qualitative Nachweis der Indels 1 und 3 kann mittels einer einfachen PCR und anschließender AGE erfolgen. Für die PCR werden die Primerpaare B.1_FW/RV und C_FW/RV (siehe Tabelle 24 sowie Abbildung 56 und Abbildung 57) eingesetzt. Diese amplifizieren jeweils den Bereich um die Insertionen herum, sodass diese, falls vorhanden, vollständig erfasst werden. Detaillierte Informationen zu der PCR sind Abschnitt 7.11.1.1 zu entnehmen.

Die Insertionen weisen eine Länge von 118 bp im Bereich B und 115 bp im Bereich C auf. Damit besitzen die mittels PCR amplifizierten DNA-Abschnitte im Bereich B eine Länge von 592 bzw. 710 bp und im Bereich C 494 bzw. 609 bp. Für die vorangehende DNA-Isolierung kann sowohl die CTAB-Methode als auch die *Rapid Extraction* herangezogen werden. Mit beiden Methoden konnten DNA-Isolate generiert werden, die für eine qualitative PCR eingesetzt werden können. Für ein schnelles Screening großer Probenmengen ist aufgrund einer Zeitersparnis von über 50% die *Rapid Extraction* zu bevorzugen. Handelt es sich bei den zu analysierenden Proben jedoch um prozessierte Produkte, die neben Haselnüssen weitere Bestandteile enthalten, ist die DNA-Isolierung mittels CTAB-Methode zu empfehlen, da bei diesem Vorgehen eine bessere Aufreinigung der DNA erfolgt und störende Begleitsubstanzen, die die Amplifikation beeinträchtigen können, entfernt werden.

Aufgrund der Größe der gewählten Fragmente und der Größe der Insertion ist eine AGE für die Detektion ausreichend. Anhand einer geeigneten DNA-Leiter kann visuell ermittelt werden, ob die analysierte Probe die Insertion aufweist oder nicht. Abbildung 14 zeigt beispielhaft ein Agarosegel der Amplifikate von Proben unterschiedlicher Sorte bzw. Herkunft von Bereich C. Die Differenzen in den Längen der Fragmente können visuell detektiert werden. Dies verdeutlicht bspw. der Vergleich der georgischen und aserbaidchanischen Proben mit den Proben europäischer Herkunft der Sorten Hallsche Riesen und Segorbe. Die Probe der Sorte Tombul weist eine Doppelbande auf, die darauf hindeutet, dass es sich bei den für die Herstellung des DNA-Isolats verwendeten Nüssen um eine Mischung unterschiedlicher Sorten handeln muss.

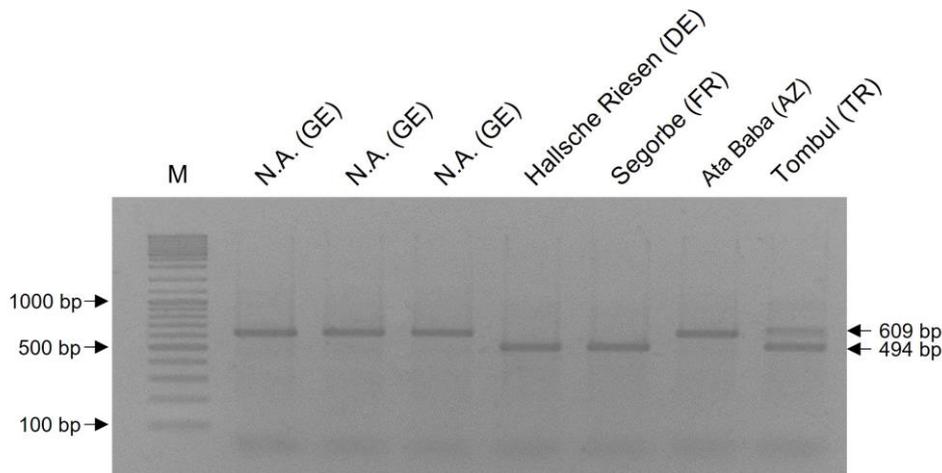


Abbildung 14: Agarosegel der PCR-Amplifikate von Bereich C sieben ausgewählter Proben unterschiedlicher Sorte und Herkunft. Die Proben aus Georgien und Aserbaidschan weisen ein längeres Fragment (609 bp) auf, was zeigt, dass diese Proben die Insertion aufweisen. Die Proben aus Deutschland und Frankreich hingegen zeigen ein kürzeres Amplifikat (494 bp). Die Probe der Sorte Tombul weist eine Doppelbande auf, weshalb davon auszugehen ist, dass in diesem Fall eine Mischung vorliegt. M = DNA-Leiter, GE = Georgien, DE = Deutschland, FR = Frankreich, AZ = Aserbaidschan, TR = Türkei.

Da im Rahmen der Sequenzierung und der anschließenden Resequenzierung die Insertionen, wie unter 6.2 erläutert, lediglich bei den Sorten Ata Baba und Palaz sowie georgischen Haselnüssen nachgewiesen werden konnten, deuten diese Ergebnisse auf eine geografische Gruppierung hin. Daher wurde die entwickelte Methode für die Analyse der Indels auf ein erweitertes Sorten- und Probenspektrum angewendet. Der Probenpool wurde v.a. in Bezug auf Haselnussproben aus den östlichen Anbaugebieten Türkei, Aserbaidschan und Georgien erhöht. Der Fokus lag im Folgenden ausschließlich auf der Analyse von Haselnusskernen, da diese den zukünftigen Anwendungsschwerpunkt darstellen. Insgesamt wurden 109 Nussproben (siehe Tabelle 27) bestehend aus 19 unterschiedlichen Sorten (Ata Baba, Barcelona, Butler, Çakıldak, Corabel, Ennis, Hallsche Riesen, Karayağlı, Mortarella, Negret, Palaz, Pauetet, Sariyağlı, Segorbe, Tonda di Giffoni, T. G. Romana, T. G. Trilobata, Tombul und Yomra) sowie Proben unbekannter Sorte aus sieben Ländern (Deutschland, Frankreich, Spanien, Italien, Türkei, Aserbaidschan und Georgien) untersucht. Zusätzlich lagen vier Proben von Wildtypen aus der Türkei vor. Einige der Proben wurden im Rahmen eines anderen geförderten FEI-Projektes (AiF 18040 N) in der Türkei vor Ort gesammelt. Hierbei wurden die Proben jeweils strauchweise gezogen und die GPS-Koordinaten der beprobten Sträucher notiert, damit in dem darauf folgenden Jahr genau derselbe Strauch wieder beprobt werden kann. Es lagen somit 21 Proben türkischer Haselnüsse vor, von denen jeweils Probenmaterial aus den Erntejahren 2014 und 2015 untersucht werden konnte, das von demselben Strauch stammte. In Tabelle 4 sind die Ergebnisse der Analyse auf das Vorhandensein der Indels dieser doppelten Proben zusammengefasst.

Tabelle 4: Übersicht der Analyse türkischer Haselnussproben aus den Erntejahren 2014 und 2015 in Bezug auf die Indels. Nebeneinander stehende Proben stammen jeweils von demselben Strauch. + Insertion vorhanden, - Insertion nicht vorhanden.

Haselnussorte	Erntejahr 2014		Erntejahr 2015	
	Probennr.	Ergebnis	Probennr.	Ergebnis
Çakıldak	213	-	215	-
Karayağlı	185	-	187	-
	133	-	188	-
	154	-	190	-
	153	-	189	-
	193	-	197	-
Palaz	139	+	167	+
	140	+	171	+
Sariyağlı	155	+	198	+
	211	+	199	+
	131	+	200	+
	157	+	201	+
	158	+	202	+
	149	+	203	+
	159	+	205	+
	160	+	206	+
	161	+	210	+
Tombul	141	+	176	+
Yomra	180	+	184	+
	143	+	182	+
	142	+	181	+

Alle Probenpaare zeigten identische Ergebnisse. Dies bestätigt die vermutete maternale Vererbung des cpGenoms im Falle von Haselnüssen. Bisher gibt es keine Veröffentlichung, welche sich speziell mit der Art *Corylus* befasste. Da allerdings, wie unter 4.2.3 bereits erläutert, 80% der Gattung Angiospermen und auch weitere Vertreter der Ordnung Fagales, der auch die Haselnuss angehört, wie z.B. die Stieleiche (*Quercus robur*) eine maternale Vererbung des cpGenoms aufweisen, ist davon auszugehen, dass dies auch bei der Haselnuss der Fall ist.^{[89],[92]} Zudem lässt sich daraus schließen, dass die erhaltenen Ergebnisse auch auf andere Jahrgänge übertragen werden können.

Im Folgenden wird für die weitere Auswertung nur jeweils eine der beiden Proben berücksichtigt, um durch die zusätzlichen Proben nicht die Probenanzahl fälschlicherweise zu erhöhen und das Ergebnis zu beeinflussen. Die verbleibenden 88 Proben setzen sich gemäß ihres Herkunftslandes wie folgt zusammen: 13x Europa, 5x Aserbaidschan, 7x Georgien und 63x Türkei. Die Ergebnisse der Analyse, aufgeschlüsselt auf die einzelnen Proben, sind im Anhang unter 9.10.3 in Tabelle 39 zusammengefasst.

Sowohl die europäischen, als auch die georgischen und aserbaidischen Proben zeigten eindeutige Ergebnisse, die sich mit den Resultaten der Sanger-Sequenzierung (siehe 6.2) deckten. Allen Haselnussproben aus Europa fehlt das Insert, wohingegen die Proben aus Aserbaidschan und Georgien das Insert enthalten. Somit konnte die Gruppierung der Haselnussorten nach ihrem Anbaugebiet, die sich während der Sequenzierung und Resequenzierung andeutete, bestätigt werden. Es konnten keine Unterschiede innerhalb des Probenpools eines der genannten Gebiete festgestellt werden. Anders hingegen war das Ergebnis der türkischen Proben (siehe Tabelle 5). Der Probenpool setzte sich hier folgendermaßen zusammen: 8x Çakildak, 16x Karayağlı, 9x Palaz, 13x Sariyağlı, 10x Tombul und 7x Yomra.

Tabelle 5: Ergebnis der qualitativen Analyse türkischer Haselnussproben auf Vorhandensein der Insertionen 1 und 3 innerhalb der Bereiche B und C des cpGenoms mittels PCR.

Sorte	Probenanzahl	Ergebnis
Çakildak	8 von 8	(-) Insertion
Sariyağlı	13 von 13	(+) Insertion
Yomra	7 von 7	(+) Insertion
Karayağlı	5 von 16	(+) Insertion
	11 von 16	(-) Insertion
Palaz	7 von 9	(+) Insertion
	1 von 9	(-) Insertion
	1 von 9	Nicht auswertbar, da Doppelbande
Tombul	1 von 10	(+) Insertion
	5 von 10	(-) Insertion
	4 von 10	Nicht auswertbar, da Doppelbande
Wildtypen	2 von 4	(+) Insertion
	2 von 4	(-) Insertion

Bei den Sorten Çakildak, Sariyağlı und Yomra wiesen alle untersuchten Proben einer Sorte dasselbe Ergebnis auf. Die Sorte Çakildak ist dabei der Gruppe der europäischen Proben zuzuordnen, da die Insertionen nicht vorhanden sind. Die Proben der Sorten Sariyağlı und Yomra hingegen wiesen alle die Insertionen auf und clustern damit mit den Proben aus Aserbaidschan und Georgien. Die Sorten Karayağlı, Palaz und Tombul hingegen zeigten ambivalente Ergebnisse. So wiesen von 16 untersuchten Proben der Sorte Karayağlı fünf die Inserts auf und elf nicht. Daher kann für diese Sorte keine genaue Zuordnung getroffen werden. Die Sorte Palaz kann ebenfalls keiner der beiden Gruppen zugeordnet werden. Eine Probe (Nr. 116) zeigte auf dem Agarosegel eine Doppelbande, was den Rückschluss zulässt, dass es sich um eine Mischung unterschiedlicher Sorten handeln muss, da sowohl Amplifikate mit als auch ohne Insert durch die PCR vervielfältigt wurden. Von den verbliebenen acht Proben wiesen sieben Proben das Insert auf und nur eine Probe (Nr. 169) nicht. Daher ist davon auszugehen, dass die Sorte Palaz der Gruppe der aserbaidischen und georgischen Proben zuzuordnen ist. Eine abschließende Aussage kann jedoch anhand dieser Ergebnisse nicht getroffen werden. Bei der Sorte Tombul wiesen insgesamt vier Proben (Nr. 114, 115, 119 und 120) eine Doppelbande auf. Es handelt sich bei diesen um Probenmaterial, das von

Industriepartnern zur Verfügung gestellt wurde. Alle anderen Proben (Nr. 127, 141, 173-175, 129) wurden über das Projekt AiF 18040 N vor Ort in der Türkei gesammelt. Diese zeigten ein eindeutiges Ergebnis, verhielten sich jedoch ebenfalls ambivalent. Eine Probe (Nr. 141) wies das Insert auf und fünf nicht. Daher kann tendenziell von einer Zuordnung der Sorte zu der Gruppe der Haselnüsse ohne Insertion ausgegangen werden. Allerdings ist eine abschließende Bewertung ohne Verifizierung anhand weiterer authentischer Proben nicht möglich.

Da die Haselnuss eine Selbstinkompatibilität aufweist und damit auf eine Fremdbestäubung angewiesen ist, befinden sich häufig auf den Plantagen verteilt Bäume unbekannter Sorte i.d.R. Wildtypen, die der Bestäubung der anderen Haselnusssträucher bzw. -bäume dienen. Die vier untersuchten Wildtypen (TR) zeigten ein ambivalentes Ergebnis. Die Proben Nr. 135 und 137 wiesen die Insertionen nicht auf, die Proben Nr. 136 und 138 hingegen schon. Im weiteren Verlauf werden die vier untersuchten Wildtypen nicht weiter betrachtet, da sie keiner Sorte zuzuordnen sind.

Zusammenfassend ist auf Basis der Indels 1 und 3 eine Unterteilung der analysierten Sorten in folgende Gruppen möglich (siehe Tabelle 6):

Tabelle 6: Gruppierung der analysierten Haselnussproben basierend auf den Ergebnissen der Indels.

Insertionen vorhanden	Insertionen nicht vorhanden	Zuordnung nicht möglich
<u>Europa:</u>	<u>Türkei:</u>	<u>Türkei:</u>
Barcelona, Butler, Corabel, Ennis, Hallsche Riesen, Mortarella, Negret, Pauetet, Segorbe, Tonda di Giffoni, T. G. Romana, T. G. Trilobata	Yomra, Sariyağlı	Karayağlı, Palaz, Tombul
<u>Türkei:</u>	<u>Aserbaidsschan:</u>	
Çakildak	Ata Baba, NA	
	<u>Georgien:</u>	
	NA	

Anhand der identifizierten Indels ist also eine eindeutige Unterscheidung zwischen europäischen und georgischen bzw. aserbaidsschanischen Haselnüssen möglich. Türkische Sorten zeigten ein ambivalentes Verhalten. Die Sorte Çakildak ist der Gruppe der europäischen Haselnüsse und die Sorten Yomra und Sariyağlı sind der Gruppe der georgischen und aserbaidsschanischen Sorten zuzuordnen. Drei weitere türkische Haselnussorten (Karayağlı, Palaz und Tombul) konnten bisher nicht eindeutig einer Gruppe zugeordnet werden, da die Ergebnisse der Proben innerhalb einer Sorte differenzierten und deshalb keine genaue Zuordnung zuließen. Dies verdeutlicht die Schwierigkeit authentisches Probenmaterial aus den östlichen Anbaugebieten zu erhalten. Speziell die Türkei spielte hierbei eine große Rolle, da sie im Vergleich zu Georgien und Aserbaidsschan einen deutlich größeren Marktanteil an der Welternte hat und gerade auch für den europäischen Markt eine wichtige Rolle spielt. Der Haselnussanbau in der Türkei ist sehr traditionell und durch viele kleine familiengeführte Betriebe geprägt. Die Analyse von Mikrosatelliten von Gökirmak et al. (2009) und Gürçan et al. (2010) zeigte bereits eine hohe Diversität innerhalb von Haselnüssen der Schwarzmeer-Region. Bei vielen Sortenbezeichnung handelt es sich lediglich um Synonyme, die nicht auf genetischen Differenzen beruhen, sondern viel mehr durch Anbauregionen oder morphologische Merkmale geprägt werden.^{[8],[218]} Die türkische Haselnussorte Tombul ist eine der wichtigsten und

am häufigsten angebauten Sorten.^[13] Sie gehört bspw. aufgrund ihrer charakteristischen Merkmale, wie z.B. ihre Form, zu den höherpreisigen Sorten und wird als „Giresun Qualität“ verkauft. Allerdings zeigten Untersuchungen auf Basis der Mikrosatelliten bereits, dass Haselnüsse der Sorte Tombul bzw. der Giresun Qualität genetische Unterschiede aufweisen.^[218] Dies deckt sich mit der Beobachtung, dass v.a. bei der Sorte Tombul viele Proben, gerade kommerziell erworbene, Mischungen sind. Spanische und italienische Haselnüsse hingegen zeigen auf Basis von Mikrosatelliten einen hohen Verwandtschaftsgrad, weshalb sie i.d.R. in einer Gruppe zusammengefasst werden.^{[8],[218]}

Eine zu den Ergebnissen der Indels vergleichbare Gruppierung zeigte sich auch bei Untersuchungen von Boccacci und Botta (2009) basierend auf den Mikrosatelliten des cpGenoms.^[28] Analysiert wurden Proben aus Spanien, Italien, Türkei und Iran. Hierbei identifizierten sie vier Chlorotypen A-D. Chlorotyp A ist dabei am weitesten verbreitet, nimmt jedoch von den westlichen in Richtung der östlichen Anbaugebiete ab. So weisen fast alle spanischen und italienischen Sorten diesen Chlorotypen auf. Haselnüsse aus der Türkei wiesen zu 40% den Typ A und zu 60% den Typ B auf. Typ C war nur in iranischen Haselnüssen nachzuweisen und Typ D nur in zwei italienischen Proben der Sorten Tonda bianca und Tonda rossa.^[28] Diese Beobachtungen decken sich allgemein mit der Gruppierung, welche auf Basis der Indels möglich ist, und auch dem ambivalenten Verhalten türkischer Haselnüsse.

Zusammenfassend ist zu sagen, dass sich die entwickelte Methode sehr gut für eine Prüfung der Authentizität europäischer Haselnussproben eignet. Anhand der Indels können Beimischungen georgischer und aserbaidtschanischer Proben eindeutig nachgewiesen werden. Von den analysierten türkischen Sorten, zeigte lediglich die Sorte Çakıldak eine eindeutige Übereinstimmung. Es können daher nicht alle möglichen Beimischungen nachgewiesen werden. Dennoch bietet die entwickelte PCR eine sehr einfache Möglichkeit den Großteil der östlich kultivierten Haselnussorten in europäischen Proben nachzuweisen. Dies ist ein wichtiger Aspekt, da bspw. Haselnüsse aus Georgien und Aserbaidschan zu deutlich niedrigeren Preisen auf dem Weltmarkt gehandelt werden und daher eine wirtschaftlich motivierte Verfälschung denkbar ist. Solche Beimischungen können anhand der Indels nachgewiesen werden. Bei den Analysenergebnissen der kommerziell erworbenen Proben der Sorte Tombul zeigte sich bereits ein Vorteil dieser Methode, denn anhand einer einfachen PCR konnte gezeigt werden, dass es sich häufig nicht um sortenreine Proben handelt. Damit ist in Kombination mit der *Rapid Extraction* in ca. 2,5 h eine erste Beurteilung der Authentizität einer Probe möglich. Und da sich die beiden Bereiche B und C des cpGenoms in Bezug auf „Insertion vorhanden“ oder „Insertion nicht vorhanden“ stets synchron verhielten, ist auch nur die Analyse einer der beiden Bereiche notwendig.

6.3.1.2 Nachweis der Indels 1 und 3 in prozessierten Haselnüssen

Häufig werden die Haselnüsse bereits in den Anbauländern ersten Prozessierungsschritten unterzogen bevor sie an die weiterverarbeitende Industrie ausgeliefert werden. So werden die Nüsse beispielsweise bereits geröstet und/oder gemahlen. Aufgrund von Temperaturen $> 100\text{ }^{\circ}\text{C}$ während des Röstvorgangs kann eine Degradation der DNA auftreten. Deshalb wurde der zuvor unter 6.3.1.1 geschilderte qualitative Nachweis auf geröstete Haselnussproben sowie selbst gebackene Kekse angewendet.

Der Nachweis der Indels wurde mit kommerziell erworbenen Proben gerösteter Haselnüsse der Sorten Tombul (V2) und T. G. Trilobata (V3) sowie selbst hergestellten Keksen (V1) mit französischen Haselnüssen durchgeführt. In allen Proben konnten die Amplifikate der Bereiche B und C generiert werden (siehe Abbildung 15). Die Sorte T. G. Trilobata sowie die Kekse mit französischen Haselnüssen zeigten wie zu erwarten keine Inserts. Die Sorte Tombul hingegen wies eine Doppelbande auf. Daher ist davon auszugehen, dass es sich bei diesen Nüssen ebenfalls um eine Mischung handelt.

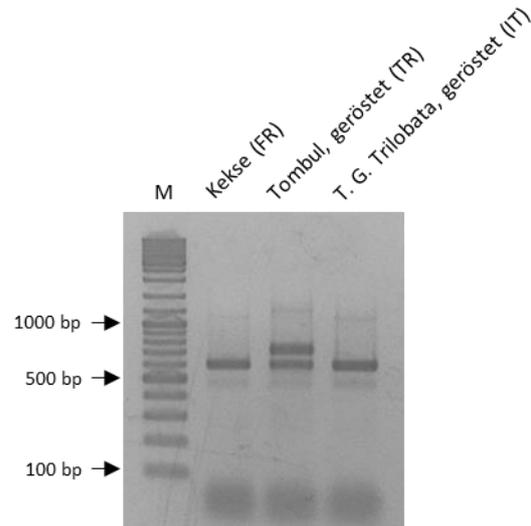


Abbildung 15: Anwendung der entwickelten *targeted*-Methoden für den qualitativen Nachweis der identifizierten Indels 1 und 3 im Bereich B des cpGenoms. M = DNA-Leiter, F = Frankreich, TR = Türkei, IT = Italien.

Das Ergebnis zeigt, dass eine Anwendbarkeit auf verarbeitete Produkte gegeben ist. Damit ist der vorgestellte qualitative Nachweis der identifizierten Indels für die geografische Einordnung der Herkunft von Haselnüssen auch für die verarbeitende Industrie, die bereits geröstete Haselnüsse erhält, anwendbar sowie für den Authentizitätsnachweis von Haselnüssen in verarbeiteten Produkten durch die Lebensmittelbehörden einsetzbar.

6.3.1.3 Quantifizierung von Haselnussmischungen

Die identifizierten Indels sollten nicht nur für die qualitative Unterscheidung zwischen den Gruppen „Insertion vorhanden“ und „Insertion nicht vorhanden“ herangezogen werden, sondern dienen ebenfalls als Grundlage für eine Methode zur Identifizierung von Mischungen von Haselnüssen unterschiedlicher Herkunft.

Für die Methodenentwicklung wurden Mischungen von georgischen (Nr. S4) und italienischen Haselnüssen der Sorte T. G. Trilobata (Nr. 9) eingesetzt. Hierbei handelt es sich um ein typisches Beispiel einer Verfälschung teurer durch preisgünstigere Ware. Für die Erstellung der Kalibrierreihe wurden insgesamt drei Standardreihen mit 2,5%, 5%, 10%, 20%, 40%, 60% und 80% Beimischung georgischer Haselnüsse analysiert. Zur Überprüfung der Kalibriergerade und zur Beurteilung der

Genauigkeit der Ergebnisse wurde zusätzlich noch ein Test-Set bestehend aus Mischungen derselben Haselnüsse mit 5%, 10%, 30%, 50% und 70% Beimischung georgischer Haselnüsse jeweils als Dreifachbestimmung hergestellt. Die Durchführung erfolgte wie unter 7.11.1.2 beschrieben. Die Detektion wurde sowohl mittels CGE, als auch AGE mit anschließender Graustufenanalyse durchgeführt. Für die Graustufenanalyse der Agarosegele wurde das Bildbearbeitungsprogramm ImageJ verwendet.

Detektion mittels CGE

Bei der Analyse der ersten Kalibrierreihe auf Basis des Bereichs C waren jedoch im Elektropherogramm drei anstatt zwei Peaks zu erkennen (siehe Abbildung 16A). Die Länge des undefinierten Fragments war nicht eindeutig zu bestimmen, da hier Schwankungen zwischen den Messungen auftraten. Bei Vorversuchen, bei denen ein erster linearer Zusammenhang zwischen Peakfläche und Beimischung untersucht wurde, wies das Elektropherogramm nur zwei Peaks (Fragment I: 494 bp, Fragment II: 609 bp) auf. Hierbei wurden Amplifikate des Bereichs C der beiden genannten Sorten separat hergestellt und anschließend in den genannten Verhältnissen gemischt und mittels CGE analysiert (siehe Abbildung 16B).

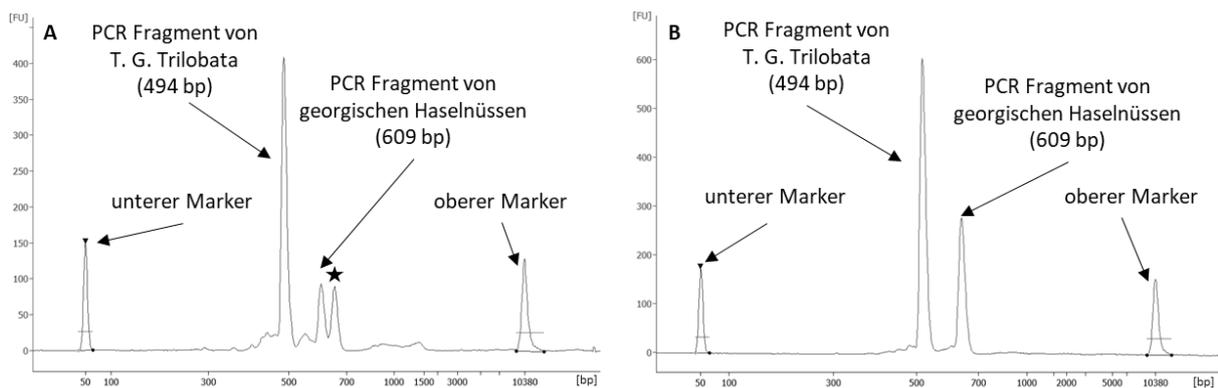


Abbildung 16: A: Elektropherogramm der Amplifikate von Bereich C einer Matrixmischung mit einem Anteil von 40% georgischer Haselnüsse. Neben den beiden gewünschten Amplifikaten der georgischen Haselnüsse und der *T. G. Trilobata* ist ein dritter Peak (*) eines unbekanntes Nebenprodukts zu erkennen. B: Elektropherogramm einer Mischung separat hergestellter Amplifikate der genannten Haselnussproben von Bereich C. Es sind nur die beiden zu erwartenden Amplifikate zu erkennen. Beide Elektropherogramme weisen zusätzlich zwei Marker (50 bp und 10380 bp) für die Größenbestimmung auf.

Die Mischung der Amplifikate zeigte damit das gewünschte Ergebnis. Das PCR-Produkt der Matrixmischung hingegen wies ein unerwünschtes Nebenprodukt auf. Die gleichen Versuche wurden auch mit Amplifikaten des Bereichs B durchgeführt. Hier zeigten sich dieselben Ergebnisse. Damit kann ein sequenz- oder primerspezifisches Problem ausgeschlossen werden. Daher wurde vermutet, dass es sich bei dem dritten Peak um eine Tertiärstruktur handelt, deren Bildung durch das Temperaturprogramm der PCR begünstigt wird. Für die Bestätigung dieser Annahme wurde erneut eine 50/50 Mischung der zuvor separat hergestellten Amplifikate der beiden Sorten angesetzt. Eine Hälfte des Ansatzes wurde kühl bis zur weiteren Verwendung gelagert und die andere Hälfte wurde für 5 Minuten auf 95 °C erhitzt und langsam auf Raumtemperatur abgekühlt. Anschließend wurden beide Ansätze mittels CGE analysiert (siehe Abbildung 17).

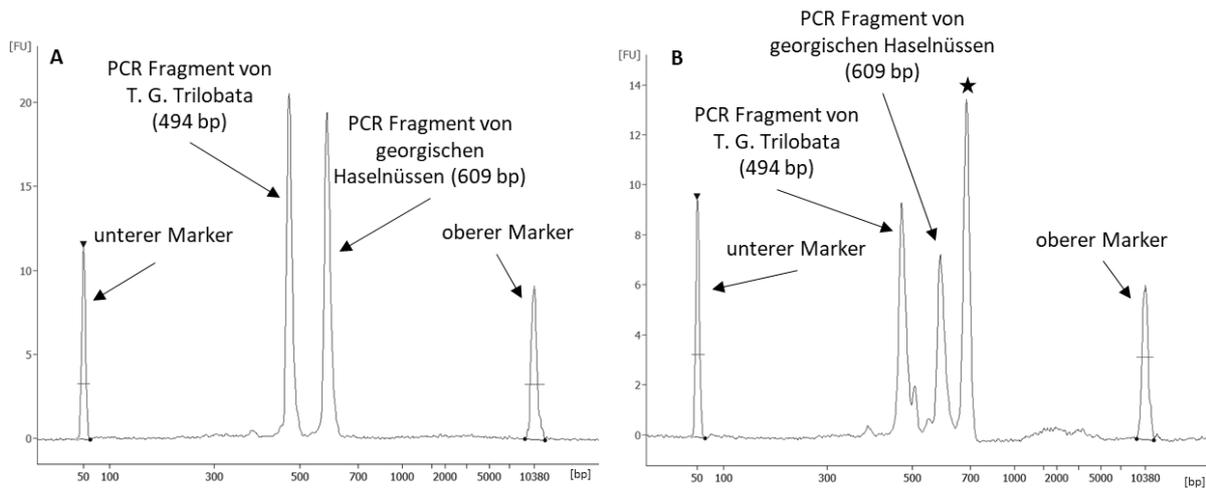


Abbildung 17: A: Elektropherogramm einer 50/50 Mischung separat hergestellter Amplifikate der italienischen und georgischen Haselnussproben von Bereich C. Zu erkennen sind zwei Peaks, die den beiden Amplifikaten zuzuordnen sind. B: Elektropherogramm einer 50/50 Mischung separat hergestellter Amplifikate der genannten Haselnussproben von Bereich C. Vor der Messung wurde die Mischung für 5 Minuten auf 95 °C erhitzt und anschließend langsam auf Raumtemperatur abgekühlt. Neben den erwarteten Fragmenten wurde zusätzlich ein drittes Fragment detektiert (*). Beide Elektropherogramme weisen zusätzlich jeweils zwei Marker (50 bp und 10380 bp) für die Größenbestimmung auf.

Wie deutlich zu erkennen ist, wurde durch das kurzzeitige Erhitzen der Mischung der beiden Amplifikate das unbekannte Fragment gebildet. Es wird angenommen, dass sich eine Art *hairpin*-Struktur, wie in Abbildung 18 dargestellt, ausbildet. Durch die langen homologen Bereiche, die das Insert flankieren, kann sich ein Einzelstrang der Amplifikate der georgischen Haselnuss-DNA und ein Einzelstrang der Amplifikate der DNA-Isolate italienischer Haselnüsse aneinander lagern, wobei das Insert eine Ausstülpung bildet. Werden die homologen Arme jedoch verkürzt, sollte dies zu einer Instabilität dieser Sekundärstruktur führen, wodurch die Bildung dieser nicht mehr begünstigt wird. Deshalb wurde für den Bereich B ein neues Primerpaar designt, mit dem Amplifikate mit einer Länge von 84 bzw. 202 bp amplifiziert wurden (siehe Abbildung 56). Als Basis wurde hierbei der Bereich B ausgewählt, da die Definition neuer sequenzspezifischer Primer aufgrund der A/T-reichen Sequenz nur eingeschränkt möglich war. Erste Test-Messungen mit dem neuen Primerpaar bestätigten die Theorie der geschilderten Sekundärstruktur, da keine zusätzlichen Peaks im Elektropherogramm zu erkennen waren. Die fortführenden Analysen wurden daher mit dem neuen Primerpaar (B.2_FW/RV) auf Basis des Bereichs B durchgeführt. Dieses ist natürlich ebenfalls für die unter 6.3.1.1 vorgestellte qualitative Analyse geeignet.

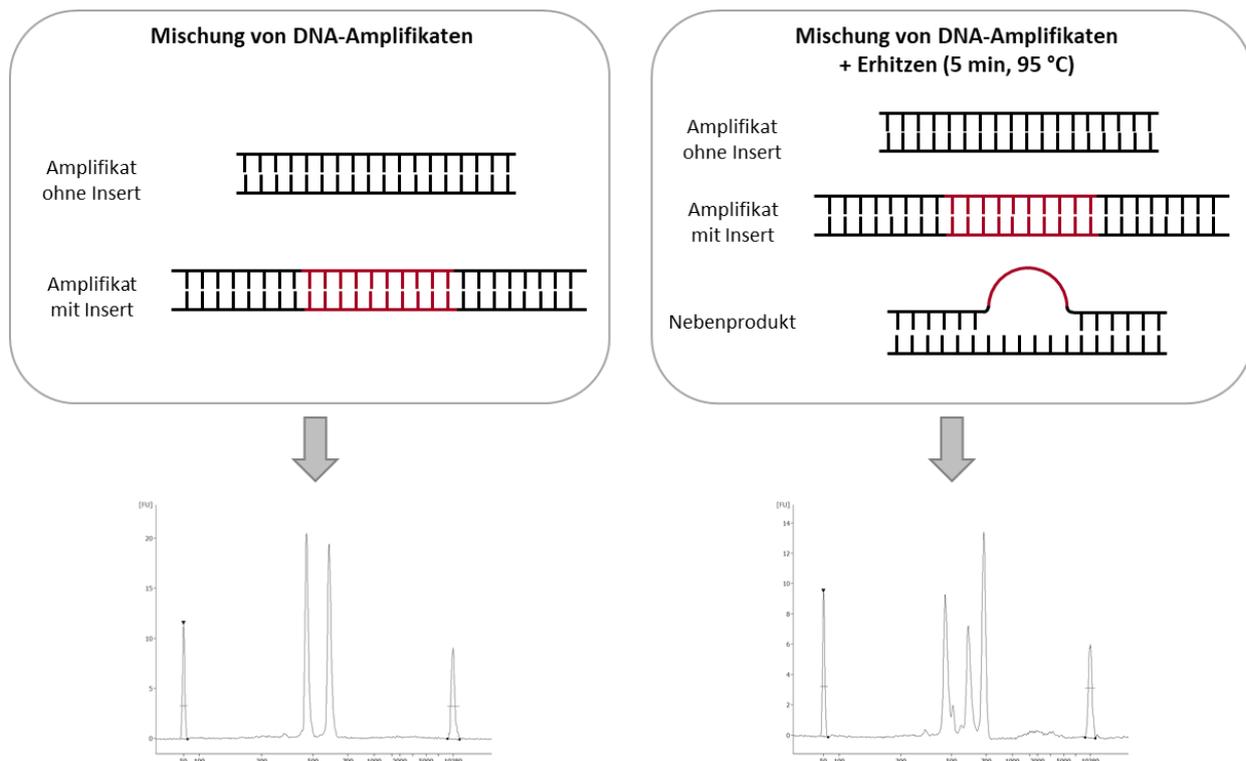


Abbildung 18: Schematische Darstellung der Bildung einer Sekundärstruktur. Links: Mischung von Amplifikaten mit und ohne Insert und das daraus resultierende Elektropherogramm nach Messung mittels CGE. Rechts: Mischung derselben Amplifikate + kurzzeitiges Erhitzen (5 min) auf 95 °C. Das daraus resultierende Elektropherogramm zeigt einen zusätzlichen Peak, der auf eine Sekundärstruktur mit der gezeigten vorgeschlagenen Struktur zurückzuführen ist.

Wie zu Beginn des Kapitels und unter 7.11.1.2 erläutert, wurden insgesamt drei Kalibrierreihen angesetzt und analog aufgearbeitet, um die Reproduzierbarkeit der Messungen zu verifizieren. Die Detektion erfolgte sowohl mittels CGE, als auch AGE mit anschließender Graustufenanalyse. Im Folgenden werden zunächst die Ergebnisse der Messung mittels CGE und anschließend die der Graustufenanalyse vorgestellt. Neben der Entwicklung des neuen Primerpaares wurden auch die Reaktionsbedingungen optimiert. Dies geschah v.a. hinsichtlich der Detektion mittels AGE, daher wird darauf im Folgenden bei der „Detektion mittels AGE und anschließender Graustufenanalyse“ detaillierter eingegangen.

Die Kalibriergerade wurde erstellt, indem die prozentuale relative Area des längeren Fragments (202 bp), das den georgischen Haselnüssen zuzuordnen ist, gegen den prozentualen Anteil an georgischen Haselnüssen in der Mischung aufgetragen wird. Die relative Area wird durch das Verhältnis der Area des langen Fragments zur Gesamt-Area erhalten. Dieser Schritt der Normalisierung ist notwendig, um mögliche Konzentrationsunterschiede vor der Messung auszugleichen. Die erhaltene Kalibriergerade ist Abbildung 19 zu entnehmen. Die Grafik enthält sowohl die einzelnen Messpunkte der Dreifachbestimmung (grau) sowie den zugehörigen Mittelwert (rot) und die daraus resultierende Kalibriergerade. Eine tabellarische Zusammenfassung der Messwerte ist der Tabelle 40 im Anhang unter 9.10.4.2 zu entnehmen.

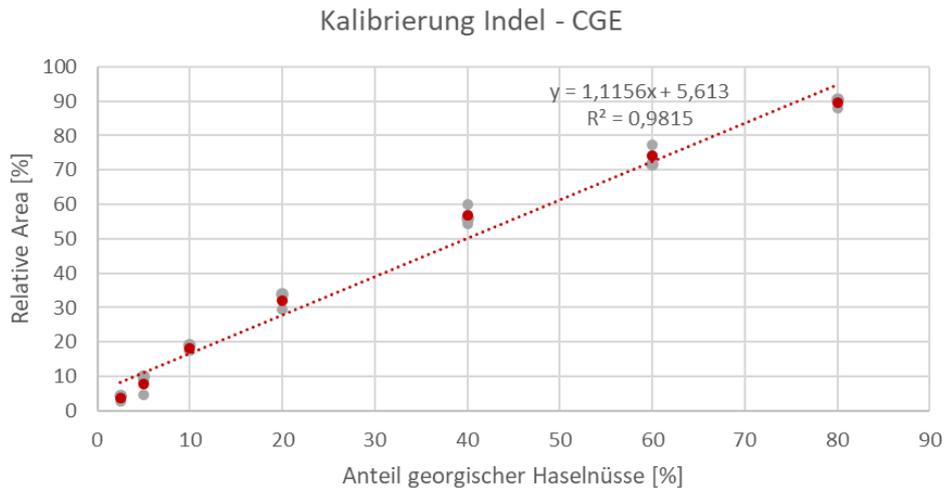


Abbildung 19: Kalibriergerade resultierend aus der Analyse der drei Kalibriersets und Detektion mittels CGE. Aufgetragen ist die prozentuale relative Area des langen Fragments (zugehörig zu Haselnüssen aus Georgien) gegen den Anteil georgischer Haselnüsse. In Grau dargestellt sind die einzelnen Messpunkte der Dreifachbestimmung. In Rot dargestellt sind die zugehörigen Mittelwerte mit der daraus resultierenden Kalibriergerade.

Für die Bewertung der Güte der erstellten Kalibriergerade wurden einige statistische Kenngrößen berechnet. Die zugehörigen Formeln sind im Anhang unter 9.9 aufgeführt. Zunächst wurde der Korrelationskoeffizient r berechnet, der ein Maß für den linearen Zusammenhang zweier Messwerte darstellt. Im Falle eines eindeutigen linearen Zusammenhangs ist der Korrelationskoeffizient $+1$ bzw. -1 , abhängig von der Steigung der Kalibriergerade. Für die in Abbildung 19 dargestellten Kalibriergerade ist $r = 0,9907$ und zeigt damit, dass ein linearer Zusammenhang zwischen dem Anteil georgischer Haselnüsse und der prozentualen relativen Area des zugehörigen Fragmentpeaks besteht. Die Standardabweichung (σ) innerhalb der Dreifachbestimmungen liegt im Bereich von 0,69 bis 2,36. Der Variationskoeffizient liegt bei den Mischungen mit 10-80% georgischer Haselnüsse zwischen 1,25 und 5,55%. Die Mischungen mit den zwei geringsten Anteilen georgischer Haselnüsse (2,5 und 5%) weisen jedoch stark erhöhte Variationskoeffizienten von 19,11 und 28,09% auf. Diese Werte verdeutlichen eine deutlich erhöhte Schwankung innerhalb der Dreifachbestimmung im Bereich niedrig konzentrierter Mischungen. Zusätzlich wurden die Reststandardabweichung $s_y = 5,0$, die Verfahrensstandardabweichung $s_{x0} = 4,5$ sowie der Verfahrensvariationskoeffizient bzw. relative Verfahrensstandardabweichung $v_{x0} = 14,5\%$, berechnet. Die Reststandardabweichung stellt dabei ein Maß für die Streuung der Messwerte um die Regressionsgerade dar. Darauf aufbauend können, bezogen auf die gesamte Methode, die Verfahrensstandardabweichung und der Verfahrensvariationskoeffizient berechnet werden. Aufgrund der kalkulierten Werte ist davon auszugehen, dass anhand der Kalibrierung lediglich eine semi-quantitative Abschätzung von Mischungen möglich ist.

Um besser beurteilen zu können wie sich diese Größen auf die Genauigkeit der Quantifizierung auswirken wurde ein Test-Set bestehend aus fünf unterschiedlichen Mischungsverhältnissen (5%, 10%, 30%, 50% und 70%) ebenfalls in Dreifachbestimmung analog zu den Kalibriersets analysiert und anhand der dargestellten Kalibriergerade der Anteil georgischer Haselnüsse quantifiziert. In Tabelle 7 sind die Ergebnisse zusammengefasst.

Tabelle 7: Ergebnisse des Test-Sets. Die Berechnung erfolgte auf Basis der Kalibriergerade in Abbildung 19 nach Detektion mittels CGE. GE = georgische Haselnüsse, TGT = italienische Haselnüsse der Sorte T. G. Trilobata.

	GE [%]	TGT [%]	Berechneter Gehalt georgischer Haselnüsse [%]	Mittelwert [%]	Variationskoeffizient [%]	Wiederfindungsrate [%]
T5-1			2,1			
T5-2	5	95	4,1	4,9	54,8	98,0
T5-3			8,6			
T10-1			11,1			
T10-2	10	90	9,3	11,2	14,8	112,0
T10-3			13,3			
T30-1			39,9			
T30-2	30	70	33,2	37,4	8,0	124,7
T30-3			39,2			
T50-1			59,9			
T50-2	50	50	55,5	56,3	4,6	112,6
T50-3			53,6			
T70-1			69,3			
T70-2	70	30	70,9	68,9	2,6	98,4
T70-3			66,6			

Die Variationskoeffizienten zeigen, dass auch beim Test-Set im Bereich gering konzentrierter Beimischungen eine deutlich höhere Schwankung innerhalb der Dreifachbestimmung vorliegt. Die Schwankungen werden mit zunehmendem Anteil georgischer Haselnüsse kleiner. Die Wiederfindungsrate liegt zwischen 98,0 und 124,7%.

Sowohl die Ergebnisse der Beurteilung der Kalibriergerade sowie die Resultate für das Test-Set zeigen, dass im Bereich gering konzentrierter Beimischungen die Schwankungen deutlich größer sind. Dennoch zeigen die Ergebnisse auch, dass eine semi-quantitative Abschätzung von Mischungen möglich ist und diese Methodik somit eine bessere Beurteilung von Verfälschungen erlaubt. Grundsätzlich ist für die quantitativen Ansätze die vorherige DNA-Isolierung mittels der unter 7.2.2.1 beschriebenen CTAB-Methode zu empfehlen, da mit dieser Verunreinigungen durch PCR-störende oder -inhibierende Substanzen besser entfernt und damit reproduzierbarere Ergebnisse erzielt werden.

Für die Bestimmung des *Limit of Detection* (LOD) und *Limit of Quantitation* (LOQ) wurden Messungen mittels CGE von PCR-Produkten durchgeführt, die nur auf der DNA italienischer Haselnüsse beruhen. Es handelt sich damit in Bezug auf die Quantifizierung georgischer Haselnüsse um Blank-Messungen. Diese wiesen damit dieselbe Matrix, aber nicht das lange Fragment mit Insert auf. Bestimmt wurde das Grundrauschen im Bereich von ca. 180-210 bp, also an der Stelle, wo sonst das lange Fragment detektiert wird. Die Analyse erfolgte ebenfalls in Dreifachbestimmung. Die erhaltenen Areas wurden anschließend ebenfalls in die entsprechenden relativen Areas umgerechnet.

Die Berechnung des LOD und LOQ erfolgte anhand der unter 7.11.1.2 dargestellten Formeln. Daraus resultieren der LOD und LOQ ebenfalls als prozentuale relative Area. Für die Umrechnung in den prozentualen Anteil georgischer Haselnüsse wurde aufgrund der geringen Genauigkeit im Bereich niedriger Konzentrationen nicht die Kalibriergerade genutzt. Stattdessen wurden die erhaltenen Werte mit den Daten der drei Kalibriersets (siehe Tabelle 40) verglichen. Daraus ergab sich ein LOD im Bereich von 2,5% Beimischung und ein LOQ zwischen 2,5 und 5%. Unter Berücksichtigung der visuellen Bewertung der Elektropherogramme und den Ergebnissen des Test-Sets wurde der LOD auf 5% und der LOQ auf 10% bezogen auf den Anteil georgischer Haselnüsse erhöht. Wird auf Basis dieser Ergebnisse die Kalibriergerade auf den Bereich von 10-80% georgischer Haselnüsse angepasst, so bestätigt sich der Korrelationskoeffizient von 0,993 und s_y verbessert sich leicht auf 4,0, s_{x0} auf 3,9 und v_{x0} auf 9,4%.

Bei der Bewertung dieser Ergebnisse ist zu beachten, dass die verwendeten Chips für die kapillargelelektrophoretische Trennung, wie unter 7.5.2 beschrieben, gereinigt und für mehrere Messungen wiederverwendet wurden. Sowohl für die Messungen der Kalibriergeraden, als auch des Test-Sets und die Bestimmung des LOD und LOQ wurden mehrfach genutzte Chips eingesetzt, um vergleichbare Reaktionsbedingungen zu erhalten. Es ist jedoch davon auszugehen, dass sich durch eine einmalige Nutzung der Chips die Vergleichbarkeit der Messungen erhöhen und der LOD und LOQ verringern würden, da im Rahmen der Analysen eine Zunahme des Grundrauschens durch die wiederholte Nutzung zu beobachten war. Da jedoch auch mit den wiederverwendeten Chips gute Ergebnisse erzielt wurden, stellte die Reinigung eine gute und einfache Möglichkeit dar, die Analysenkosten deutlich zu reduzieren.

Detektion mittels AGE und anschließender Graustufenanalyse

Neben der Detektion mittels CGE wurde zusätzlich auch die AGE mit anschließender Graustufenanalyse der Gelbilder für die semi-quantitative Abschätzung herangezogen.

Die Durchführung erfolgte wie unter 7.11.1.2 beschrieben mit dem optimierten Primerpaar B.2_FW/RV. Für eine optimale Detektion über ein Agarosegel wurde der PCR-Ansatz dahingehend optimiert, dass die Konzentration an Primern deutlich verringert wurde. Dieser Schritt wurde durchgeführt, da das kurze PCR-Fragment, das den italienischen Haselnüssen zuzuordnen ist, sich aufgrund seiner Länge nur schwer von den überschüssigen Primern abtrennt und von diesen überlagert wird. Eine Aufreinigung des PCR-Ansatzes vor der AGE würde dieses Problem natürlich lösen. Doch das Ziel der Quantifizierung über ein Agarosegel mit entsprechender Auswertung war, eine zeit- und v.a. auch kostengünstige Alternative zu der CGE zu entwickeln. Aus diesem Grund wurde der Fokus auf die Optimierung des Reaktionsansatzes gelegt, sodass eine vorherige Aufreinigung zumindest für die AGE nicht notwendig ist. Des Weiteren wurde eine DNA Polymerase mit einer höheren Genauigkeit eingesetzt und die Anzahl an PCR-Zyklen auf 45 erhöht. Die PCR-Produkte wurden mittels AGE getrennt. Die optimierten Reaktionsbedingungen wurden für alle PCR-Ansätze der Kalibriersets sowie des Test-Sets sowohl für die Detektion mittels CGE, als auch AGE

verwendet. Die Gelbilder wurden schließlich mit Hilfe des Bildbearbeitungsprogramms ImageJ wie unter 7.5.1 beschrieben ausgewertet. Abbildung 20 zeigt beispielhaft das Gelbild eines der drei Kalibriersets und die zugehörigen grafischen Darstellungen für die Mischungen mit 2,5%, 5%, 10% und 20% georgischen Haselnüssen.

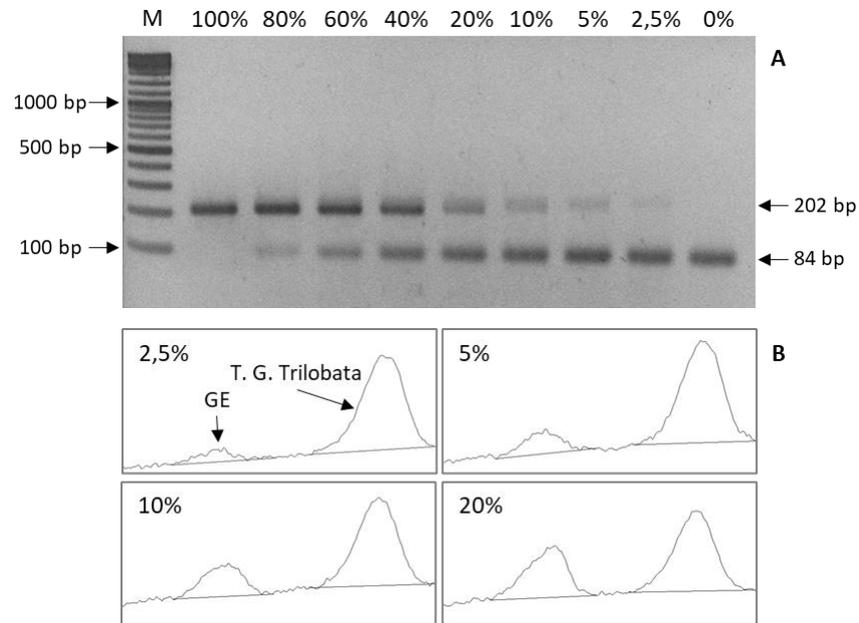


Abbildung 20: Graustufenanalyse eines Gelbildes. A: Gelbild des Kalibriersets 1, B: Grafische Darstellung der Graustufen für die vier niedrigsten Konzentrationen georgischer Haselnüsse. Es sind jeweils zwei Peaks zu erkennen. Der linke Peak entspricht der oberen Bande (202 bp) auf dem Gelbild und ist den georgischen Haselnüssen zuzuordnen und der rechte Peak ist der unteren Bande (84 bp) und damit den italienischen Haselnüssen der Sorte T. G. Trilobata zuzuordnen. GE = Georgien, M = DNA-Leiter

Alle drei Kalibriersets wurden analog vorbereitet, amplifiziert und anschließend wie beschrieben ausgewertet. Die erhaltenen Werte für die relative Area wurden im Anhang unter 9.10.4.2 in Tabelle 41 zusammengefasst. Für die Kalibriergerade wurden, wie bereits bei der Detektion mittels CGE beschrieben, die prozentualen relativen Areas des langen Fragments (+ Insert) gegen den Anteil georgischer Haselnüsse aufgetragen. Die daraus resultierende Kalibriergerade ist der Abbildung 21 zu entnehmen.

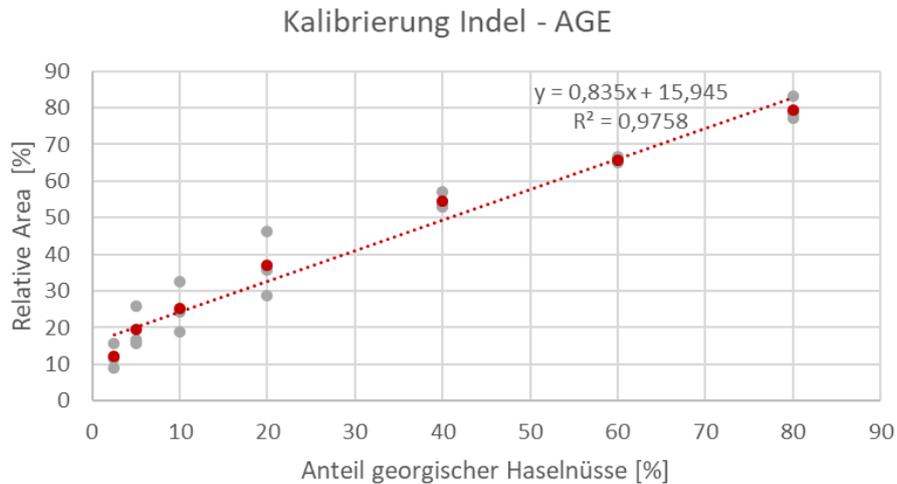


Abbildung 21: Kalibriergerade resultierend aus der Analyse der drei Kalibriersets und Detektion mittels AGE + Graustufenanalyse. Aufgetragen ist die prozentuale relative Area des langen Fragments mit 202 bp (zugehörig zu Haselnüssen aus Georgien) gegen den Anteil georgischer Haselnüsse. In Grau dargestellt sind die einzelnen Messpunkte der Dreifachbestimmung. In Rot dargestellt sind die zugehörigen Mittelwerte mit der daraus resultierenden Kalibriergerade.

Die Kalibriergerade zeigt mit einem Korrelationskoeffizienten von $r = 0,9878$ ebenfalls einen linearen Zusammenhang zwischen dem Anteil georgischer Haselnüsse und der resultierenden relativen Area. Es ist jedoch auch zu erkennen, dass in dem Bereich von 2,5-20% Beimischung die Schwankungen innerhalb der Dreifachbestimmungen deutlich größer sind, als im oberen Bereich der Kalibriergerade. Dies verdeutlicht auch der Variationskoeffizient, der im unteren Bereich 19,62-23,85% beträgt wohingegen er bei 40-80% Beimischung nur bei 0,91-3,53% liegt. Die Reststandardabweichung s_y beträgt 4,3, die Verfahrensstandardabweichung s_{x0} 5,2 und der Verfahrensvariationskoeffizient v_{x0} liegt bei 16,6%. Basierend auf diesen Ergebnissen, ist bei Berechnungen anhand dieser Kalibriergeraden ebenfalls von einer semi-quantitativen Abschätzung auszugehen.

Für eine genauere Bewertung quantitativer Ergebnisse basierend auf der erstellten Kalibriergerade, wurde ebenfalls das bereits zuvor beschriebene Test-Set analog zu den Kalibriersets analysiert. Die Ergebnisse sind in Tabelle 8 zusammengefasst.

Tabelle 8: Ergebnisse des Test-Sets. Die Berechnung erfolgte auf Basis der Kalibriergerade in Abbildung 21 nach Detektion mittels AGE + Graustufenanalyse. GE = georgische Haselnüsse, TGT = italienische Haselnüsse der Sorte Tonda Gentile Trilobata.

	GE [%]	TGT [%]	Berechneter Gehalt georgischer Haselnüsse [%]	Mittelwert [%]	Variations- koeffizient [%]	Wieder- findungsrate [%]
T5-1			-4,9			
T5-2	5	95	0,1	-1,5	-163,5	-29,7
T5-3			0,4			
T10-1			8,4			
T10-2	10	90	9,1	9,5	11,1	95,0
T10-3			10,9			
T30-1			36,2			
T30-2	30	70	32,5	35,1	5,3	117,0
T30-3			36,6			
T50-1			52,0			
T50-2	50	50	49,7	49,8	3,6	99,6
T50-3			47,6			
T70-1			64,4			
T70-2	70	30	64,2	64,1	0,5	91,5
T70-3			63,7			

Die Ergebnisse zeigen, dass eine ausreichende Genauigkeit im Bereich kleiner Konzentration nicht gegeben ist. So ergibt sich ein negativer Mittelwert für die Mischungen mit 5% georgischen Haselnüssen. Im Gegensatz dazu zeigen jedoch die Mischungen mit 10% georgischen Haselnüssen einen Variationskoeffizient von nur 11,1% und eine Wiederfindungsrate von 95%. Damit ist die Schwankung innerhalb der Dreifachbestimmung deutlich geringer im Vergleich zu den drei Matrixmischungen mit 10% georgischen Haselnüssen der drei Kalibriersets. Hier betrug der Variationskoeffizient 22,5%. Die Schwankungen innerhalb der Dreifachbestimmungen bei höheren Konzentrationsbereichen sind vergleichbar zu den Ergebnissen der Kalibriersets. Ohne Betrachtung der Proben des Test-Sets mit 5% Beimischung liegen die Wiederfindungsraten zwischen 91,5 und 117,0%.

Der LOD und LOQ für die semi-quantitative Abschätzung mittels Graustufenanalyse wurde wie bereits zuvor für die Detektion mittels CGE beschrieben durchgeführt. Der Mittelwert ($\bar{x}_{0\%}$) der relativen Area der Dreifachbestimmung, die durch die Matrix verursacht wurde, betrug 2,88% und die Standardabweichung $s_{0\%}$ lag bei 0,63. Daraus folgte ein berechneter LOD von 4,8% und ein LOQ von 9,2% relative Area. Auch hier erfolgte die Umrechnung in den Anteil georgischer Haselnüsse nicht anhand der Kalibriergeraden, aufgrund einer zu hohen Ungenauigkeit im Bereich niedriger Konzentration, sondern ebenfalls durch den Vergleich der Werte mit den Daten der Kalibriersets (siehe Tabelle 41). Auf Basis derer würde sowohl der LOD, als auch der LOQ im Bereich von 2,5% Beimischung liegen. Unter Berücksichtigung der visuellen Bewertung der Gelbilder sowie der Ergebnisse des Test-Sets wurde der LOD auf 5% und der LOQ auf 10% erhöht. Erfolgt auch hier eine

Anpassung der Kalibrierung auf den Bereich von 10-80% verbessert sich der Korrelationskoeffizient auf 0,9931, s_y auf 2,9, s_{x0} auf 3,9 und v_{x0} auf 9,3%.

Die Ergebnisse verdeutlichen, dass auch über die Graustufenanalyse nach Detektion mittels AGE eine semi-quantitative Abschätzung erfolgen kann und durchaus vergleichbare Ergebnisse zu der CGE erhalten werden. Diese Detektionsmethode kann damit als zeit- und kosteneffiziente Alternative genutzt werden um die Menge beigemischter nicht deklarerter Haselnüsse abzuschätzen.

6.3.2 PCR-RFLP für den Nachweis georgischer Haselnüsse

6.3.2.1 Qualitativer Nachweis

Wie bereits unter 6.2 erläutert wurden im Rahmen der Resequenzierung im Bereich D zwei Einzelbasenaustausche (SNP 5 und 7) identifiziert, die die geografische Gruppierung widerspiegeln wie sie auch auf Basis der Indels 1 und 3 in den Bereichen B und C möglich ist. Zusätzlich wurde ein C/A-Austausch (SNP 6) nachgewiesen, der bei Haselnüssen georgischen Ursprungs in der Erkennungssequenz des Restriktionsenzym *MfeI* resultiert. Damit ist ein spezifischer Nachweis georgischer Haselnüsse mittels PCR-RFLP möglich. Abbildung 22 zeigt einen Ausschnitt eines Alignments des Bereichs D, dem die genannten Unterschiede zu entnehmen sind. Das vollständige Alignment ist unter 9.10.2 aufgeführt.

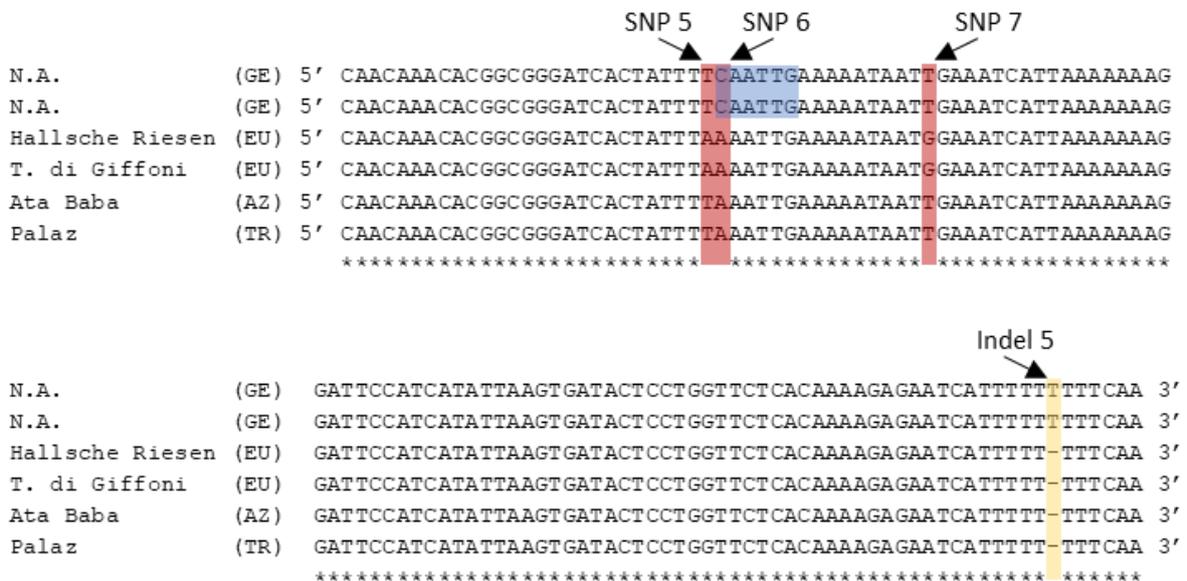


Abbildung 22: Ausschnitt des Alignments von Bereich D der Sorten Palaz, Ata Baba, Tonda di Giffoni, Hallsche Riesen sowie zwei georgischen Haselnussproben. Die Einzelbasenaustausche SNP 5-7 sind rot, Indel 5 gelb und die Restriktionsschnittstelle blau hinterlegt. GE = Georgien, EU = Europa, AZ = Aserbaidschan, TR = Türkei.

Für die qualitative Analyse wurden sowohl DNA-Isolate verwendet, die mittels der CTAB-Methode isoliert wurden, als auch Isolate, die mittels der *Rapid Extraction* hergestellt wurden. Detaillierte Informationen zu den Reaktionsbedingungen der PCR sowie des enzymatischen Verdau sind Abschnitt 7.11.2.1 zu entnehmen.

Das Ausgangstemplate, welches mit dem Primerpaar D.2_FW/RV hergestellt wurde, ist 619 bp lang. Durch den anschließenden enzymatischen Verdau mit dem Restriktionsenzym *MfeI* wird dieses geschnitten, wodurch zwei Fragmente entstehen. Fragment 1 mit einer Länge von 206 bp und Fragment 2 mit einer Länge von 413 bp. Abbildung 23 zeigt, dass nur die PCR-Produkte, die aus DNA-Isolaten georgischer Haselnüsse (Nr. S5, 162-165, 151, 152) amplifiziert wurden, von dem Enzym geschnitten wurden. Amplifikate der Proben anderer Herkunft und Sorte (Nr. S2, 109, 140) blieben hingegen intakt. Es ist zudem zu erkennen, dass DNA-Isolate beider Isolierungsmethoden für die Analyse geeignet sind und auch die Isolate der *Rapid Extraction* (Nr. 162-165) trotz verringerter Reinigungsschritte keine störenden Begleitsubstanzen aufweisen, die den enzymatischen Verdau hemmen.

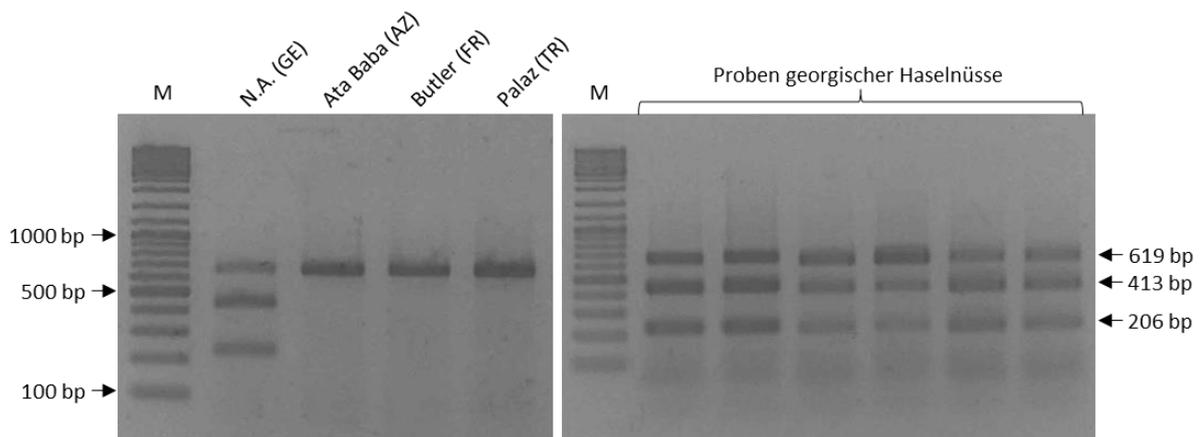


Abbildung 23: Ergebnisse des enzymatischen Verdau von Amplifikaten des Bereichs D mit dem Restriktionsenzym *MfeI*. Bei den Proben georgischer Haselnüsse sind neben dem Template (619 bp) jeweils die zwei Fragmente mit einer Länge von 206 und 413 bp zu erkennen. Proben anderer Sorte/Herkunft werden von dem Enzym nicht geschnitten. M = DNA-Leiter, GE = Georgien, AZ = Aserbaidschan, FR = Frankreich, TR = Türkei.

Die Abbildung zeigt jedoch auch, dass der enzymatische Verdau nicht vollständig abläuft, da bei allen georgischen Haselnussproben neben den beiden Fragmenten auch noch das Ausgangstemplate (Länge: 619 bp) zu erkennen ist.

Einige der analysierten Proben basierten dabei auf DNA-Isolaten, die aus nur einer Nuss gewonnen wurden. Damit können Mischungen von Haselnüssen unterschiedlicher Sorte bzw. Herkunft als Ursache für den unvollständigen enzymatischen Verdau ausgeschlossen werden. Eine Überprüfung

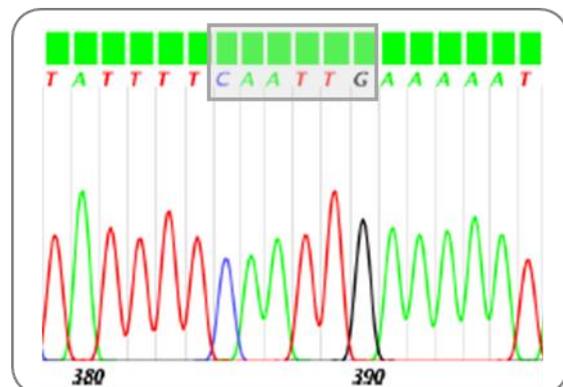


Abbildung 24: Ergebnis der Sanger-Sequenzierung der Probe Nr. S5. Ausschnitt des Elektropherogramms. Die Restriktionsschnittstelle ist grau hervorgehoben.

der Sanger-Sequenzierung anhand der Elektropherogramme (siehe Abbildung 24) zeigte, dass im Bereich der Restriktionsschnittstelle keine Signalüberlagerungen zu erkennen sind und damit das Vorliegen eines heterozygoten SNPs, als weitere mögliche Ursache, ebenfalls ausgeschlossen werden konnte. Da auf Basis des Probenmaterials und der Sequenz keine Ursache für den unvollständigen enzymatischen Verdau identifiziert werden konnten, wurden anschließend die Reaktionsbedingungen optimiert. Hierbei wurden folgende Optimierungen einzeln und in Kombination getestet:

- Verlängerung der Inkubationszeit auf bis zu 24 Stunden
- Variation der eingesetzten Menge des PCR-Produkts für den enzymatischen Verdau
- Variation des Reaktionsvolumens
- Variation der Enzym-Konzentration
- Aufreinigung des PCR-Produkts vor dem enzymatischen Verdau
- Verkürzung des Amplifikats

Keiner der Optimierungsschritte zeigte eine wesentliche Verbesserung. Ein vollständiger Verdau des Templates konnte nicht erzielt werden.

Neben den Reaktionsbedingungen kann auch die Sequenz und die Position des Sequenzunterschieds die Güte des enzymatischen Verdau beeinflussen und letztendlich zu einem unvollständigen Verdau führen. Bereits 1990 beschäftigten sich R.S. Jack und H. Eggert mit den Auswirkungen der Position einer Restriktionsschnittstelle im Genom auf die Zugänglichkeit für Restriktionsenzyme und der damit verbundenen Effizienz des enzymatischen Verdau.^[219] Des Weiteren gibt es bereits verschiedene Veröffentlichungen, auch im Bereich der Authentizitätsbestimmung von Lebensmitteln, die ebenfalls das Phänomen eines unvollständigen Verdau beschreiben ohne hierfür eine Ursache finden zu können oder das Problem durch eine Optimierung der Reaktionsbedingung beheben zu können.^{[220]–[223]} So entwickelte Meyer et al. (1995) eine PCR-RFLP basierend auf einem Fragment des Cytochrom B Gens auf dem mitochondrialen Genom für die Unterscheidung verschiedener Fleischarten. Hierbei zeigte sich ein unvollständiger Verdau bei dem Amplifikat von Pferdefleisch mit dem Enzym *AluI*.^[221] Céspedes et al. (1998) nutzte ebenfalls Sequenzunterschieden innerhalb des Cytochrom B Gens für die Unterscheidung verschiedener Plattfischarten, wie Seezunge, Flunder und Scholle. Hierbei wurde nach dem enzymatischen Verdau des Amplifikats der Seezunge mit dem Enzym *Sau3AI* stets noch das Ausgangstemplate detektiert.^[222] Und auch Fernández et al. (2001) berichteten bei der Entwicklung einer PCR-RFLP zur Unterscheidung dreier zum Verzehr geeigneter Muschelarten von diesem Phänomen. Es war ebenfalls eine der untersuchten Spezies betroffen und in diesem Fall der Verdau mit dem Enzym *HinfI*.^[223] In allen Publikationen konnte das Vorliegen von nicht authentischem Probenmaterial ausgeschlossen werden. Das beschriebene Problem zeigte sich also bei unterschiedlichen Matrizes und damit auch verschiedenen Sequenzen unterschiedlicher Längen sowie variierenden Restriktionsenzymen. Dies spricht dafür, dass es weitere Faktoren wie z.B. sterische Einflüsse, geben muss, die die Effizienz eines enzymatischen Verdau beeinflussen.

Da jedoch trotz des bisher beobachteten unvollständigen Schneidens durch das Restriktionsenzym bei Haselnussproben georgischen Ursprungs eine Fragmentierung zu beobachten ist, ist die vorgestellte PCR-RFLP für den Nachweis dieser geeignet. Dies bestätigten auch die Ergebnisse des im Folgenden analysierten erweiterten Probenpools. Dabei wurde die entwickelte Methode auf insgesamt 30 Haselnussproben, davon 15 Proben aus Georgien und 15 Proben anderer Herkunft angewendet. Ziel war zum einen die Bestätigung der Sequenzierungsergebnisse anhand weiterer Proben und zum anderen die Prüfung, ob bei allen georgischen Proben trotz des bisher beobachteten unvollständigen Verdau eine Fragmentierung zu beobachten ist. Das Ergebnis der Sequenzierung konnte bestätigt werden. Alle Proben georgischer Herkunft wurden erfolgreich geschnitten und alle anderen Proben hingegen nicht, so wie es basierend auf den Sequenzierungsergebnissen zu erwarten war. Es bestätigte sich jedoch auch die bisherige Beobachtung, dass die Haselnussproben aus Georgien nicht vollständig geschnitten werden. Bei allen analysierten Proben war auf dem Agarosegel neben den beiden Fragmenten auch das ungeschnittene Template zu sehen. Damit ist die Methode nicht geeignet, eine Streckung georgischer Haselnüsse durch die Zugabe von Haselnüssen anderer Herkunft nachzuweisen. Da es sich jedoch bei georgischen Haselnüssen im Vergleich zu anderen Anbauländern um günstigere Ware handelt, ist rein wirtschaftlich betrachtet eine Verfälschung georgischer Haselnüsse unwahrscheinlich. Die entwickelte Methode eignet sich aber sehr gut für den qualitativen Nachweis georgischer Haselnüsse, da gezeigt werden konnte, dass das Enzym das Template aufgrund des SNPs selektiv nur bei diesen angreift und die beiden Teilfragmente zu erkennen sind. Damit kann die Methode auch angewendet werden, um Beimischungen georgischer Haselnüsse zu Proben anderer Sorten bzw. Herkunft zu identifizieren, denn der Nachweis der Teilfragmente ist immer auf das Vorhandensein georgischer Haselnüsse zurückzuführen.

6.3.2.2 Nachweis in prozessierten Haselnüssen

Die entwickelte PCR-RFLP wurde ebenfalls auf ihre Anwendbarkeit auf verarbeitete Haselnüsse überprüft. Sie wurde mit zwei selbst gerösteten Haselnussproben aus Georgien (V4) und Aserbaidschan (V5) sowie kommerziell erworbenen Haselnüssen der Sorte T. G. Trilobata (V3) und Nüssen aus einer Schokolade (V6), die laut Herstellerangabe aus der Türkei stammen, durchgeführt. Das gewünschte Amplifikat des Bereichs D konnte bei allen Proben erfolgreich generiert werden. Der anschließende enzymatische Verdau ergab lediglich bei den gerösteten Haselnüssen aus Georgien das erwartete Schnittmuster. Die Amplifikate der anderen Proben blieben intakt (siehe Abbildung 25). Um einen erfolgreichen enzymatischen Verdau sicherzustellen, wurde parallel eine Probe ungerösteter georgischer Haselnüsse als Kontrolle mit analysiert.

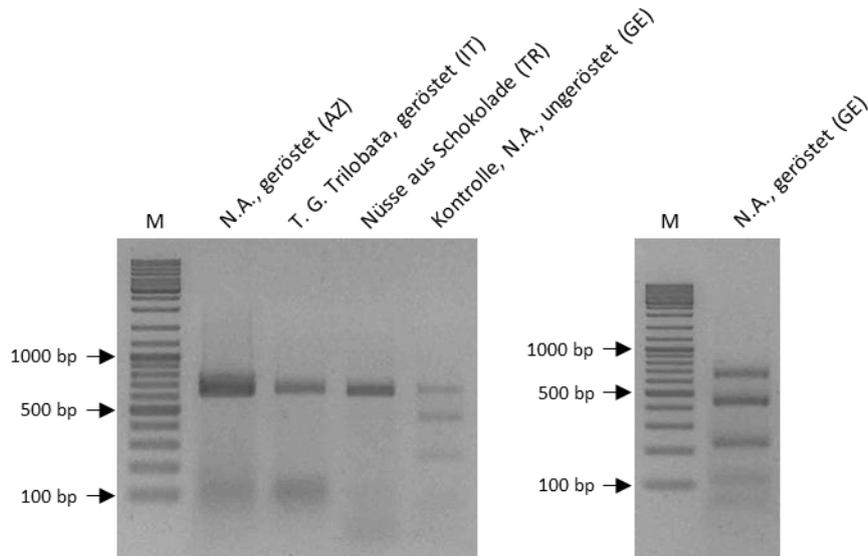


Abbildung 25: Anwendung der entwickelten *targeted*-Methode für den qualitativen Nachweis georgischer Haselnüsse mittels PCR-RFLP. M = DNA-Leiter, AZ = Aserbaidschan, IT = Italien, TR = Türkei, GE = Georgien.

Diese Ergebnisse zeigen, dass die entwickelte PCR-RFLP für den Nachweis georgischer Haselnüsse auch auf geröstete Haselnüsse und damit auch auf prozessierte Produkte anwendbar ist.

6.3.2.3 Quantifizierung des Anteils georgischer Haselnüsse in Mischungen

Methodenoptimierung

Auf Basis der entwickelten PCR-RFLP sollte ebenfalls eine Methode speziell für die Quantifizierung von georgischen Haselnüssen in Mischungen entwickelt werden. Die qualitative Untersuchung von Haselnüssen zeigte jedoch, dass das Amplifikat des Bereichs D mit dem spezifischen SNP nicht vollständig durch das Restriktionsenzym geschnitten wird.

Voraussetzung für eine Quantifizierung ist, dass der Anteil an ungeschnittenem Template konstant ist. Deshalb wurde eine Probe georgischer Haselnüsse (Nr. S4) in Vierfachbestimmung und zusätzlich noch vier weitere Proben (Nr. S6, 15, 162 und 163) analysiert. Die PCR und der anschließende enzymatische Verdau wurden wie unter 7.11.2.2 beschrieben durchgeführt. Die Detektion erfolgte mittels CGE. Im Mittel betrug bei der Vierfachbestimmung der Haselnussprobe Nr. S4 der Anteil an ungeschnittenem Fragment $23,1 \pm 2,4\%$. Der Variationskoeffizient lag bei 10,5%. Der Anteil an ungeschnittenem PCR-Produkt gemittelt über alle acht Proben betrug $24,21 \pm 3,5\%$ mit einem Variationskoeffizient von 14,3%. Anhand dieser Ergebnisse zeigt sich, dass der Anteil des langen Fragments, das nicht durch das Enzym verdaut wird, weitestgehend konstant ist und damit eine Abschätzung bzgl. der Beimischung georgischer Haselnüsse möglich ist. Es muss jedoch von einer rein semi-quantitativen Abschätzung ausgegangen werden.

Für die Erstellung der Kalibriergeraden wurden dieselben Kalibriersets verwendet wie für die Quantifizierung anhand der Indels (siehe 7.11.1.2). Die Durchführung der PCR-RFLP ist unter 7.11.2.2

erläutert. Die Detektion erfolgt ebenfalls mittels CGE sowie AGE mit anschließender Graustufenanalyse der Gelbilder.

Detektion mittels CGE

Die Daten, die für die Erstellung der Kalibriergeraden genutzt wurden, sind im Anhang unter 9.10.5 der Tabelle 42 zu entnehmen. Die Erstellung der Kalibriergerade erfolgte nach demselben Prinzip, wie für die Analyse auf Basis der Indels (siehe 6.3.1.3). Als Ziel-Fragment diente das Fragment 2, da dieses aufgrund seiner Länge von 413 bp im Vergleich zu Fragment 1 (206 bp) ein stärkeres Fluoreszenzsignal aufweist und damit auch im Bereich kleiner Konzentrationen besser detektiert und integriert werden kann. Unter Berücksichtigung der Gesamt-Area aller drei Peaks (Fragment 1 und 2 sowie das ungeschnittene Template) wurde die prozentuale relative Area berechnet und diese anschließend gegen den Anteil georgischer Haselnüsse aufgetragen. Die daraus resultierende Kalibriergerade ist in Abbildung 26 dargestellt.

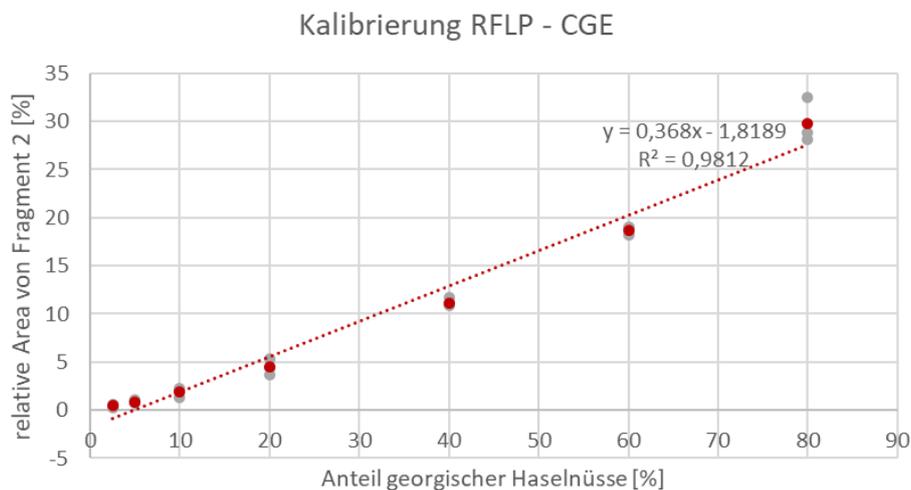


Abbildung 26: Kalibriergerade resultierend aus der Analyse der drei Kalibriersets mittels der entwickelten PCR-RFLP und Detektion mittels CGE. Aufgetragen ist die prozentuale relative Area des Fragments 2 (zugehörig zu Haselnüssen aus Georgien) gegen den Anteil georgischer Haselnüsse. In Grau dargestellt sind die einzelnen Messpunkte der Dreifachbestimmung. In Rot dargestellt sind die zugehörigen Mittelwerte mit der daraus resultierenden Kalibriergerade.

Die Kalibriergerade zeigt mit einem Korrelationskoeffizienten von 0,99, dass ein linearer Zusammenhang zwischen dem Anteil georgischer Haselnüsse und der relativen Area des Fragments 2 besteht. Die Variationskoeffizienten (15,4-29,11%) der Dreifachbestimmungen zeigen jedoch deutliche Schwankungen bei den Mischungen im Bereich von 2,5-20% georgischer Haselnüsse. Im oberen Bereich der Kalibrierung liegen diese hingegen $\leq 6,3\%$. Auch die Reststandardabweichung (s_y), die Verfahrensstandardabweichung (s_{x0}) sowie der Verfahrensvariationskoeffizient (v_{x0}) zeigen mit Werten von $s_y = 1,7$, $s_{x0} = 4,5$ und $v_{x0} = 14,6\%$, dass innerhalb der Kalibrierung Unsicherheiten vorliegen und deshalb lediglich von einer semi-quantitativen Abschätzung ausgegangen werden kann.

Der LOD und LOQ wurden in Anlehnung an das unter 6.3.1.3 beschriebene Vorgehen ermittelt. Das Grundrauschen wurde in dem Bereich bestimmt, in dem bei georgischen Haselnüssen das längere

Fragment (413 bp) nach enzymatischem Verdau detektiert wurde. Der Mittelwert ($\bar{x}_{0\%}$) der relativen Area, die aufgrund der Matrix berechnet wurde, betrug 0,48% und die Standardabweichung ($s_{0\%}$) der Dreifachbestimmung lag bei 0,4. Daraus errechnet sich ein LOD von 1,65% und ein LOQ von 4,37% bezogen jeweils auf die relative Area. Da die Kalibriergerade der PCR-RFLP ebenfalls keine ausreichende Genauigkeit im Bereich niedriger Konzentrationen aufweist, erfolgte die Umwandlung der relativen Area in den zugehörigen prozentualen Anteil georgischer Haselnüsse ebenfalls anhand eines Vergleichs mit den Daten der Kalibrierung. Anhand dessen kann von einem LOD von 10% und einem LOQ von 20% bezogen auf den prozentualen Anteil georgischer Haselnüsse ausgegangen werden. Wird auf Basis dieser Ergebnisse die Kalibrierung auf den Bereich von 20-80% angepasst, verbessern sich s_y leicht auf 1,6, s_{x0} auf 3,9 und v_{x0} auf 7,7%. Diese Werte beruhen allerdings nur auf einer Vier-Punkt-Kalibrierung. Für eine genauere quantitative Aussage, sollte die Anzahl an Kalibrierpunkten im Bereich $> 20\%$ erweitert werden.

Detektion mittels AGE und anschließender Graustufenanalyse

Für die Detektion mittels AGE und anschließender Graustufenanalyse der Gelbilder wurden dieselben Reaktionsansätze verwendet, die auch mittels CGE analysiert wurden. Wie unter 7.11.2.2 beschrieben, wurde ein Drittel der Reaktionslösung gelelektrophoretisch getrennt und zwei Drittel aufgereinigt und für die CGE eingesetzt. Die Kalibriergerade (siehe Abbildung 27) wurde ebenfalls basierend auf den relativen Areas des Fragments 2 erstellt.

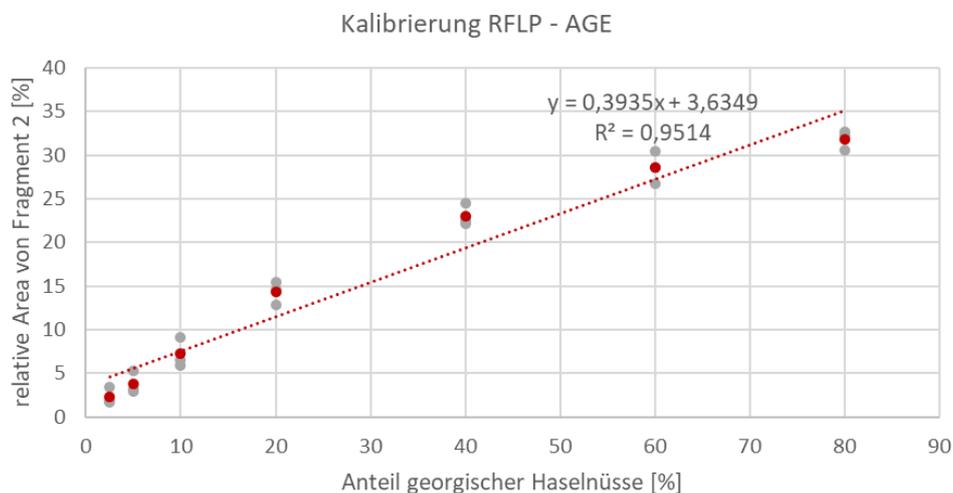


Abbildung 27: Kalibriergerade resultierend aus der Analyse der drei Kalibriersets mittels der entwickelten PCR-RFLP und Detektion mittels AGE + Graustufenanalyse. Aufgetragen ist die prozentuale relative Area des Fragments 2 (zugehörig zu Haselnüssen aus Georgien) gegen den Anteil georgischer Haselnüsse. In Grau dargestellt sind die einzelnen Messpunkte der Dreifachbestimmung und in Rot die zugehörigen Mittelwerte mit der daraus resultierenden Kalibriergerade.

Bereits visuell ist zu erkennen, dass sowohl innerhalb der Dreifachbestimmungen Schwankungen vorliegen und die Kalibrierung insgesamt Unsicherheiten aufweist. Dies zeigt auch das Bestimmtheitsmaß von 0,9514. Dennoch ist auch deutlich, dass ein gewisser linearer Zusammenhang vorliegt ($r = 0,98$). Vergleichbar zu den Ergebnissen nach der Messung mittels CGE sind die

Variationskoeffizienten im unteren Bereich der Kalibrierung (2,5-10%) deutlich höher (19,35-33,52 %) als im oberen Bereich (< 7,68%). Die Abweichungen der einzelnen Kalibrierpunkte von der Regressionsgeraden und damit die Unsicherheiten der Kalibrierung verdeutlichen auch die Reststandardabweichung (s_y) von 2,9, die Verfahrensstandardabweichung (s_{x0}) von 7,4 und der Verfahrensvariationskoeffizient (v_{x0}) von 23,9%. Daher handelt es sich hier ebenfalls um eine rein semi-quantitative Abschätzung bzgl. des Anteils georgischer Haselnüsse. Zudem zeigen die Werte deutlich, dass die Schwankungen bei der Detektion mittels AGE und anschließender Graustufenanalyse größer ausfallen als im Vergleich zu der Detektion mittels CGE.

Auch für diese Möglichkeit der Abschätzung von Beimischungen wurden der LOD und LOQ wie zuvor auch für die CGE bestimmt. Der Mittelwert ($\bar{x}_{0\%}$) der relativen Area verursacht durch die Matrix betrug 1,7% und die zugehörige Standardabweichung ($s_{0\%}$) 0,25. Mit diesen Werten wurde mit den unter 7.11.1.2 angegebenen Formeln ein LOD von 2,46% und ein LOQ von 4,22% bezogen auf die relative Area berechnet. Diese Werte entsprechen etwa einem LOD von 4% und einem LOQ von 7% bezogen auf den Anteil georgischer Haselnüsse. Unter Berücksichtigung der visuellen Bewertung der Gelbilder ist von einem LOD von 10% und einem LOQ von 20% auszugehen. Erfolgt auch hier eine Anpassung der Kalibriergeraden auf den Bereich von 20-80%, bestätigt sich das Bestimmtheitsmaß von 0,98 und es ist eine Verbesserung von s_y auf 2,0, s_{x0} auf 6,8 und v_{x0} auf 13,5% zu erkennen.

6.3.3 HRMA für die Differenzierung von Haselnüssen unterschiedlichen geographischen Ursprungs

6.3.3.1 Ausgangssituation für die Methodenentwicklung

Ziel war die Entwicklung einer HRMA auf Basis der im Bereich D des cpGenoms identifizierten SNPs. Wie in Abbildung 28 zu erkennen ist, unterscheiden sich die Gruppen der georgischen, aserbaidisch/türkischen und europäischen Proben in der Kombination der SNPs. Es ist dabei dieselbe Gruppierung zu erkennen, wie sie bereits unter 6.2 geschildert wurde. Bei der in Abbildung 28 aufgeführten Probe der Sorte Palaz aus der Türkei handelt es sich um eine der Proben, die ebenfalls das Insert aufweist und daher derselben Gruppe zugeordnet werden kann, der auch die Haselnüsse aus Aserbaidisch angehören. Im Folgenden wird daher auch die Unterteilung in Gruppe 1-3 (siehe Abbildung 52) für die Probenzuordnung und -benennung herangezogen.

Zu Beginn der Analyse des Schmelzpunktes T_m wurde daher basierend auf den Sequenzen eine Clusterung der Haselnüsse in drei Gruppen erwartet.

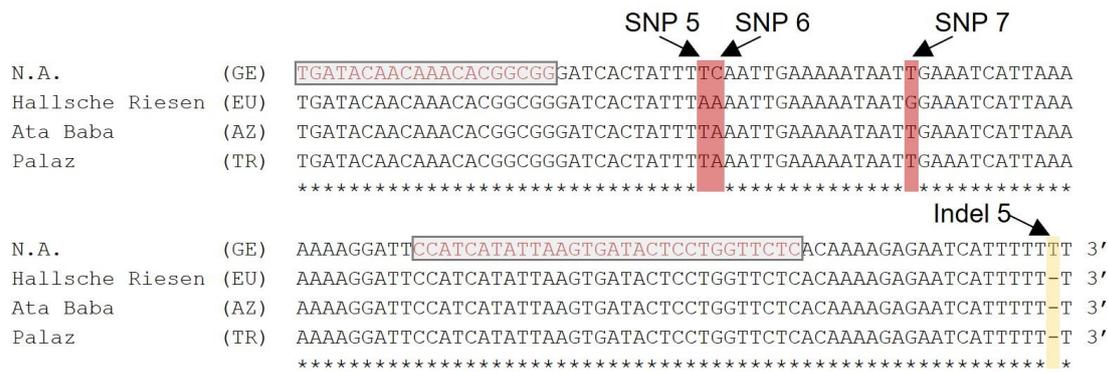


Abbildung 28: Ausschnitt eines Alignments des Bereichs D des cpGenoms vier ausgewählter Sorten. Grau hinterlegt sind die Primerbindungsstellen des Primerpaares D.HRMA_FW/RV. Rot hinterlegt sind die identifizierten SNPs.

Haselnüsse aus Georgien (Gr. 3) unterscheiden sich von Nüssen aus Gruppe 2 (Aserbaidschan und Türkei mit Insert) lediglich anhand eines C/A-Austauschs (SNP 6). Die europäischen Haselnüsse sowie Sorten aus der Türkei ohne das Insert (Gruppe 1) weisen im Vergleich zu den georgischen Nüssen an allen drei Positionen andere Basen und im Vergleich zu Gruppe 2 zwei SNPs (A/T-Austausch bei SNP 5 und ein G/T-Austausch bei SNP 7) auf.

6.3.3.2 Berechnung von Schmelzkurven

Mit Hilfe des Web-Tools uMelt^{Quartz} wurden die zu erwartenden Schmelzkurven berechnet.^[224] Hierfür müssen Parameter bzgl. der Reaktionsbedingungen wie z.B. die Konzentration freier Mg^{2+} -Ionen sowie monovalenter Kationen angegeben werden. Zudem müssen die für die Berechnung angewendeten thermodynamischen Grundlagen ausgewählt werden. Um den Einfluss dieser Parameter auf das Ergebnis der Berechnung abschätzen zu können, wurde folgende Sequenz mit einer Länge von 70 bp und einem GC-Gehalt von 40% beispielhaft herangezogen und deren Schmelzkurve basierend auf variierenden Parametern berechnet:

5' – ACGACGTTGTAAAACGACAGAAGCATAGTATAGAAGAAACGTACGATTCGCGCCAGATCGACTTATTAA –3'

Veränderungen der Reaktionsbedingungen (Konzentration freier monovalenter Kationen, freier Mg^{2+} -Ionen und Zugabe von DMSO) zeigten die zu erwartenden Verschiebungen der Schmelzkurve (siehe Abbildung 61 unter 9.10.6.1). So stabilisieren Kationen die DNA und erhöhen damit den Schmelzpunkt. Dieser Effekt zeigt sich bei Mg^{2+} -Ionen deutlich stärker im Vergleich zu monovalenten Kationen wie Na^+ .^{[225],[226]} DMSO hingegen führt zu einer Verringerung des Schmelzpunkts, was sich ebenfalls auch in den berechneten Schmelzkurven zeigte.^[227] Der Vergleich der Schmelzkurven, die basierend auf den unterschiedlichen thermodynamischen Berechnungsgrundlagen erhalten wurden, zeigte, dass diese einen enormen Einfluss auf das Ergebnis haben (siehe Abbildung 29). Die Schmelzpunkte schwanken in einem Bereich von 78-86,5 °C.

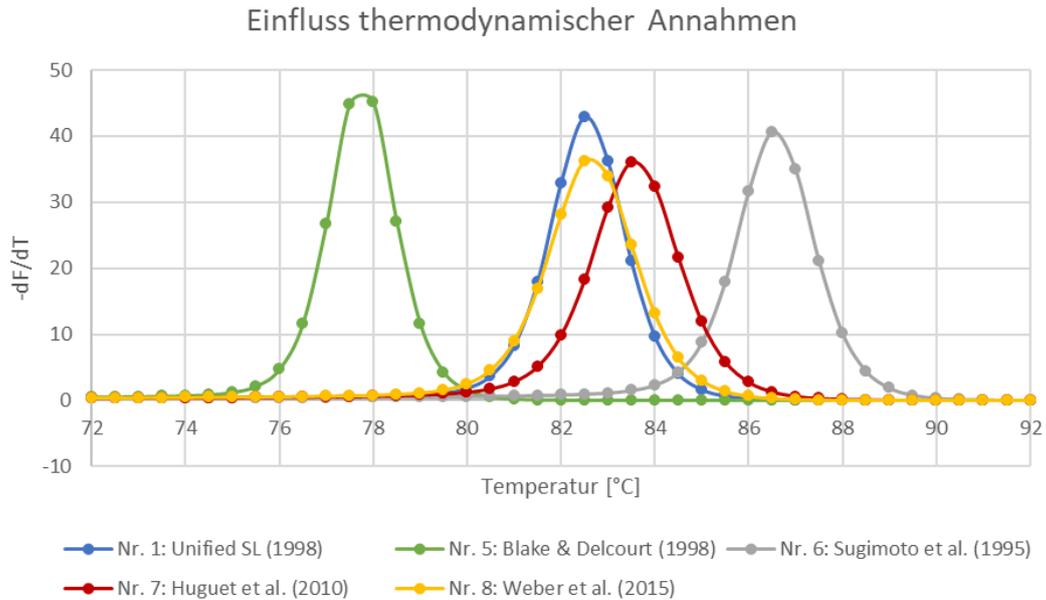


Abbildung 29: Berechnete Schmelzkurven für die angegebene 70 bp lange DNA-Sequenz. Die Berechnung basiert auf unterschiedlichen thermodynamischen Berechnungsgrundlagen.

Die Ergebnisse dieser Vorversuche zeigten, dass mit der Berechnung von Schmelzkurven die final zu erwartende Schmelztemperatur nur sehr grob abgeschätzt werden kann. Allerdings eignet sich das Web-Tool gut, um den Einfluss von Sequenzunterschieden auf den Schmelzpunkt einzuschätzen und ggf. die Methodenparameter bereits vor Beginn experimenteller Versuche optimieren zu können. So kann zum Beispiel die Länge des zu analysierenden DNA-Abschnitts sowie die Position eines SNPs hinsichtlich des Einflusses auf die Schmelztemperatur optimiert werden, um größtmögliche Unterschiede in T_m zu erhalten.

Für die im Anschluss angewendete Berechnung im Zuge der Methodenentwicklung der HRMA wurden die unter 7.11.3.1 in Tabelle 20 angegebenen Parameter herangezogen. Abbildung 30 zeigt die berechneten Schmelzkurven der drei Gruppen.

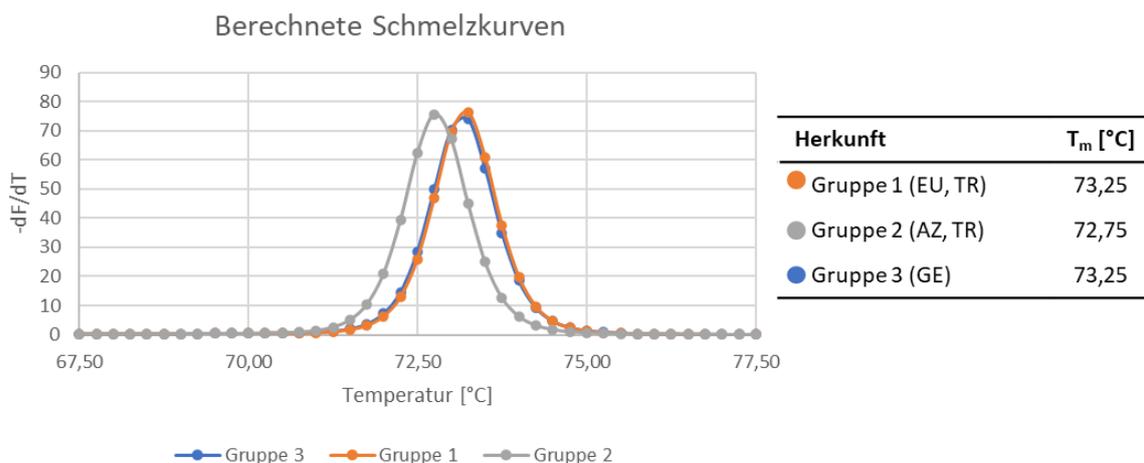


Abbildung 30: Links: Mittels uMelt^{Quartz} berechnete Schmelzkurven auf Basis der in Abbildung 28 dargestellten Sequenzen für Haselnüsse aus Georgien (GE), Europa (EU), Aserbaidschan (AZ) und Türkei (TR), Rechts: Zusammenfassung der berechneten Schmelzpunkte T_m .

Es ist deutlich zu erkennen, dass die Sequenz der Haselnüsse aus Gruppe 1 und 3 dasselbe berechnete Schmelzverhalten zeigen. Es ist möglich, dass dies auf der SNP-Kombination beruht. Die georgischen Nüsse weisen, rein die SNPs betrachtet, die Basen TCT und die europäischen und türkischen Sorten ohne Insert die Basen AAG auf. Da es sich hierbei um die jeweils komplementären Basen handelt, könnte das der Grund für das ähnliche Schmelzverhalten sein. Beide Sequenzen weisen einen berechneten Schmelzpunkt von 73,25 °C auf. Haselnüsse der Gruppe 2 aus Aserbaidschan und Türkei (mit Insert) weisen hingegen ein abweichendes Schmelzverhalten auf. Sie können laut der Berechnung durch eine Temperaturdifferenz von 0,5 °C von den Haselnüssen anderer Herkunft abgetrennt werden. Auf Grundlage der Berechnung ist mit einer möglichen Aufteilung in zwei Gruppen zu rechnen. Eine Gruppe bilden dabei Haselnüsse aus Aserbaidschan und Türkei ($T_m = 72,75$ °C) und die zweite Gruppe Haselnüsse aus Europa, Türkei und Georgien ($T_m = 73,25$ °C).

6.3.3.3 Methodenentwicklung und -validierung

Bei der Entwicklung einer HRMA spielt die Länge des gewählten PCR-Fragments eine entscheidende Rolle. Sie sollte relativ kurz gewählt werden, damit das Verhältnis zwischen den Sequenzunterschieden zur Gesamtlänge möglichst groß ist. Sind die PCR-Amplifikate zu lang, kann ggf. die Temperaturdifferenz nicht mehr detektiert werden.^{[200],[203]} In der Literatur werden meist PCR-Amplifikate zwischen 50-500 bp verwendet.^[197] Gemäß der Herstellerangaben des verwendeten *AccuMelt™ HRM SuperMix* sollte die Fragmentlänge 250 bp nicht überschreiten.^[228] Die Wahl der Primer für die Amplifikation sollte möglichst so gewählt werden, dass der zu detektierende Sequenzunterschied sich in der Mitte des Fragments befindet. Neben der Fragmentlänge beeinflusst auch der GC-Gehalt den Schmelzpunkt, da sich zwischen den Basen Guanin und Cytosin drei Wasserstoffbrückenbindungen ausbilden, die zu einer stärkeren Bindungen führen.^[200] Damit steigt mit zunehmendem GC-Gehalt die Schmelztemperatur. Ein weiterer Faktor sind, wie bereits bei der Berechnung von Schmelzkurven verdeutlicht, die thermodynamischen Umgebungsbedingungen wie beispielsweise die Salzkonzentration der Reaktionslösung.^[199]

Die finale Analyse der Schmelzpunkte wurden wie unter 7.11.3 beschrieben durchgeführt. Es wurden insgesamt 49 Proben, bestehend aus zehn georgischen, sechs aserbaidshianischen, 14 türkischen und 19 europäischen Haselnussproben analysiert. Das Amplifikat, dessen Schmelztemperatur T_m gemessen wurde, wurde zuvor mit dem Primerpaar D.HRMA_FW/RV hergestellt. Die Ergebnisse sind in Tabelle 9 zusammengefasst.

Bezüglich der analysierten Proben aus der Türkei ist noch zu erwähnen, dass es sich nur um Proben handelt, die auch hinsichtlich der Indels untersucht wurden und dabei eindeutige Ergebnisse lieferten. Die Zuordnung der Proben zu den Gruppen 1 und 2 erfolgte lediglich auf Basis des Kriteriums „Insert vorhanden / nicht vorhanden“. So wurden auch Proben der Sorten Tombul (Nr. 127), Palaz (Nr. 128, 139, 140, 168 und 170) und Karayağlı (Nr. 134) analysiert, die in Bezug auf die

Indels eindeutige Ergebnisse lieferten und daher basierend auf ihrer Sequenz eindeutig den Gruppen 1 und 2 zuzuordnen sind.

Tabelle 9: Ergebnisse der HRMA zur Identifizierung der Schmelztemperatur des kurzen Amplifikats (generiert mit dem Primerpaar D.HRMA_FW/RV) von Bereich D (Länge: 99 bp) von Haselnüssen aus Georgien (GE), Aserbaidshan (AZ), Türkei (TR) und Europa (EU). *Proben, die für die Erstellung der Abbildung 31 herangezogen wurden.

Gruppe 1			Gruppe 2			Gruppe 3		
Probe	Herkunft	T _m [°C]	Probe	Herkunft	T _m [°C]	Probe	Herkunft	T _m [°C]
F10*	EU	72,8	S2*	AZ	72,2	S5*	GE	72,8
7*	EU	72,8	25	AZ	72,2	S6*	GE	72,8
8*	EU	72,6	102*	AZ	72,4	S4*	GE	72,8
9*	EU	72,8	121*	AZ	72,4	13*	GE	73,0
22	EU	72,6	122*	AZ	72,2	15*	GE	72,8
31*	EU	72,6	123*	AZ	72,2	19*	GE	73,0
32*	EU	72,8	128*	TR	72,4	118*	GE	72,6
37*	EU	72,6	131	TR	72,2	124*	GE	72,8
40*	EU	72,8	134	TR	72,2	151*	GE	72,8
51*	EU	72,8	139*	TR	72,4	152*	GE	72,8
52*	EU	72,6	140*	TR	72,4			
64	EU	72,6	142*	TR	72,4			
80	EU	72,6	143*	TR	72,4			
84	EU	72,6	155	TR	72,2			
85	EU	72,6	168	TR	72,2			
106*	EU	72,8	170	TR	72,2			
107	EU	72,4	177	TR	72,2			
108*	EU	72,6	178	TR	72,0			
109	EU	72,6						
127	TR	72,8						
132	TR	72,8						
Mittelwert:		72,7 ± 0,1	Mittelwert:		72,3 ± 0,1	Mittelwert:		72,8 ± 0,1
Variationsk.		0,16%	Variationsk.		0,13%	Variationsk.		0,15%

Die erhaltenen Werte für T_m bestätigen das nach der Berechnung zu erwartende Ergebnis. Die Haselnussproben aus den Gruppen 1 und 3 weisen im Mittel nur eine Temperaturdifferenz von 0,1 °C auf und die Einzelwerte überschneiden sich deutlich. Daher ist bei diesen beiden Sequenzen von einem sehr ähnlichen Schmelzverhalten auszugehen, weshalb hier auf Basis von T_m keine Trennung möglich ist. Die Proben aus Gruppe 2 weisen im Mittel einen Schmelzpunkt von 72,3 °C auf. Lediglich eine Probe (Nr. 178) weicht mit nur 72,0 °C leicht ab. Es treten zum Teil gewisse Schwankungen innerhalb der Messreihen einer Herkunft auf, dennoch liegen alle Variationskoeffizienten < 0,2%. Es gibt lediglich eine europäische Haselnussprobe (Nr. 107), die mit einer Schmelztemperatur von 72,4 °C im Bereich der Gruppe 2 liegt. Im Mittel unterscheiden sich die georgischen Haselnussproben um 0,5 °C von den Nüssen aus Aserbaidshan und auch einigen Sorten aus der Türkei (TR[+] = Insert vorhanden). Diese analytisch ermittelte Temperaturdifferenz stimmt mit der berechneten überein.

Damit zeigte sich erneut, dass sich die Berechnung von Schmelzkurven aufgrund der Variablen wie die hinterlegten Grundlagen der Thermodynamik sowie den Ionenkonzentrationen in Bezug auf T_m nicht immer bestätigen lässt, jedoch zu erwartende Temperaturdifferenzen gut abgeschätzt werden können. uMelt^{Quartz} stellt damit ein sehr hilfreiches Tool im Rahmen der Methodenentwicklung einer HRMA dar.

Abbildung 31 zeigt beispielhaft die basierend auf den normalisierten Schmelzkurven (siehe 9.10.6, Abbildung 62) berechneten Differenzkurven einiger Proben. Als Referenz dienten dabei die europäischen Proben. Es ist deutlich zu erkennen, dass alle drei Gruppen unterschiedliche Kurvenverläufe aufweisen. Die Differenz zwischen den Proben der Gruppe 2 (blau) und die der Gruppe 3 (grün) fällt dabei sehr deutlich aus. Werden nur die Gruppen 2 und 3 betrachtet und die Differenzkurven auf Basis der Gruppe 2 berechnet, so wird ebenfalls der Unterschied beider Gruppen nochmals deutlich. Die zugehörigen Abbildungen (Abbildung 63 und Abbildung 64) sind unter 9.10.6 aufgeführt. Damit stellt auch die Bewertung des *Difference plots* einen wichtigen Bestandteil der Probenzuordnung da, da er wichtige Hinweise auf die Herkunft geben kann. Voraussetzung hierfür ist jedoch, dass immer Referenzproben bekannter Herkunft als Referenz sowohl für den Vergleich der Schmelztemperatur, als auch die Berechnung der Differenzkurven mit analysiert werden.

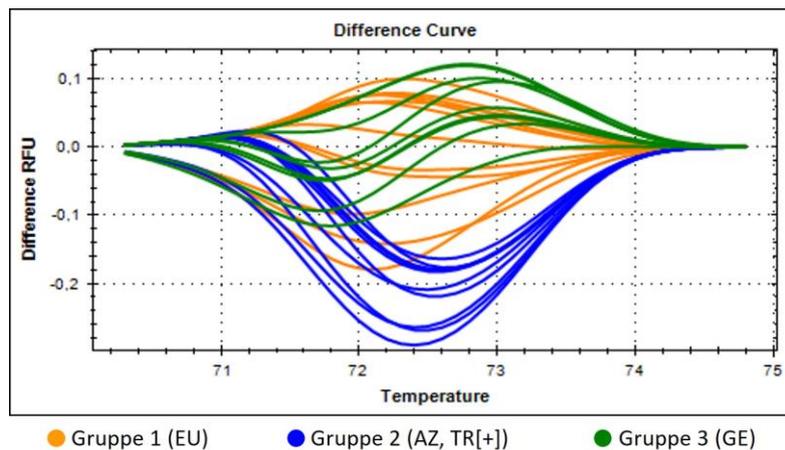


Abbildung 31: Beispielhafte Abbildung eines *Difference plots* 32 ausgewählter Proben (siehe Tabelle 9) bestehend aus zwölf europäischen, fünf aserbaidchanischen, fünf türkischen (TR[+]) und zehn georgischen Proben. Als Referenz für die Berechnung dienten die europäischen Proben.

Für die Validierung der erhaltenen Ergebnisse und zur Überprüfung der Wiederholpräzision wurden von den zuvor bereits untersuchten Proben insgesamt 26 Proben ausgewählt und an drei unterschiedlichen Tagen analysiert. Die Ergebnisse sind in Tabelle 10 zusammengefasst.

Tabelle 10: Ergebnis der Validierung der qualitativen HRMA für die Unterscheidung von Haselnüssen unterschiedlicher Herkunft auf Basis des Bereichs D des cpGenoms.

Gruppe	Probennr.	Herkunft	Schmelztemperatur T _m [°C]			Mittelwert [°C]
			Bestimmung 1	Bestimmung 2	Bestimmung 3	
Gruppe 1	F10	EU	72,8	72,8	72,6	72,7 ± 0,10
	7	EU	72,8	72,8	72,6	
	8	EU	72,8	72,8	72,6	
	9	EU	72,8	72,8	72,6	
	31	EU	72,8	72,8	72,8	
	32	EU	72,8	72,8	72,8	
	107	EU	72,8	72,8	72,6	
	108	EU	72,6	72,6	72,6	
	109	EU	72,6	72,8	72,6	
Gruppe 2	S2	AZ	72,4	72,4	72,4	72,4 ± 0,07
	102	AZ	72,4	72,4	72,4	
	121	AZ	72,4	72,4	72,4	
	122	AZ	72,4	72,4	72,4	
	123	AZ	72,6	72,6	72,4	
	128	TR	72,4	72,4	72,4	
	131	TR	72,4	72,4	72,4	
	134	TR	72,4	72,6	72,4	
	142	TR	72,4	72,6	72,4	
Gruppe 3	S5	GE	73,0	73,0	73,0	73,0 ± 0,04
	S6	GE	73,0	72,8	73,0	
	S4	GE	73,0	73,0	73,0	
	13	GE	73,0	73,0	73,0	
	15	GE	73,0	73,0	73,0	
	124	GE	73,0	73,0	73,0	
	151	GE	73,0	73,0	73,0	
	152	GE	73,0	73,0	73,0	

Innerhalb der Gruppen sind nur sehr geringe Schwankungen der Messwerte festzustellen. Die georgischen Proben (Gruppe 3) zeigen im Vergleich zu den vorherigen Messungen insgesamt einen leichten Temperaturshift um 0,2 °C nach oben. Während bei der Methodenentwicklung der Mittelwert bei 72,8 °C lag, beträgt er nun 73,0 °C. Die Proben der Gruppe 1, im Probenpool für die Validierung rein europäischer Herkunft, sind bzgl. ihres Schmelzpunktes konstant geblieben (72,7 °C). Bei Gruppe 2 ist ebenfalls ein leichter Temperaturshift nach oben zu erkennen und damit ein um 0,1 °C erhöhter Mittelwert von 72,4 °C. Solche Temperaturshifts wurden im Rahmen der Methodenentwicklung mit zunehmendem Alter des *AccuMelt™ HRM SuperMix* beobachtet. Da anderweitig keinerlei Parameter verändert wurden und die Werte stets reproduzierbar waren, wird davon ausgegangen, dass dieses Phänomen mit dem Alter des *AccuMelt™ HRM SuperMix* zusammenhängt. Unter anderem aus diesem Grund sollte für eine Bestimmung der Herkunft jeweils

eine authentische Referenzprobe jeder Herkunft parallel analysiert werden, um einen exakten Vergleich der Werte vornehmen zu können.

Vergleich Gruppe 1 und 3:

Es ist eine Temperaturdifferenz von 0,3 °C zu erkennen, die zuvor nur bei 0,1 °C lag. Die Standardabweichung bei Gruppe 1 beträgt jedoch 0,1 °C. ΔT fällt damit sehr gering aus. Die Differenzkurven (siehe Abbildung 32) können jedoch zudem einen Hinweis geben, da hier leichte Unterschiede in den Kurvenverläufen zu erkennen sind. Dennoch sollte hier das Ergebnis ggf. durch eine weitere Methode wie die PCR-RFLP (siehe 6.3.2) für den Nachweis georgischer Proben bestätigt werden.

Vergleich Gruppe 1 und 2:

Bei der Validierung ergab sich eine Temperaturdifferenz von 0,3 °C zwischen den Gruppen 1 und 2. Die jeweiligen Mittelwerte der Dreifachbestimmungen weisen keine Überschneidungen auf. Die geringste Temperaturdifferenz (Probe Nr. 123 (Gr. 2) mit $T_m = 72,5$ °C und Nr. 108 (Gr. 1) mit $T_m = 72,6$ °C) liegt jedoch bei nur 0,1 °C. Auch hier können die Differenzkurven einen Hinweis auf die Herkunft geben, allerdings sollte auch hier im Zweifelsfall das Ergebnis durch weitere Analysen bestätigt werden.

Vergleich Gruppe 2 und 3:

Die Unterscheidung dieser beiden Gruppen war das primäre Ziel dieser HRMA. Mit $\Delta T = 0,6$ °C ist eine verlässliche Trennung von Haselnüssen aus Georgien von denen aus Aserbaidshan oder Türkei (mit Insert) möglich. Einen wichtigen Anhaltspunkt bietet dabei auch der *Difference plot*, welcher bei diesen beiden Gruppen visuell Unterschiede deutlich erkennbar macht (siehe Abbildung 32 und Abbildung 67).

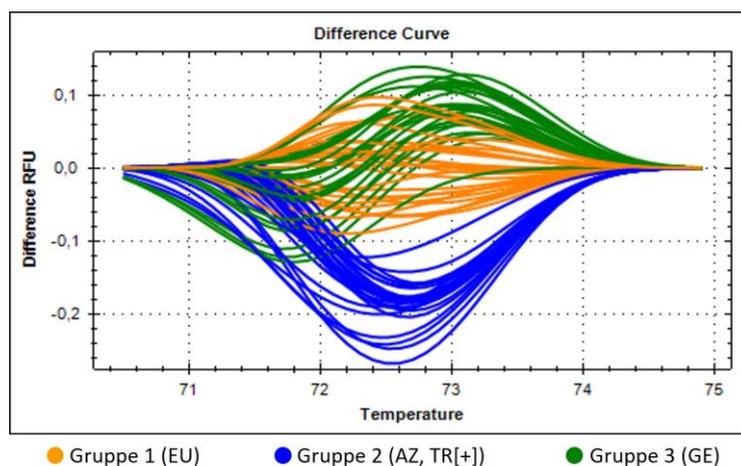


Abbildung 32: *Difference plot* aller Proben (siehe Tabelle 10) der Prüfung auf Reproduzierbarkeit der entwickelten HRMA. Die zugehörigen normalisierten Schmelzkurven sind unter 9.10.6 in Abbildung 65 dargestellt.

Da die Verläufe der Schmelzkurven, die über die *Difference plots* visualisiert werden können, bei der Identifizierung der Herkunft sehr hilfreich sind, wird empfohlen immer Referenzproben authentischer Herkunft aus allen drei Regionen bei der Analyse mitlaufen zu lassen. Sie können nicht nur erforderlich für den Abgleich von T_m , sondern auch relevant für die Berechnung der

Differenzkurven sein. Für die Unterscheidung der Gruppen 2 und 3 sollten Proben europäischer Haselnüsse als Referenzproben für den *Difference plot* verwendet werden, da damit der größte visuell erfassbare Unterschied erreicht wird.

Die Gruppen 2 und 3, die sich anhand der Inserts nicht unterscheiden lassen, können deutlich voneinander getrennt werden. Die hier vorgestellte HRMA stellt damit eine Alternative zu der PCR-RFLP dar, mit der ebenfalls georgische Haselnüsse von Nüssen anderer Herkunft unterschieden werden können. Nachteil der HRMA im Vergleich zu der PCR-RFLP ist die mögliche Überlagerung mit Sorten der Gruppe 1. Vorteil ist jedoch, dass es sich um einen *closed tube*-Ansatz handelt. Das heißt, es muss lediglich die PCR pipettiert werden, um direkt nach der Amplifikation inklusive der automatisch angehängten Schmelzpunktanalyse ein aussagekräftiges Ergebnis zu erhalten ohne, dass weitere Schritte wie der enzymatische Verdau und die anschließende Detektion mittels AGE notwendig sind. Damit kann die benötigte Analysenzeit verringert werden.

6.3.3.4 Überprüfung der Anwendbarkeit auf prozessierte Proben

Die zuvor validierte Methode wurde schließlich auf ihre Anwendbarkeit auf prozessierte Produkte hin untersucht. Dafür wurden selbst geröstete georgische und aserbaidische sowie kommerziell erworbene geröstete Haselnüsse der Sorte Tonda Gentile Trilobata aus Italien und zusätzlich Nüsse aus einer handelsüblichen Schokolade, die laut Herstellerangaben aus der Türkei stammen, analysiert. Als Vergleichsprobe und zur Überprüfung einer einwandfreien Analyse wurde jeweils eine Probe aus Georgien (Nr. S4), Aserbaidschan (Nr. 102) und Europa (Nr. 7) als Vergleich parallel untersucht. Die Ergebnisse sind in Tabelle 11 zusammengefasst.

Tabelle 11: Ergebnisse der HRMA prozessierter Proben sowie drei Vergleichsproben.

	Nr.	Probe	Herkunft	Schmelztemperatur [°C]
Prozessierte Proben	V4	Geröstete Haselnüsse	Georgien	73,0
	V5	Geröstete Haselnüsse	Aserbaidschan	72,4
	V3	Geröstete T. G. Trilobata	Europa	72,6
	V6	Geröstete Haselnüsse aus Schokolade	Türkei	73,0
Vergleichs-proben	S4	Ungeröstete Haselnüsse	Georgien	73,0
	102	Ungeröstete Haselnüsse	Aserbaidschan	72,4
	7	Ungeröstete T. G. Romana	Europa	72,8

Bei allen Proben konnte das gewünschte Amplifikat hergestellt werden. Ein Vergleich der ungerösteten und gerösteten Proben, die jeweils aus einem Land stammen, zeigte keine Unterschiede. Alle Proben wiesen die gleichen Schmelztemperaturen auf, wie sie auch zuvor während der Methodenentwicklung und -validierung erhalten wurden. Die gerösteten Haselnüsse aus der Schokolade, die laut Herstellerangaben aus der Türkei, stammen, clustern mit einem

Schmelzpunkt von 73,0 °C zu den Gruppen 1 und 3. Da mittels PCR-RFLP (siehe Abschnitt 6.3.2.2) keine Fragmentierung detektiert wurde, kann ausgeschlossen werden, dass es sich um georgische Haselnüsse handelt. Sind die Angaben des Herstellers korrekt, muss es sich um eine Sorte handeln, die die Inserts nicht aufweisen wie z.B. die Sorte Çakildak. Der zugehörige *Difference plot*, ist in Abbildung 33 dargestellt. Dieser zeigt ebenfalls sehr anschaulich die Clusterung der Proben gemäß ihrer Herkunft.

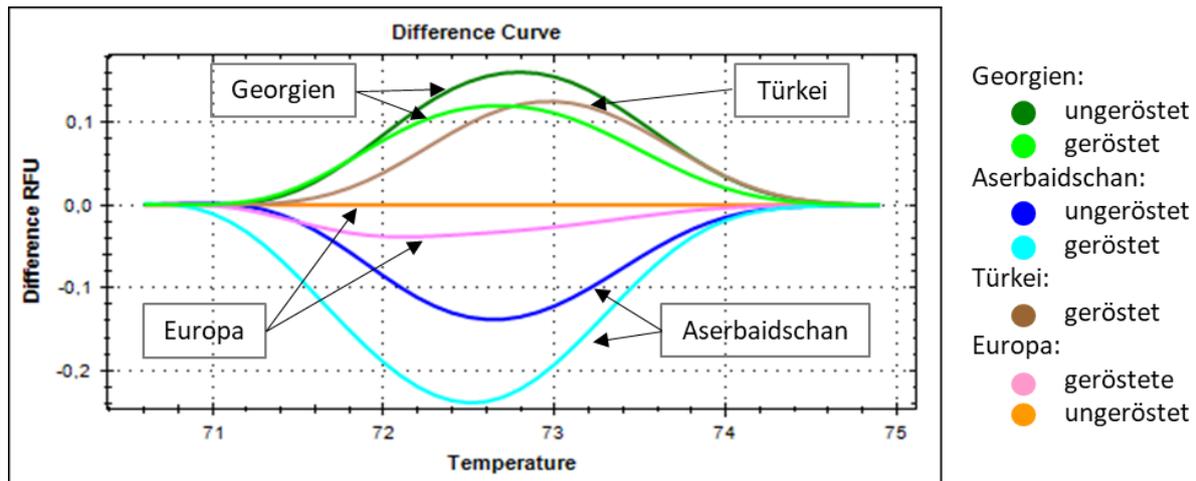


Abbildung 33: *Difference plot* der HRMA der prozessierten Haselnussproben. Für die Berechnung wurde die ungeröstete Probe europäischer Haselnüsse als Referenzprobe herangezogen.

Die entwickelte Methode konnte folglich erfolgreich auf geröstete Haselnüsse übertragen werden und ist damit für die Analyse dieser anwendbar.

6.3.3.5 Quantifizierung mittels HRMA

Im Folgenden sollte geprüft werden, ob auf Basis der entwickelten HRMA auch eine Quantifizierung von Mischungen möglich ist. Der Fokus lag dabei auf der Mischungen von Haselnüssen aus Georgien und Aserbaidschan bzw. Türkei (TR[+]). Um zunächst festzustellen, ob dieser Ansatz denkbar ist, wurden Mischungen von DNA-Isolaten der Proben Nr. S4 (Georgien) und Nr. 123 (Aserbaidschan) hergestellt und wie unter 7.11.3 erläutert analysiert. Abbildung 34 zeigt die erhaltenen normalisierten Schmelzkurven sowie die zugehörigen Schmelzpunkte T_m .

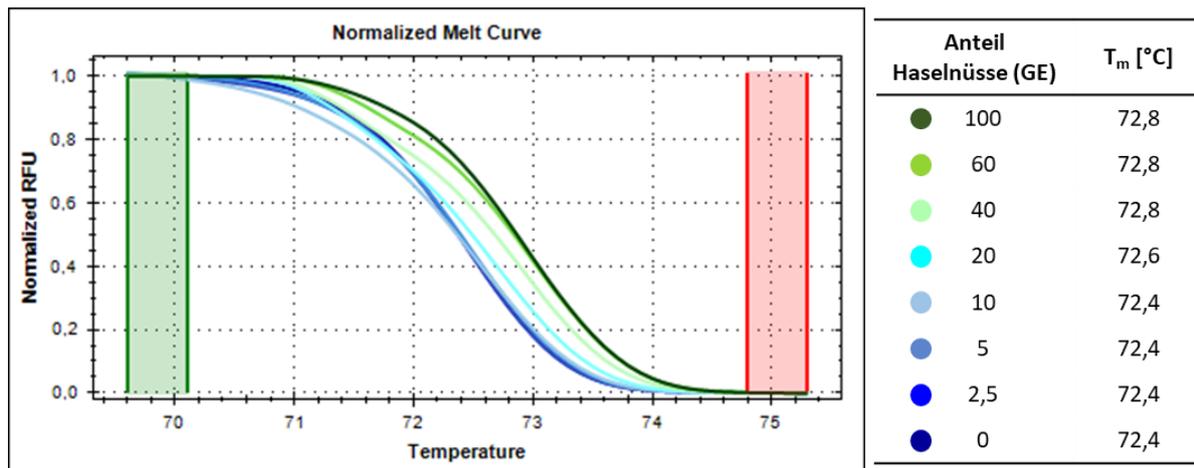


Abbildung 34: Links: Normalisierte Schmelzkurven von Mischungen aus DNA-Isolaten georgischer und aserbaidchanischer Haselnüsse. Rechts: Die zugehörigen detektierten Schmelzpunkte T_m .

Die Verschiebung der Schmelzkurve in den Bereich höherer Schmelztemperaturen mit zunehmendem Anteil georgischer Haselnüsse ist deutlich zu erkennen. Da die Erhöhung der Temperatur während der Analyse in Schritten von 0,2 °C erfolgte, verschiebt sich die Temperatur auch in diesen Intervallen. Die erste Erhöhung um 0,2 °C wird ab einem Anteil von 20% georgischen Haselnüssen detektiert, die Zweite ab 40%. Abbildung 35 zeigt die entsprechenden Differenzkurven basierend auf den normalisierten Schmelzkurven (Abbildung 34). Als Referenz für die Auswertung diente ein Ansatz bestehend aus 100% aserbaidchanischen Haselnüssen (0% GE).

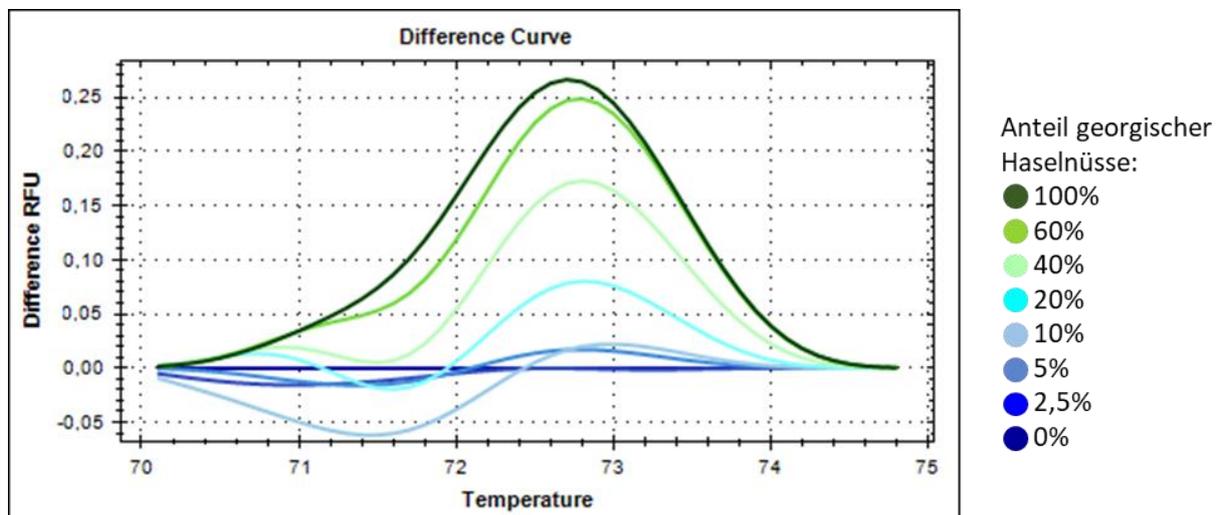


Abbildung 35: *Differenz plot* der Schmelzpunktanalyse von DNA-Mischungen georgischer und aserbaidchanischer Haselnüsse.

Wie in Abbildung 35 deutlich zu erkennen ist, nimmt mit zunehmendem Anteil georgischer Haselnüsse auch die Höhe des Peaks im Bereich zwischen 72 und 73 °C zu. Bereits 5 und 10% Beimischung zeigen einen positiven Wert für ΔF und ab 20% ist ein sehr deutlicher Peak zu erkennen. Da Haselnüsse der Gruppe 2 (AZ, TR[+]) stets einen negativen Wert für ΔF aufwiesen, kann von einer möglichen Beimischung ausgegangen werden, sobald im Bereich zwischen 72 und 73 °C ein positiver Wert erhalten wird. Voraussetzung für diese Bewertung ist jedoch, dass eine Probe, bestehend zu 100% aus aserbaidchanischen Haselnüssen als Referenzkurve dient.

Wird ΔF (entspricht der absoluten Höhe der Differenzkurve) gegen den Anteil an Haselnüssen aus Georgien aufgetragen, so folgt daraus die in Abbildung 36 dargestellte Kalibriergerade. Bestimmt wurde die Höhe im Bereich der zuvor bestimmten Schmelzpunkte, d.h. die maximale Höhe zwischen 72 und 73 °C. Die Werte, die für die Erstellung der Kalibriergeraden herangezogen wurden, sind der Tabelle 45 (siehe 9.10.6.3) zu entnehmen.

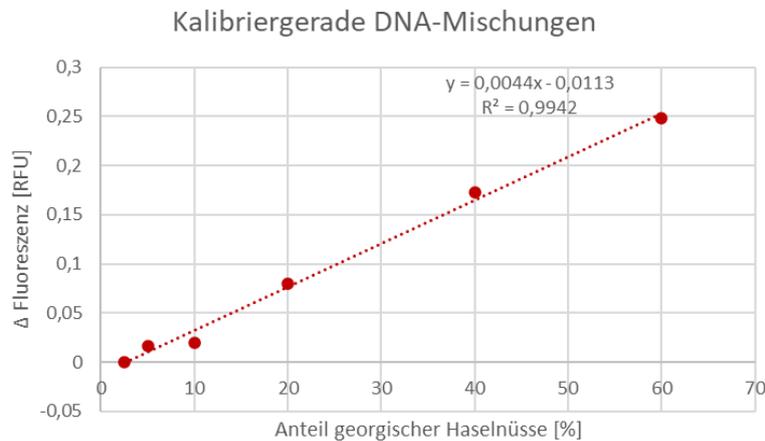


Abbildung 36: Kalibriergerade der HRMA von Mischungen von DNA-Isolaten georgischer und aserbaidischer Haselnüsse. Aufgetragen wurde Δ Fluoreszenz gegen den Anteil georgischer Haselnüsse.

Bereits rein visuell ist ein linearer Zusammenhang zwischen den Größen ΔF und dem Anteil georgischer Haselnüsse zu erkennen. Dies wird auch durch den Korrelationskoeffizienten von 0,9971 bestätigt. Die Reststandardabweichung s_y beträgt 0,009, die Verfahrensstandardabweichung s_{x0} 1,95 und der Verfahrensvariationskoeffizient v_{x0} 8,52%. v_{x0} zeigt, dass eine gewisse Unsicherheit innerhalb der Kalibriergeraden vorhanden ist. Dies ist zum Beispiel auch visuell an der 10%-Mischung zu erkennen. Die Ergebnisse zeigen jedoch auch, dass die Möglichkeit einer quantitativen Abschätzung von Mischungen gegeben ist.

Daher wurden im nächsten Schritt Matrixmischungen der georgischen Haselnüsse Nr. S4 und der aserbaidischen Haselnüsse Nr. 102 hergestellt und wie unter 7.11.3 beschrieben analysiert. Erste Versuche erfolgten mit Mischungen, die 10%, 30%, 50% und 70% georgische Haselnüsse enthielten. Die resultierenden Differenzkurven sind der Abbildung 37 zu entnehmen.

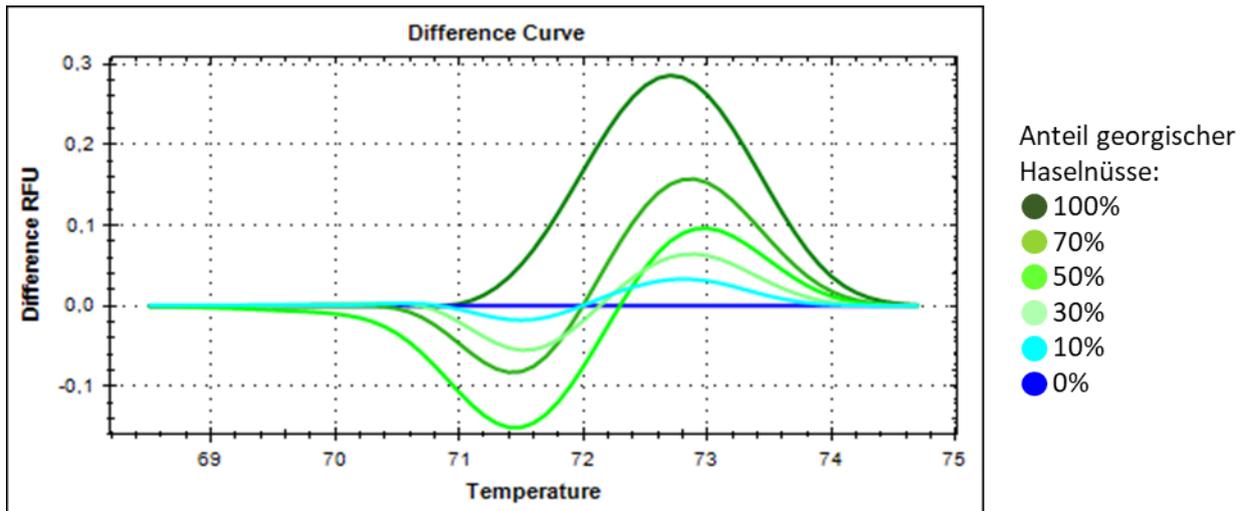


Abbildung 37: *Difference plot* von Matrixmischungen mit 0%, 10%, 30%, 50% und 70% georgische Haselnüsse. Als Referenzkurve diente die Schmelzkurve des 0%-Ansatzes.

Den Differenzkurven ist eine kontinuierliche Zunahme der Peak-Höhe mit zunehmendem Anteil an georgischen Haselnüssen vergleichbar zu den Vorversuchen mit Mischungen von DNA-Isolaten zu erkennen. Die 10%-Mischung weist bereits einen deutlich zu erkennenden Peak und einen Wert von $\Delta F > 0$ auf. Für die Erstellung der Kalibriergerade wurde ΔF unter Berücksichtigung der maximalen Peakhöhe, die der Ansatz mit 100% georgischen Haselnüssen in dem entsprechenden Lauf erreichte, in den prozentualen Anteil dieser umgerechnet und gegen den Anteil an Haselnüssen aus Georgien aufgetragen. Dieser Schritt der Normalisierung wurde durchgeführt, da die Intensitäten einer Mischung bei verschiedenen Durchläufen trotz unveränderter Reaktionsbedingungen schwankten und daher die Läufe untereinander nur schwer zu vergleichen waren. Durch die Umrechnung ist es möglich mehrere Kalibriergeraden miteinander zu vergleichen, um letztendlich eine möglichst konstante Gerade für die Quantifizierung zu erhalten. Die daraus resultierende Kalibriergerade ist Abbildung 38 zu entnehmen. Die zugehörigen Werte für die Erstellung der Kalibrierung sind unter 9.10.6 in der Tabelle 46 zusammengefasst.

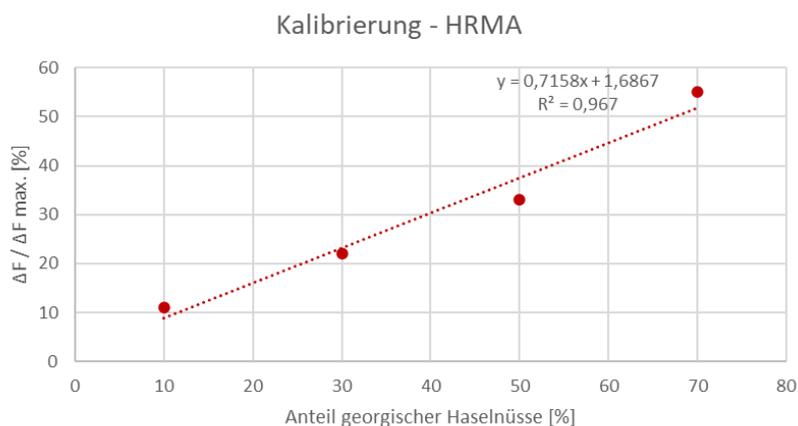


Abbildung 38: Kalibriergerade, die durch Auftragen des prozentualen Anteils von ΔF in Bezug auf ΔF des 100%-Ansatzes gegen den Anteil georgischer Haselnüsse erhalten wurde.

Bereits visuell sind bei der dargestellten Kalibriergerade Unsicherheiten festzustellen. Diese wurden durch weitere Parameter, die für die Bewertung der Kalibrierung bestimmt wurden, bestätigt ($R = 0,9833$, $s_y = 4,2$, $s_{x0} = 5,8$ und $v_{x0} = 14,6\%$). Aufgrund dessen sollte lediglich von einer semi-quantitativen Abschätzung ausgegangen werden. Um die Möglichkeit der Quantifizierung im Bereich kleiner Konzentrationen besser beurteilen zu können, wurde die bestehende Kalibrierung um zwei Matrixmischungen mit einem Anteil von 2,5% und 5% georgischer Haselnüsse erweitert. Die Messung erfolgte letztendlich im Bereich von 2,5-50% Beimischung. Der zugehörige *Difference plot* und die daraus resultierende Kalibriergerade sind im Anhang unter 9.10.6 (Abbildung 68 und Abbildung 69) dargestellt. Die Parameter zur Bewertung der Regressionsgerade sind vergleichbar zu der ersten Kalibrierung ($R = 0,99$, $s_y = 3,1$, $s_{x0} = 2,9$ und $v_{x0} = 15,0\%$). Die 10%-Mischung ist auch hier anhand eines Peaks im Bereich von 72-73 °C zu erkennen. Die beiden niedriger konzentrierten Mischungen weisen jedoch kein $\Delta F > 0$ auf. Aufgrund dieser Ergebnisse ist davon auszugehen, dass eine semi-quantitative Abschätzung ab 10% detektiert werden kann. Für die Verifizierung dieser Ergebnisse sollten jedoch weitere Kalibriersets hergestellt und analysiert werden, um die Reproduzierbarkeit der Kalibrierung und die Vergleichbarkeit der Messungen abschließend bewerten zu können. Anschließend sollte zudem ein Test-Set auf Basis der Kalibrierung quantifiziert werden, um eine Aussage über die Genauigkeit der entwickelten Methode treffen zu können.

Ein Blick auf die bereits veröffentlichten Arbeiten im Gebiet der Authentizitätsbestimmung und quantitativen HRMA im Bereich der Lebensmittel zeigt, dass es bereits verschiedene Publikationen diesbezüglich gibt. So entwickelte Sakaridis et al. (2013) eine HRMA für den Nachweis von Kuhmilch in Büffelmilch und entsprechenden verarbeiteten Produkten wie Büffelmozzarella.^[229] Für die Bewertung einer möglichen Quantifizierung wurden Mischungen aus Kuh- und Büffelmilch herangezogen. Die erhaltene Kalibriergerade weist ein im Vergleich zu der hier vorgestellten Methode leicht besseres Bestimmtheitsmaß von $R^2 = 0,988$ auf, dennoch werden auch hier die Schwankungen innerhalb der Kalibrierung gerade im Bereich niedrig konzentrierter Beimischungen diskutiert. Diese Abweichungen sind auch visuell festzustellen, werden jedoch nicht näher beurteilt. Es wird eine mögliche Detektion bereits bei 1% Beimischung angegeben.^[229] Zu beachten ist jedoch, dass Sakarides et al. (2013) mit zwei Primer-Paaren gearbeitet haben, um je ein spezifisches PCR-Produkt für Kuh- und Büffelmilch zu generieren. Da diese Amplifikate grundsätzlich bereits unterschiedliche Schmelzverhalten aufweisen, fallen hier die detektierten Unterschiede im Verlauf der Schmelzkurven größer aus. In der hier vorgestellten Arbeit ist der Unterschied der beiden Amplifikate, die sich nur aufgrund von drei SNPs ergeben bereits deutlich kleiner, weshalb auch hier die Verschiebung der Schmelzkurve erst durch die Verwendung des *Difference plots* deutlich wird und auch dort erst im Bereich höherer Konzentrationen in Bezug auf die Beimischung.

Häufig finden sich jedoch auch Veröffentlichungen, bei denen bereits durch eine rein visuelle Beurteilung der Kalibriergeraden die beschriebene lineare Korrelation, welche für eine Quantifizierung herangezogen werden soll, in Frage gestellt werden muss.^{[200],[230]} Insgesamt fehlen in der Literatur grundsätzlich die Prüfungen, ob mit den erhaltenen Kalibriergeraden verlässliche Ergebnisse erzielt werden können. So erfolgte i.d.R. auch keine Beurteilung der Schwankungen der

Kalibrierpunkte über die Berechnung der Reststandardabweichung. Die Korrelation wird lediglich anhand des Bestimmtheitsmaßes bewertet, was für eine abschließende Beurteilung der Güte einer Kalibrierung nicht ausreicht.

Zusammenfassend stellt die HRMA eine Methode mit viel Potential dar. Verschiedene Publikationen zeigten bereits die Anwendung auch im Bereich der Authentizitätsbestimmung von Lebensmitteln. Die Vorteile der HRMA sind vielseitig. So besteht die Möglichkeit eine Vielzahl an Proben gleichzeitig analysieren zu können. Des Weiteren werden aufgrund des *closed-tube*-Ansatzes die Gefahr von Kreuzkontaminationen deutlich verringert und es sind nach der PCR keine weiteren Schritte wie bspw. enzymatischer Verdau oder die Detektion mittels AGE notwendig.^[229] Zudem werden auch im Vergleich zu anderen Methoden für den Nachweis von Mutationen keine markierten Sonden benötigt, weshalb die HRMA auch eine vergleichsweise günstige Analysenmethode darstellt.^[231] Bezüglich der möglichen Quantifizierung fehlen jedoch gerade auch im Bereich der Lebensmittel ausführliche Beurteilungen inwieweit die vorgestellten Methoden für die Quantifizierung geeignet sind, da die Kalibrierungen häufig nicht ausreichend bewertet und ihre Anwendung auf Mischungen evaluiert werden.

6.4 Analyse definierter Bereiche des Karyoms

Auf dem Karyom gibt es definierte Bereiche bzw. bestimmte Gene, die sich im Bereich der Authentizitätsbestimmung und der Identifizierung von Sequenzunterschieden als Grundlage für die Entwicklung von *targeted*-Methoden etabliert haben. Die Ergebnisse der Analyse zwei dieser Bereiche werden im Folgenden vorgestellt.

6.4.1 Analyse der ITS-Bereiche

Ziel war die Identifizierung von Sequenzunterschieden innerhalb des ITS-Bereichs. Wie bereits unter 4.2.2.1 erläutert, wird dieser Bereich aufgrund seines polymorphen Charakters häufig zur Artendifferenzierung herangezogen. Die Arbeiten erfolgten in Anlehnung an die Veröffentlichung von Erdogan und Mehlenbacher (2000), welche phylogenetische Untersuchungen an verschiedenen *Corylus* Arten anhand der ITS-Bereiche des Karyoms sowie des *matK* Gens auf dem plastidären Genom durchführten.^[3] Die Amplifikation der ITS-Bereiche wurde mit dem veröffentlichten Primerpaar durchgeführt. Dieses bindet auf der 18S- bzw. 25S-Geneinheit. Das generierte PCR-Produkt umfasst damit neben den Bereichen ITS-1 und ITS-2 auch die 5,8s rDNA. Mit einer Länge von 666 bp liegt es in dem für Angiospermen üblichen Bereich von 500-700 bp.^[58]

Da der Schwerpunkt dieser Arbeit auf der *C. avellana* lag, wurde das Probenspektrum der drei veröffentlichten Sorten Casina, Barcelona und Tombul Ghiagli um elf Sorten erweitert. Dazu zählten vier europäische Proben (Butler, Ennis, Pautet, Segorbe) sowie sieben Proben aus den östlichen Anbaugebieten Aserbaidschan, Georgien sowie Türkei. Details zu dem eingesetzten Probenmaterial

ist der Tabelle 27 im Anhang unter 9.6 zu entnehmen. Informationen bzgl. der drei veröffentlichten Sequenzen sind in Tabelle 12 zusammengefasst.

Tabelle 12: Informationen bzgl. der bereits veröffentlichten Sequenzen der ITS-Bereiche dreier Haselnusssorten der Art *C. avellana*.^[3]

Sorte	Herkunft	Accession No.
<i>C. avellana</i> Casina	Spanien, Asturias	AF297333.1
<i>C. avellana</i> Barcelona	Spanien, Tarragona	AF297334.1
<i>C. avellana</i> Tombul Ghiagli	Griechenland	AF297335.1

Nach erfolgreicher Sequenzierung wurden die erhaltenen Sequenzen mit Hilfe des *Clustal Omega Alignment Tools* des EMBL-EBI miteinander verglichen. Das vollständige Alignment ist Abschnitt 9.10.7 zu entnehmen. Abbildung 39 zeigt einen Ausschnitt des Alignments, das mit der Software *Genome Compiler* grafisch dargestellt wurde.

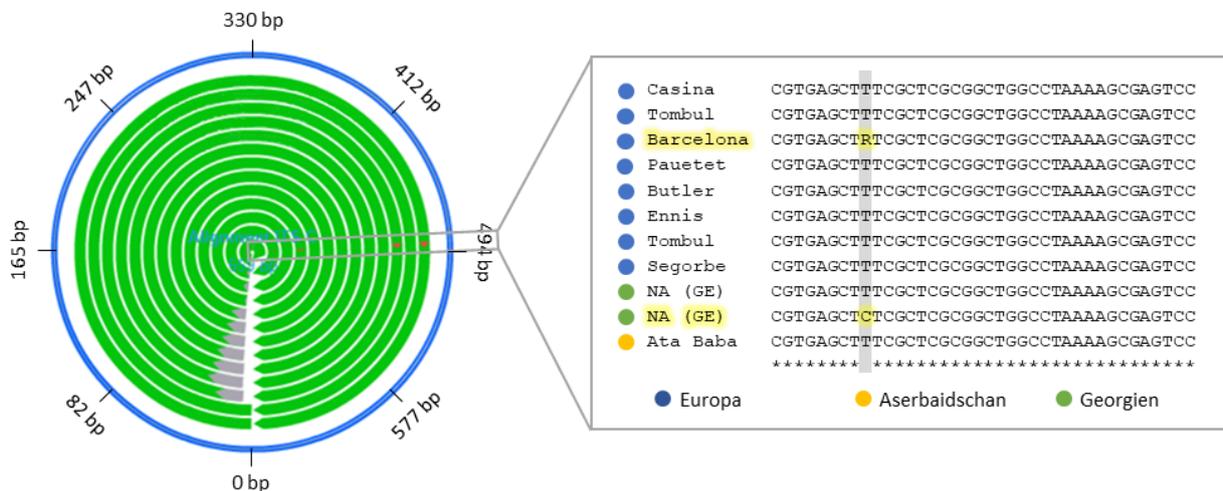


Abbildung 39: Ausschnitt des Alignments des ITS-Bereichs. Die Herkunft ist farblich gekennzeichnet. Blau: Europa, Grün: Georgien und Gelb: Aserbaidschan. Abweichende Proben sind hervorgehoben. Die Position des SNPs ist grau umrandet bzw. hinterlegt.

Links in der grafischen Darstellung des Alignments steht jeder grüne Kreis für eine der Sequenzen. Der blaue Kreis gibt die Länge in Basenpaaren an. Die Sequenzen beginnen unten bei 0 bp. Der graue Bereich zu Beginn mancher Sequenzen zeigt die getrimmten Bereiche an. Sequenzunterschiede werden durch rote Dreiecke symbolisiert. Der entsprechende Abschnitt ist rechts vergrößert dargestellt. Es ist deutlich zu erkennen, dass sich die analysierten Haselnusssorten in ihren Sequenzen der ITS-Bereiche kaum unterscheiden. Lediglich ein SNP an Position 489 wurde detektiert. An derselben Position weist auch eine der veröffentlichten Sequenzen (*C. avellana* Barcelona, AF297334.1) ein R auf. Dieses steht gemäß der IUPAC Nomenklatur für ein A oder G. Alle anderen veröffentlichten Sequenzen sowie die selbst sequenzierten Proben weisen an dieser Position ein Thymin auf. Vier der fünf Proben aus Georgien hingegen zeigten an der Position ein Cytosin. Die Sequenzierungsdaten wiesen insgesamt eine hohe Qualität auf, was an einer deutlichen Basislinientrennung der Peaks zu erkennen ist. Bei genauer Betrachtung der Position 489 ist jedoch eine Signalüberlagerung der Basen Cytosin und Thymin zu erkennen (siehe Abbildung 40).

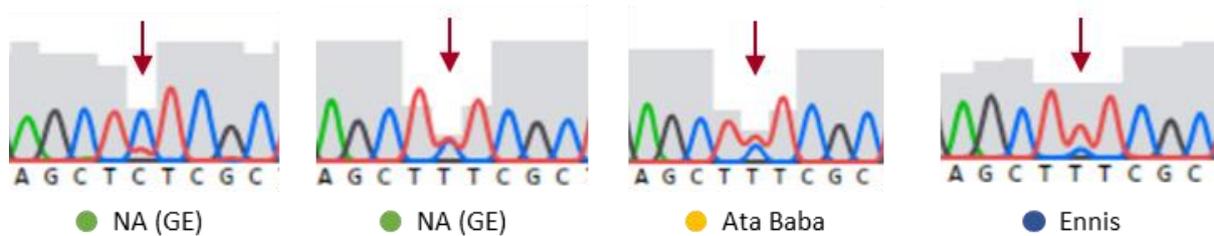


Abbildung 40: Ausschnitte aus den Sequenzierungsergebnissen von vier Haselnussorten. Die Position mit den überlagerten Signalen (Rot: Thymin, Blau: Cytosin) ist mit einem Pfeil gekennzeichnet. Die Herkunft der Proben ist farblich gekennzeichnet. Grün: Georgien, Gelb: Aserbaidschan, Blau: Europa.

Bei Überprüfung aller Sequenzierungen zeigte sich, dass die Signalüberlagerung nicht nur bei den Proben georgischer Herkunft vorhanden ist. Aufgrund der insgesamt hohen Qualität der Sequenzierungen, kann ein Sequenzierfehler ausgeschlossen werden. Vielmehr ist in diesem Fall von einem heterozygoten SNP auszugehen. Vergleichbare Ergebnisse wurden bspw. bereits von Chu et al. (2001) veröffentlicht, die ebenfalls heterozygote Basen bei der Analyse des Bereichs ITS-1 bei Crustaceae identifizierten.^[62] Generell handelt es sich jedoch bei den ITS-Bereichen um einen etablierten Ansatzpunkt für phylogenetische Untersuchungen und Analysen mit dem Ziel der Identifizierung interspezifischer Sequenzunterschiede z.B. als Grundlage für die Entwicklung von *targeted*-Methoden. Im Falle der *C. avellana* zeigte sich jedoch bereits bei den Analysen von Erdogan und Mehlenbacher (2000) auf der interspezifischen Ebene ein hoher Verwandtschaftsgrad.^[3] Intraspezifisch konnten keine Sequenzunterschiede für den Einsatz zur Authentizitätsbestimmung identifiziert werden, was auf den auf Sortenebene noch höheren Verwandtschaftsgrad zurückzuführen ist.

6.4.2 Analyse der WRKY-Gene

Neben dem ITS-Bereich wurden auch die WRKY-Gene auf Sequenzunterschiede hin untersucht. Ausgewählt wurde dieser Bereich basierend auf den Untersuchungen von Borrone et al. (2004) im Zuge dessen Sequenzunterschiede innerhalb des WRKY-typischen Introns identifiziert wurden.^[69]

Da für die *Corylus avellana* bisher noch keine Gensequenzen der WRKY-Proteine veröffentlicht wurden, mussten zunächst sequenzspezifische Primer entwickelt werden. Grundlage hierfür bildeten die von Borrone et al. (2004) entwickelten degenerierten Primer.^[69] Im Rahmen der genannten Publikation wurden unterschiedliche degenerierte Primer auf Basis veröffentlichter WRKY-Gene entwickelt. Das Primerpaar, das im Rahmen der vorliegenden Arbeit Anwendung fand (WRKY_FP1 und WRKY_RP2 siehe Tabelle 24), war das einzige, das in der Veröffentlichung ein aussagekräftiges Bandenmuster mit deutlichen und voneinander getrennten Banden bildete.

Bei degenerierten Primern handelt es sich um eine Mischung mehrerer unterschiedlicher Primer mit vergleichbarer Grundstruktur. Definiert werden die unterschiedlichen Möglichkeiten durch Basen, die in der Primersequenz angegeben sind, welche gemäß des IUPAC-Codes für unterschiedliche Basen stehen können. So steht beispielsweise R für die Basen Adenin oder Guanin und Y für Cytosin

oder Thymin. Es gibt auch die Angabe N, an deren Stelle jede beliebige Base eingesetzt werden kann. Eine weitere Besonderheit degenerierter Primer ist der Einsatz von Inosin (siehe Abbildung 41). Den Zuckerbestandteil bildet hierbei β -D-Ribofuranose und bei der Base handelt es sich um Hypoxanthin, welches vergleichbar zu Adenin und Guanin ebenfalls ein Purin als Grundgerüst aufweist.^{[108],[232]} Inosin ist normalerweise Teil der RNA und wird im Zuge des RNA-Editing durch Desaminierung von Adenosin mittels einer RNA-spezifischen Adenosindesaminase gebildet. Inosin wird im Zuge der Translation als

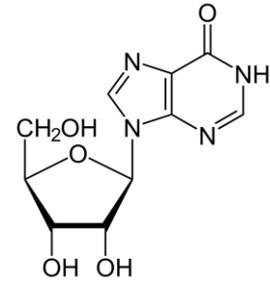


Abbildung 41:
Strukturformel von Inosin.

Guanosin gelesen, was letztendlich zu einem Einbau einer anderen Aminosäure führen kann.^{[36],[164]} Im Falle der tRNA paart es in der Wobble-Position mit Cytidin, Uridin und Adenosin. Bei degenerierten Primern wird es als „neutrale“ Base eingesetzt, welche mit allen vier Basen Wasserstoffbrückenbindungen ausbilden kann. Diese sind jedoch abhängig von der ihr gegenüberliegenden Base unterschiedlich thermodynamisch begünstigt.^[233] Durch die zusätzlich eingesetzten variablen Basen in der Primersequenz entsteht eine Mischung unterschiedlicher Primer, die sich in ihrem Grundaufbau gleichen, aber durch SNPs dennoch unterscheiden. Solche speziellen Primer kommen u.a. zum Einsatz, wenn die genaue Zielsequenz nicht bekannt ist.

Wie in Abbildung 42 dargestellt, binden die ausgewählten degenerierten Primer in den Randbereichen der WRKY-Domäne sowohl auf der Gensequenz der WRKY-Proteine der Gruppe I als auch II. Die Domäne wies im Falle von *Theobroma cacao* ein Intron auf, innerhalb dessen SNPs identifiziert wurden. Zusätzlich bindet der FW-Primer bei den Proteinen der Gruppe I auch auf der zweiten Domäne, die allerdings kein Intron aufwies.

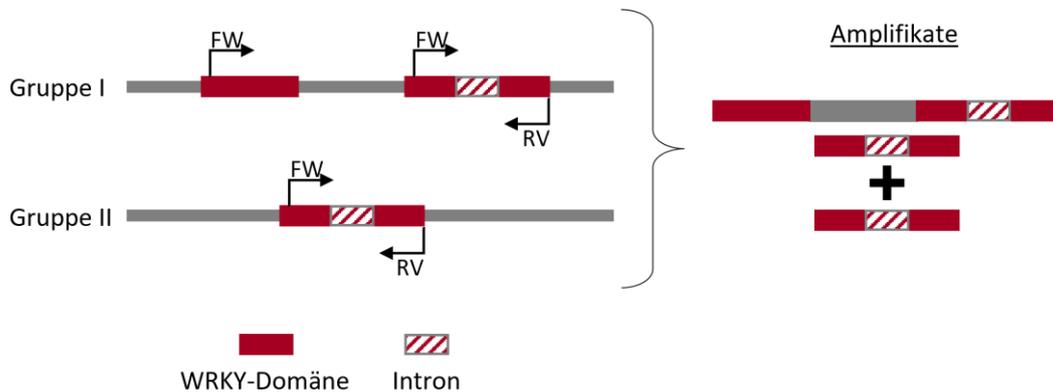


Abbildung 42: Bindungsstellen der degenerierten Primer auf den Gensequenzen der WRKY-Proteine der Gruppe I und II und den daraus resultierenden PCR-Produkten. In Anlehnung an Borrone et al. (2004).^[69]

Zunächst erfolgte die Optimierung der PCR-Bedingungen für den Einsatz der degenerierten Primer mit dem Ziel möglichst viele deutlich voneinander getrennte Banden zu erzeugen. Wichtig war hierbei der Einsatz einer *Taq* DNA Polymerase, da diese einen spezifischen Adenosin-Überhang am 3'-Ende der Amplifikate erzeugen, der essentiell für die anschließend notwendige Klonierung mit Hilfe des *TopoTM TA CloningTM Kit for Sequencing* der Firma Invitrogen (Thermo Fisher ScientificTM, Waltham, USA) ist. Der Vektor ist so konzipiert, dass er einen 3'-Thymidin-Überhang aufweist und

sich somit die Amplifikate anlagern können und die Ligation erfolgen kann. In diesem Fall wurde die *DreamTaq* DNA Polymerase der Firma Thermo Fisher Scientific™ (Waltham, USA) eingesetzt, welche die erforderlichen *sticky ends* ausbildet.

Abbildung 43 zeigt ein Agarosegel der mittels der degenerierten Primer und den optimierten PCR-Parametern erzeugten Amplifikaten drei unterschiedlicher Haselnusssorten (Nr. 9 = T. G. Trilobata, Nr. 33 = Pauetet und Nr. 102 = Ata Baba). Für alle Sorten konnte ein vergleichbares Bandenmuster generiert werden. Da die DNA-Banden sehr nah beieinander lagern, erfolgte eine doppelte elektrophoretische Auftrennung (2x 20 min).

Für die Entwicklung sequenzspezifischer Primer müssen die potentiell verschiedenen Fragmente einer Bande mittels Klonierung getrennt werden. Die erhaltenen Sequenzen können anschließend mit der Sequenzdatenbank des NCBI (*National Center for Biotechnology Information*) abgeglichen werden, um festzustellen, ob es sich um die Zielsequenz handelt. Anschließend können sequenzspezifische

Primer entwickelt werden, um schließlich die gewünschten Amplifikate sequenzieren und auf Sequenzunterschiede hin analysieren zu können. Die Banden mit den Längen 280, 320, 800 und 950 bp wurden für die Klonierung ausgewählt. Für die PCR zur Generierung des Inserts wurden DNA-Isolate der Sorte Puaetet (Nr. 31 und 33) eingesetzt. Nach der Transformation wurden pro kloniertem Fragment 18 Kolonien mittels *Colony-PCR* auf das Vorhandensein des Inserts überprüft. Jeweils sechs bis acht Kolonien wurden anschließend mittels Sanger-Sequenzierung sequenziert und die erhaltenen Sequenzen, wie unter 7.8 erläutert, miteinander verglichen und auf Unterschiede hin untersucht. Hierbei zeigte sich jeweils eine deutliche Gruppierung der Sequenzen (siehe Tabelle 13). Innerhalb einer Gruppe wiesen die Sequenzen eine sehr hohe Übereinstimmung auf.

Tabelle 13: Gruppierung der nach Klonierung erhaltenen Sequenzen der Fragmente, die mit den degenerierten Primern WRKY_FP1 und WRKY_RP2 nach Borrone et al. (2004)^[69] (siehe Tabelle 24) amplifiziert wurden.

Bande	Einteilung der Sequenzen basierend auf dem Alignment		
	Gruppe 1	Gruppe 2	Gruppe 3
280 bp	K2, K3, K4, K8, K11, K18	K1	K9
320 bp	K1, K3, K5, K6, K13, K14, K16	K18	---
800 bp	K4, K10, K11, K13, K14, K15, K16	K3	---
950 bp	K3, K12, K18	K4, K16, K15	---

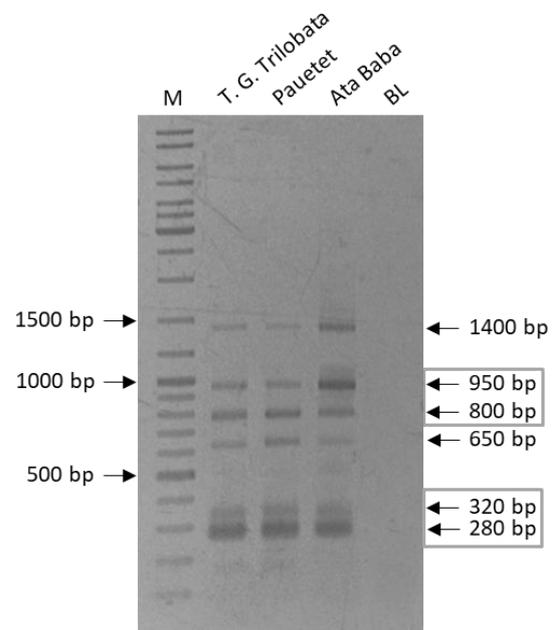


Abbildung 43: Agarosegel der Amplifikate, die mit den degenerierten Primern von Borrone et al. (2004) generiert wurden. Bei den eingerahmten Fragmentlängen handelt es sich um die Fragmente, die für die folgende Klonierung ausgewählt wurden.

Für die Überprüfung, ob es sich bei den erhaltenen Sequenzen auch um Gene von Transkriptionsfaktoren handelt, erfolgte ein Abgleich mit der Datenbank des NCBI mittels BLAST (*Basic Local Alignment Search Tool*). Dabei ergaben sich für alle Sequenzen Übereinstimmungen mit Sequenzen unterschiedlicher Transkriptionsfaktoren anderer höherer Pflanzen. Tabelle 14 zeigt beispielhaft einen Auszug der Ergebnisse für drei Kolonien der 280 bp-Bande. Der Fokus lag bei der Prüfung auf den Ergebnissen bzgl. Transkriptionsfaktoren.

Tabelle 14: Ergebnisse des Abgleichs der erhaltenen Sequenzen mit der Datenbank des NCBI mittels BLAST; beispielhaft für drei Kolonien (K1, K2, K9) der Bande mit einer Länge von 280 bp.

Bande	Kolonie	Accession number	Query Cover [%]	Percent Identity [%]
280 bp	K1	PREDICTED: <i>Quercus lobata</i> WRKY transcription factor 44	42	90,09
		PREDICTED: <i>Quercus suber</i> WRKY transcription factor 44	42	90,09
		PREDICTED: <i>Macadamia integrifolia</i> WRKY transcription factor 44-like	41	89,09
	K2	PREDICTED: <i>Juglans microcarpa x Juglans regia</i> Probable WRKY transcription factor 2	45	91,53
		PREDICTED: <i>Quercus suber</i> Probable WRKY transcription factor 2	40	90,48
		PREDICTED: <i>Carya illinoensis</i> Probable WRKY transcription factor 2	40	90,48
	K9	PREDICTED: <i>Juglans microcarpa x Juglans regia</i> Probable WRKY transcription factor 2	45	92,37
		PREDICTED: <i>Quercus suber</i> Probable WRKY transcription factor 2	40	91,43
		PREDICTED: <i>Carya illinoensis</i> Probable WRKY transcription factor 2	40	91,43

Wichtige Anhaltspunkte waren die Werte für *Query Cover* und *Percent Identity*. *Query Cover* ist ein Maß dafür, wie stark die Ausgangssequenz durch die Ergebnissequenz abgedeckt wird. Die *Percent Identity* beschreibt die Ähnlichkeit der Sequenzen. Je höher dieser Wert ist, desto ähnlicher sind sie sich. Werden diese beiden Größen gemeinsam betrachtet, so kann die Angabe der *Query Cover* zudem einen Anhaltspunkt geben, wie lange die Sequenzen im Verhältnis zueinander sind.

Für alle erhaltenen Sequenzen konnten Übereinstimmungen mit verschiedenen veröffentlichten Sequenzen von Transkriptionsfaktoren identifiziert werden. Zudem wurden ebenfalls Ergebnisse bzgl. veröffentlichter Gesamt-Genome erhalten. So war stets die 2020 veröffentlichte Sequenz der *C. avellana* der Sorte Tombul mit einem *Query Cover* von mind. 97% und einer *Percent Identity* von über 96% vorhanden. Zudem waren ebenfalls immer Gensequenzen der *Carpinus viminea* und mit Ausnahme der 800 bp-Bande der *Fagus sylvatica* enthalten, wobei die Ähnlichkeit und

Übereinstimmung mit der Sequenz von *Carpinus viminea* höher war, was durch die nähere Verwandtschaft zu erklären ist, denn diese gehört ebenfalls zur Unterfamilie der Coryloidea (Haselnussgewächse) wohingegen *Fagus sylvatica* lediglich zur selben Ordnung der Fagales (Buchenartige) gehört.

Im nächsten Schritt sollten nun Sequenzunterschiede zwischen unterschiedlichen Haselnussorten identifiziert werden. Hierfür wurden sequenzspezifische Primer entwickelt. Dabei wurde v.a. auf den GC-Gehalt und die Länge geachtet um eine für ein möglichst spezifisches Annealing ausreichend hohe Annealingtemperatur zu erhalten. Die Primer wurden zudem stets hinsichtlich möglicher Sekundärstrukturen und Dimer-Bildung mittels geeigneter Online-Tools (z.B. *Multiple Primer Analyzer* der Firma Thermo Fisher Scientific™, Waltham, USA) überprüft, da diese Faktoren die Güte einer PCR negativ beeinflussen können. Entwickelt wurden schließlich zunächst die folgenden vier Primerpaare WRKY-280bp_K1_FW/RV, WRKY-280bp_K2_FW/RV, WRKY-320bp_K3_FW/RV und WRKY-950bp_K16_FW/RV (siehe Tabelle 24 sowie Abbildung 70 bis Abbildung 73). Für die anschließende PCR mit dem Ziel das Amplifikat zu sequenzieren wurden DNA-Isolate von Blättern verwendet, um einen möglichen Bestäubereinfluss zu umgehen, da das Karyom im Gegenteil zum cpGenom nicht rein maternal vererbt wird. Nach der PCR mit dem Primerpaar WRKY-280bp_K1_FW/RV konnte lediglich eine sehr schwache Bande mit der entsprechenden Länge auf dem Gel detektiert werden. Zusätzlich war ein Amplifikat mit einer Länge von 800 bp zu erkennen. Auch Optimierungen der PCR-Bedingungen zeigten kein verbessertes Ergebnis. Mit dem Primerpaar WRKY-280bp_K2_FW/RV hingegen konnte ein Fragment der gewünschten Länge ohne störende Nebenprodukte amplifiziert werden. Das Agarosegel der Amplifikate generiert mittels PCR und dem Primerpaar WRKY-320bp_K3_FW/RV wies zunächst lediglich Schlieren ohne definierte Banden auf. Nach Optimierung des Temperaturprogramms und Erhöhung der Annealingtemperatur auf 65,0 °C konnte eine scharfe Bande mit einer Länge von 300 bp generiert werden. Mit Hilfe des vierten Primerpaares WRKY-950bp_K16_FW/RV konnte trotz unterschiedlicher Optimierungsschritte kein Amplifikat der gewünschten Länge amplifiziert werden. Die PCR für die anschließende Sequenzierung wurde schließlich mit DNA-Isolaten gewonnen aus Blattmaterial der Sorten Ata Baba, Pauetet und Palaz durchgeführt. Für die Amplifikation mit dem Primerpaar WRKY-280bp_K2_FW/RV wurden die Proben Nr. B1, B9 und B12 und für die PCR mit den Primern WRKY-320bp_K3_FW/RV die Proben B2, B9 und B11 eingesetzt. Die Sanger-Sequenzierung erfolgte über einen externen Dienstleister (GATC Biotech; heute: Eurofins Genomics). Für den anschließenden Sequenzvergleich wurde ein Alignment, wie unter 7.8 beschrieben, erstellt. Abbildung 44 zeigt einen Ausschnitt des Alignments. Gemäß diesem weist die Probe B12 der Sorte Puaetet einen SNP (A/T) auf.

```

WRKY_280K2_Pauetet  GTTGTCCCCATTTTCAGTTTGGCAACAAGTTGTCCCCATTTAAGAGGACAAACGAGAAGATAAATGAGTTCACGGTGATTATTA
WRKY_280K2_Palaz   GTTGTCCCCATTTTCAGTTTGGCAACAAGTTGTCCCCATTTAAGAGGACAAACGAGAAGATAAATGAGTTCACGGTGATTATTA
WRKY_280K2_Ata Baba GTTGTCCCCATTTTCAGTTTGGCAACAAGTTGTCCCCATTTAAGAGGACAAACGAGAAGATAAATGAGTTCACGGTGATTATTA
*****

```

Abbildung 44: Ausschnitt des Alignments der Sequenzen amplifiziert mit Hilfe des Primerpaares WRKY-280bp_K2_FW/RV und Blattisolaten der Proben Nr. B1 (Ata Baba), B9 (Palaz) und B12 (Pauetet). Ein * markiert jeweils die Übereinstimmung aller drei Sequenzen. Der vorhandene SNP ist rot hinterlegt.

Bei genauer Betrachtung weisen jedoch alle Proben an der Position des vermeintlichen SNPs eine Signalüberlagerung auf (siehe Abbildung 45). Die Proben der beiden anderen Sorten Palaz und Ata Baba weisen innerhalb derselben Sequenz eine zweite Base mit einer deutlichen Überlagerung des T- und G-Signals auf.

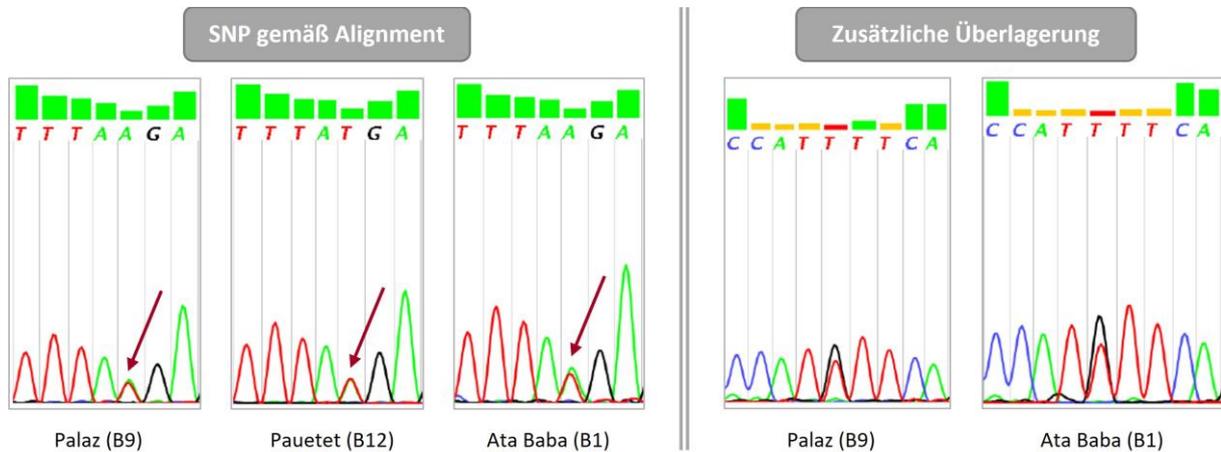


Abbildung 45: Ausschnitte der visuellen Darstellung der Sequenzierungsergebnisse der Fragmente, die mit Hilfe der Primer WRKY-280bp_K2_FW/RV amplifiziert wurden. Links: SNP gemäß des Alignments (Abbildung 44) mit der Überlagerung der A-/T-Signale. Rechts: Zusätzlicher Bereich mit einer Signalüberlagerung bei den Proben B1 und B9.

Es konnten innerhalb der Sequenzen basierend auf dem 280 bp-Fragment keine homozygoten SNPs identifiziert werden. Anhand des Alignments der PCR-Fragmente generiert mit dem Primerpaar WRKY-320bp_K3 konnten keinerlei Sequenzunterschiede identifiziert werden. Die visuelle Darstellung der Sequenzierung zeigte jedoch ebenfalls bei den Proben B2 (Ata Baba) und B5 (Pautet) eine Base mit einer Signalüberlagerung (T/G) (siehe Abbildung 74, Abschnitt 9.10.8). Die vollständigen Alignments der Sequenzierungen sind unter 9.10.8 aufgeführt.

Zusammenfassend konnten lediglich Unterschiede zwischen den Allelen also heterozygote SNPs identifiziert werden. Aufgrund dieser Ergebnisse und der als gering einzuschätzenden Erfolgsaussichten sowie des vorhandenen Einflusses des Bestäubers aufgrund der biparentalen Vererbung des Karyoms wurden keine weiteren Analysen der WRKY-Gene durchgeführt.

6.5 Analyse von SCoT Polymorphismen

Neben den bisher vorgestellten zielgerichteten Methoden, wurde auch ein *non targeted*-Ansatz untersucht. Ziel war die Identifizierung von SCoT-Markern. Die Methodik ähnelt der der RAPD-PCR. Es wird ebenfalls mit nur einem Primer gearbeitet, der als FW- und RV-Primer fungiert. Der Unterschied zu den i.d.R. nur 10 bp langen RAPD-Primern ist jedoch, dass die SCoT-Primer mit ca. 18 bp länger sind und auf Basis von Pflanzengenomen entwickelt wurden mit dem Ziel im konservierenden Bereich des ATG Startcodons zu binden.^{[163],[176]}

6.5.1 Methodenentwicklung und -optimierung

Für die Analyse von SCoT-Polymorphismen wurden insgesamt dreizehn unterschiedliche Haselnussorten mit neun SCoT-Primern analysiert. Die Gesamt-DNA-Isolate wurden mittels CTAB-Methode hergestellt. Bei der Wahl des Untersuchungsmaterials wurden, sofern möglich, pro Sorte drei Proben unterschiedlicher Herkunft ausgewählt.^[234] Bei den verwendeten Primern handelt es sich um eine Auswahl der von Collard und Mackill (2009) entwickelten Primer.^[163]

Zunächst wurden die PCR-Bedingungen optimiert. Hierfür wurde jeweils SCoT-Primer 13 verwendet. Ausgangssituation stellte dabei die Publikation von Collard und Mackill (2009) dar.^[163] Ziel war die Generierung eines möglichst umfangreichen Bandenmusters mit einer Vielzahl an Banden, die aber gleichzeitig auf dem Agarosegel deutlich voneinander getrennt werden können. Bezüglich des Temperaturprogramms zeigte das von Collard und Mackill (2009) veröffentlichte die besten Ergebnisse.^[163] Der PCR-Ansatz wurde in Bezug auf seine Zusammensetzung hin optimiert. Sehr entscheidend war dabei die Wahl der Polymerase. Zu Beginn wurde die *Taq* DNA Polymerase der Firma Biozym Scientific GmbH (Hessisch Oldendorf, Deutschland) verwendet. Diese lieferte jedoch sehr verwaschene Bandenmuster. Mit der *DreamTaq* DNA Polymerase der Firma Thermo Fisher Scientific™ (Waltham, USA) hingegen wurde ein Bandenmuster mit scharfen und deutlich voneinander getrennten Banden erhalten. Es zeigte sich zudem, dass mit der *DreamTaq* DNA Polymerase auch Amplifikate über 3000 bp generiert wurden. Die maximale Fragmentlänge mit der Biozym *Taq* DNA Polymerase war < 3000 bp. In anderen Publikation wiesen die Amplifikate Längen im Bereich von 100-2000 bp auf.^{[163],[169]} Die Länge der Banden wird auch durch das Ziel-Genom beeinflusst, da keine Banden generiert werden können, wenn die nicht kodierenden Bereiche zwischen den Genen zu groß sind und damit die Primerbindungsstellen in zu großer räumlicher Entfernung liegen. Gleichzeitig spielt aber auch, wie die Ergebnisse zeigen, die Geschwindigkeit der DNA Polymerase eine entscheidende Rolle. Die Menge an Template im PCR-Ansatz musste im Vergleich zu anderen Veröffentlichungen, in denen 10-25 ng DNA bezogen auf einen 10 µL PCR-Ansatz eingesetzt wurden, enorm reduziert werden.^{[163],[172],[235]} Mit nur 1 ng DNA wurde für den Großteil der verwendeten Primer das beste Resultat erzielt in Bezug auf eine klare Trennung der Banden und damit verbunden einer guten Auswertbarkeit des Bandenmusters. Lediglich für zwei Primer (SCoT 7 und 17) musste die Menge an Template auf 3 ng erhöht werden, damit die

Bandenintensität ausreichend war. Wie bereits erwähnt, zeigte das von Collard und Mackill (2009) veröffentlichte Temperaturprogramm mit einer Annealingtemperatur von 50,0 °C die besten Ergebnisse. Nur für Primer 21 musste die Annealingtemperatur auf 46,4 °C reduziert werden, um ein klares Bandenmuster zu erhalten.

Nach beendeter Optimierung der Reaktionsbedingungen erfolgte die PCR aller dreizehn Sorten mit den neun SCoT-Primern. Die abschließend verwendeten Parameter sind unter 7.13 zusammengefasst. Für fünf Proben konnten mit jeweils einem Primer keine Banden generiert werden. So war es mit Primer SCoT 36 für alle drei Proben der Sorte Corabel (Probennr.: F2, 32 und 35) nicht möglich ein Bandenmuster zu generieren. Zudem konnten mit Primer SCoT 24 für Probe Nr. 109 (Sorte: Butler) und mit Primer SCoT 7 für Probe Nr. 113 (Sorte: T. G. Romana) keine Amplifikate erzeugt werden. Auch eine Wiederholung der PCR brachte keine Ergebnisse. Da die eingesetzten DNA-Isolate und Chemikalien auch für andere Reaktionsansätze eingesetzt wurden, konnten diese als Ursache ausgeschlossen werden. Auch eine andere Ursache konnte nicht identifiziert werden, daher wurden für die betroffenen Proben alle Banden jeweils als nicht vorhanden gewertet. Insgesamt zeigten die Bandenmuster eine Vielzahl an Unterschieden. Es waren jedoch ebenfalls Inhomogenitäten innerhalb der Bandenmuster unterschiedlicher Proben einer Sorte zu erkennen. Aufgrund dessen war es nicht möglich einen Sortenbaum zu erstellen, der zur Sortenidentifizierung herangezogen werden kann, so wie es bspw. von Felbinger et al. (2020) oder Öztürk et al. (2018) publiziert wurde.^{[234],[236]}

Dennoch wurden die Daten dahingehend untersucht, ob sie für eine phylogenetische Verwandtschaftsanalyse geeignet sind. Hierfür wurden zwei unabhängige multivariate Methoden herangezogen. Zum einen wurde ein *Neighbour-Joining*-Diagramm erstellt und zum anderen wurde eine Hauptkomponentenanalyse (*principal component analysis* = PCA) durchgeführt. Für die Auswertung der Gelbilder mit den genannten Methoden wurden die Ergebnisse in eine 1/0-Matrix umgewandelt. „1“ steht dabei für „Bande vorhanden“ und „0“ für „Bande nicht vorhanden“. Proben für die keine Banden generiert werden konnten wurden ebenfalls als „0“ gewertet. Die Matrix ist unter 9.10.9 aufgeführt. Alle betrachteten Banden aller neun Primer bildeten letztendlich 155 Variablen, die mit Hilfe der multivariaten Methoden ausgewertet wurden.

6.5.2 Auswertung mittels *Neighbour-Joining*-Methode

Für die Bewertung der Ähnlichkeit bzw. des Verwandtschaftsgrades wurde ein *Neighbour-Joining*-Diagramm mit Hilfe des Programms PAST erstellt.^[237] Die Berechnung erfolgte zunächst basierend auf den Algorithmen (*Similarity Index*) nach Jaccard, Dice und Euclidean. Zusätzlich wurde die Auswirkung der Reduzierung des Bootstraps von 10.000 auf 1.000 bewertet. Der Bootstrap gibt dabei die Anzahl der zugrunde liegenden Subdatensets an, anhand derer der finale Verwandtschaftsbaum berechnet wird und liegt i.d.R. im Bereich von 100 bis 10.000. Die anhand der drei unterschiedlichen Algorithmen erhaltenen Diagramme wiesen alle dieselbe Grundstruktur auf. Die in Abbildung 46 gezeigten Gruppen waren stets in derselben Zusammensetzung vorhanden. Lediglich die Anordnung

der Gruppen 3-5, bei denen es sich um die überwiegend französischen Proben handelt, waren z.T. vertauscht. Die Reduzierung der Anzahl an Bootstraps zeigte keine relevanten Änderungen in Bezug auf die Bootstrap-Werte. Für die finale Auswertung wurde der *Similarity Index* nach Jaccard und ein Bootstrap von 10.000 gewählt.

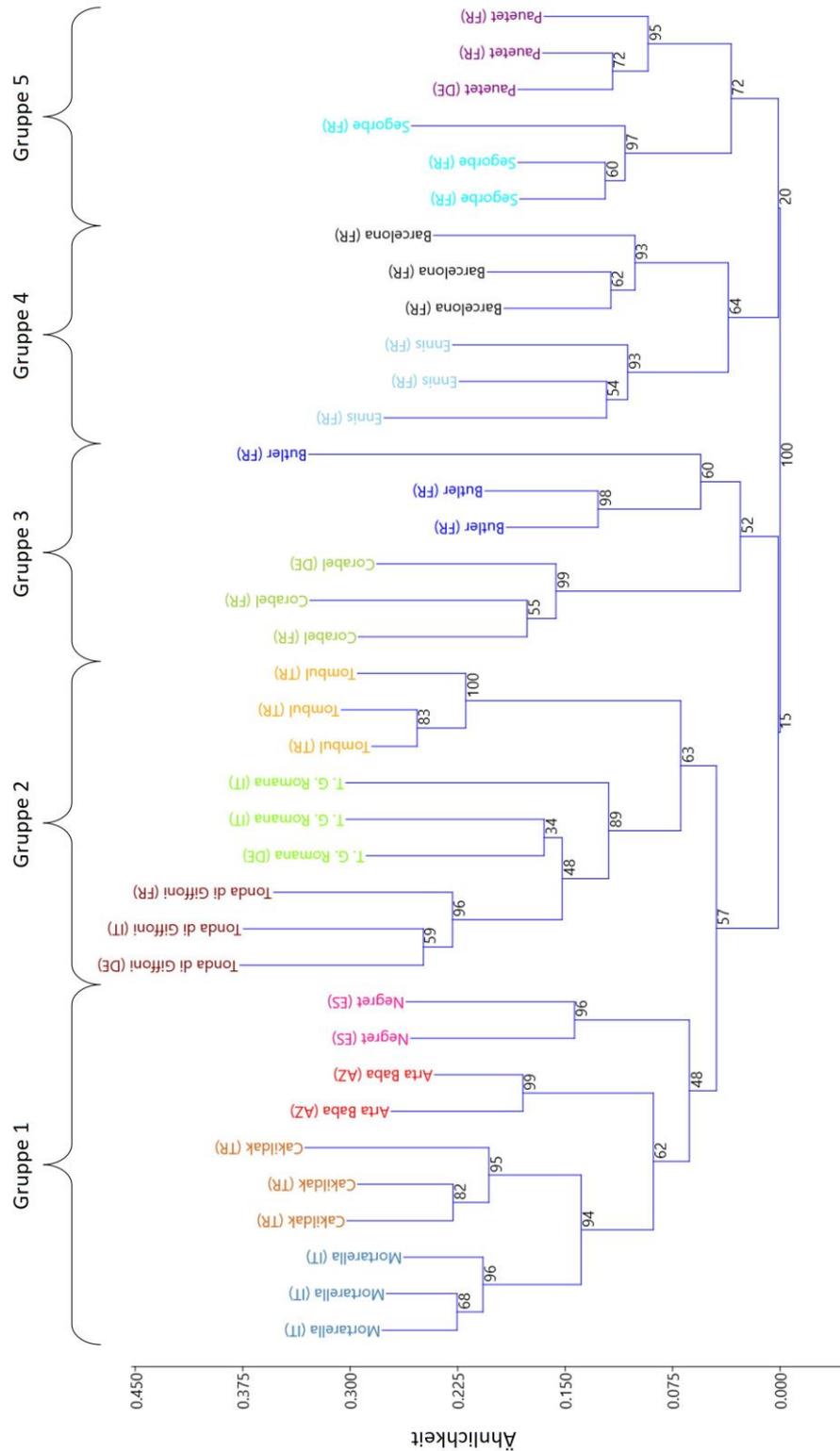


Abbildung 46: *Neighbour-Joining*-Diagramm auf Basis der 1/0-Matrix (siehe Tabelle 47 bis Tabelle 59) berechnet mit dem Jaccard-Index und einem Bootstrap von 10.000.

Auf den ersten Blick ist bereits zu erkennen, dass jeweils alle Proben einer Sorte in räumlicher Nähe zueinander liegen. Das Dendrogramm zeigt eine Unterteilung in fünf Gruppen. In Gruppe 1 clustern die Sorten Mortarella, Çakildak, Ata Baba und Negret zusammen. Die türkische Sorte Çakildak bildet dabei mit der Sorte Mortarella, die ihren Ursprung in Italien hat eine Untergruppe. Die Sorte Ata Baba aus Aserbaidshan sowie die spanische Sorte Negret bilden jeweils einen eigenen Ast aus, wobei die Sorte Ata Baba eine etwas nähere Verwandtschaft zu der zuvor genannten Untergruppe aufweist, da der entsprechende Ast einen Knotenpunkt näher liegt.^{[2],[238]} Gruppe 2 bilden die italienischen Sorten Tonda di Giffoni und T. G. Romana zusammen mit der türkischen Sorte Tombul.^{[12]–[14],[238]} Die italienischen Sorten clustern dabei jedoch näher zueinander. Innerhalb der Dreifachbestimmung der Sorte T. G. Romana zeigt sich, dass eine Probe im Vergleich stärker abweicht, da sie einen Knotenpunkt weiter entfernt liegt. Die Gruppen 3-5 setzen sich schließlich überwiegend aus Probenmaterial aus Frankreich zusammen. Gruppe 3 bilden die Sorten Corabel und Butler. Gruppe 4 Ennis und Barcelona und Gruppe 5 die Sorten Segorbe und Pauetet. Lediglich die Sorte Corabel hat ihren Ursprung in Frankreich.^[8] Die Sorten Segorbe, Barcelona und Pauetet sind ursprünglich in Spanien beheimatet und die Sorten Butler und Ennis kommen aus den USA.^[8] Dennoch sind sie nah miteinander verwandt, da, wie Untersuchungen auf Basis der Mikrosatelliten von Gökirmak et al. (2009) zeigten, sowohl die Sorte Corabel, als auch Butler und Ennis von der Sorte Barcelona abstammen. Die Sorte Pauetet hingegen weist eine enge Verwandtschaft mit der Sorte Negret auf.^{[8],[238]} Der Verwandtschaftsbaum zeigt deutlich, dass jeweils die Proben einer Sorte beieinander liegen, jedoch kein eindeutiger Einfluss des Anbaugebiets oder der ursprünglichen Sortenherkunft zu erkennen ist. So liegt bspw. die Sorte Mortarella nicht in derselben Gruppe wie die beiden anderen analysierten italienischen Sorten Tonda di Giffoni und T. G. Romana. Ebenso liegt die Sorte Negret deutlich getrennt von den anderen Sorten spanischen Ursprungs (Barcelona, Segorbe, Pauetet) sowie speziell der Sorte Pauetet bei der eine Verwandtschaft zu der Sorte Negret nachgewiesen wurde.^[8]

Die Länge der Äste sowie die Bootstrap-Werte spiegeln die bereits geschilderten Differenzen innerhalb der Dreifachbestimmung einer Sorte wieder. Bezüglich der Astlänge ist eine Probe der Sorte Butler (Nr. 109) besonders auffällig. Diese stammt aus Aquitanien (Frankreich) aus dem Jahr 2015. Die beiden anderen Proben dieser Sorte wurden 2014 ebenfalls in Aquitanien und den Midi-Pyrénées geerntet. Probe Nr. 109 clustert zwar über einen Knotenpunkt mit den anderen beiden Proben zusammen ohne, dass andere Sorten beteiligt sind, allerdings sticht die Probe mit einem sehr langen Ast im Vergleich zu dem Gesamtbild deutlich hervor. Würden alle Proben einer Sorte dieselben Ergebnisse aufweisen, wäre kein Ast zu erkennen, sondern die Proben würden direkt auf der waagrechteten Linie an dem gemeinsamen Knotenpunkt liegen und der Bootstrap-Wert, der einen Anhaltspunkt für die Wahrscheinlichkeit der Ausbildung eines Astes darstellt, betrüge 100.

Grundsätzlich zeigt das Dendrogramm, dass eine Sortendifferenzierung basierend auf der Analyse der SCoT-Polymorphismen möglich ist. Das Ergebnis sollte jedoch durch die Analyse eines größeren Probenpools verifiziert werden. Des Weiteren wäre ein Ansatz für die leichtere Einsetzbarkeit im Rahmen der Authentizitätsbestimmung die Überprüfung, ob eine Differenzierung auch mit einer verringerten Anzahl an Primern möglich wäre.

6.5.3 Auswertung mittels Hauptkomponentenanalyse

Neben dem *Neighbour-Joining*-Diagramm wurde eine Hauptkomponentenanalyse (*Principle component analysis* = PCA) als zweite unabhängige Methode für die Bewertung der Analyseergebnisse durchgeführt. Es handelt sich hierbei um eine Möglichkeit der grafischen Darstellung der Probengruppen anhand der Varianz. Das *Neighbour-Joining*-Diagramm bewertet im Vergleich dazu Ähnlichkeiten. Bei beiden Methoden handelt es sich um nicht supervisierte Ansätze.

Folgende Abbildung zeigt den 3D-Plot unter Berücksichtigung der ersten drei Hauptkomponenten.

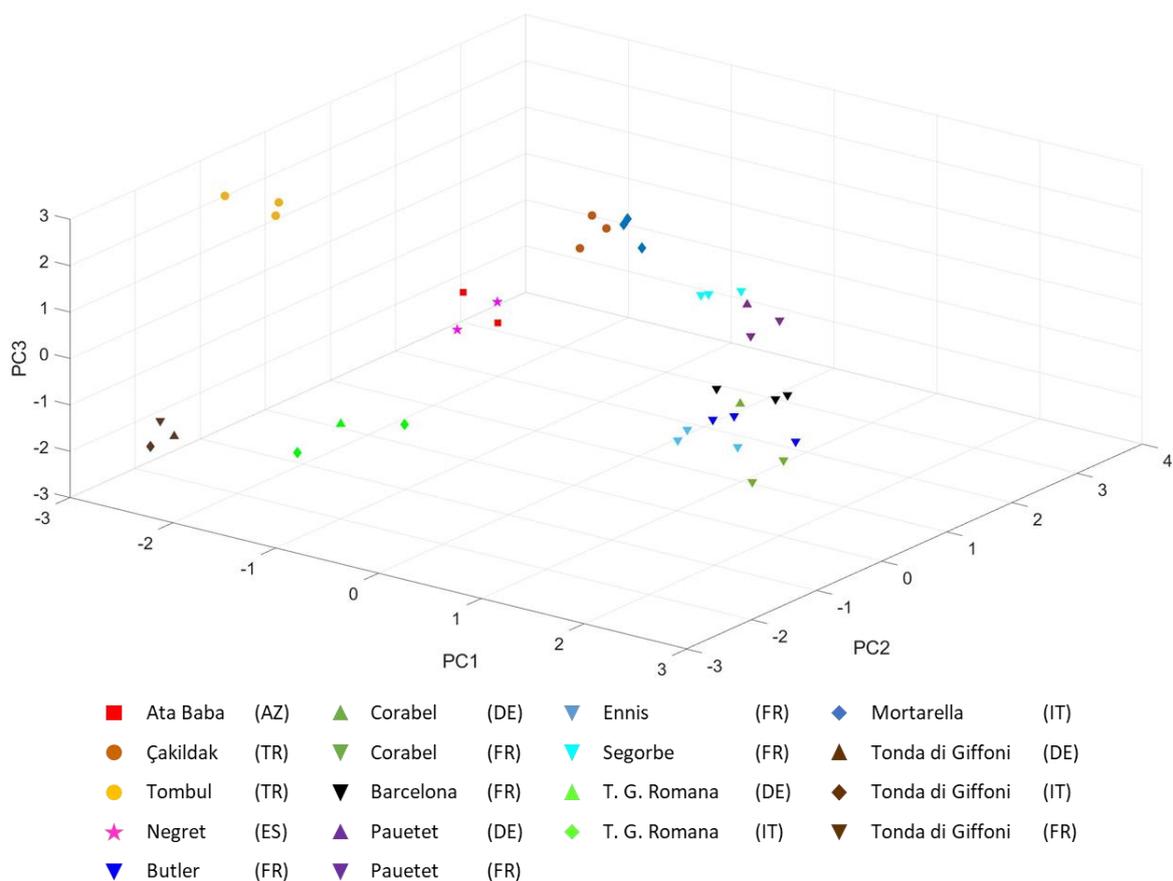


Abbildung 47: PCA - 3D-Plot der ersten drei Hauptkomponenten.

Der Anteil der erklärten Varianz der ersten drei Hauptkomponenten beträgt 15%, 14% und 9%. Der Plot zeigt, dass die Proben v.a. sortenabhängig clustern. Sehr gut ist das bspw. anhand der Sorte Tonda di Giffoni zu sehen, da die drei untersuchten Proben aus drei unterschiedlichen Anbauländern kommen und dennoch im Plot nah beieinander liegen. Die zugehörigen 2D-Plots zeigen, dass anhand der ersten Hauptkomponente (PC1) zwei Gruppen unterteilt werden können (siehe Abbildung 48). Alle Proben aus Frankreich (zugehörig zu den Sorten Ennis, Barcelona, Butler, Segorbe, Pauetet und Corabel) weisen dabei einen positiven Score in Bezug auf PC1 auf. Auch die zwei Proben aus Deutschland der Sorten Corabel und Pauetet clustern zu den anderen Proben dieser Sorten. Alle anderen Sorten weisen einen negativen PC1-Score auf. Anhand der zweiten Hauptkomponente ist eine leichte Abtrennung der Sorten Ennis und Barcelona ($PC2 < -0,25$) von den anderen französischen

Proben zu erkennen. Die zweite Gruppe mit $PC1 < 0$ zeigt basierend auf $PC2$ eine deutliche Auftrennung. Die Sorten Mortarella und Çakildak bilden dabei eine Gruppe mit einem Score von etwa 3. Mit Score-Werten von ca. 0-1 clustern die Sorten Ata Baba und Negret zusammen. Eine dritte Gruppen mit einem Score < -1 bilden die Sorten Tombul, T. G. Romana und Tonda di Giffoni.

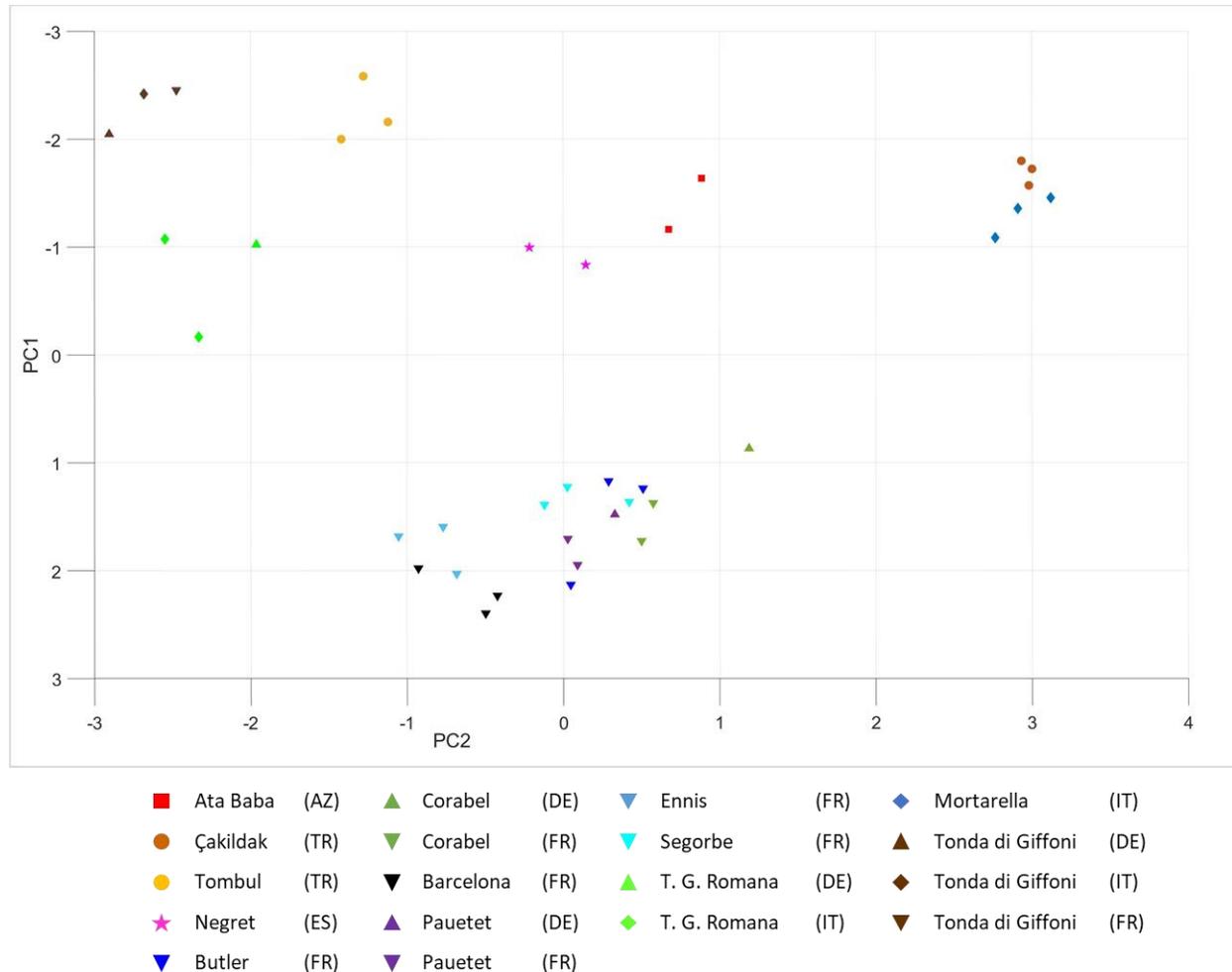


Abbildung 48: PCA - 2D-Plot der grafischen Darstellung der PC1 in Abhängigkeit von PC2.

Auch der 2D-Plot der PC2 in Abhängigkeit von PC3 (siehe Abbildung 49) zeigt noch einmal deutlich die Abtrennung der Sorten Mortarella und Çakildak. Auf Basis der PC2 bzw. PC3 ist keine weitere Trennung der beiden Sorten voneinander möglich. Bei der Sorte Tombul hingegen zeigt sich auf Basis der PC3 eine deutliche Abtrennung von den Sorten T. G. Trilobata und Tonda di Giffoni. Bei den Proben der Sorte Tombul ist zu beachten, dass es sich um Handelsproben handelt, bei welchen es sich unter Berücksichtigung der Ergebnisse der Analyse des cpGenoms häufig um Mischungen handelt.

Bei Betrachtung der Gruppe der überwiegend französischen Proben, die sich auf Basis der PC1 deutlich von den anderen Proben abtrennt, zeigt sich anhand der PC3 eine leichte Auftrennung. So bilden die Sorten Segorbe und Pautetet eine Untergruppe mit einem Score > 1 . Auch die Sorten Barcelona und Ennis, die basierend auf der zweiten Hauptkomponenten sehr nahe beieinander liegen, trennen sich. Die dritte Gruppe bilden die Sorten Butler und Corabel mit Scores von -3 bis -1.

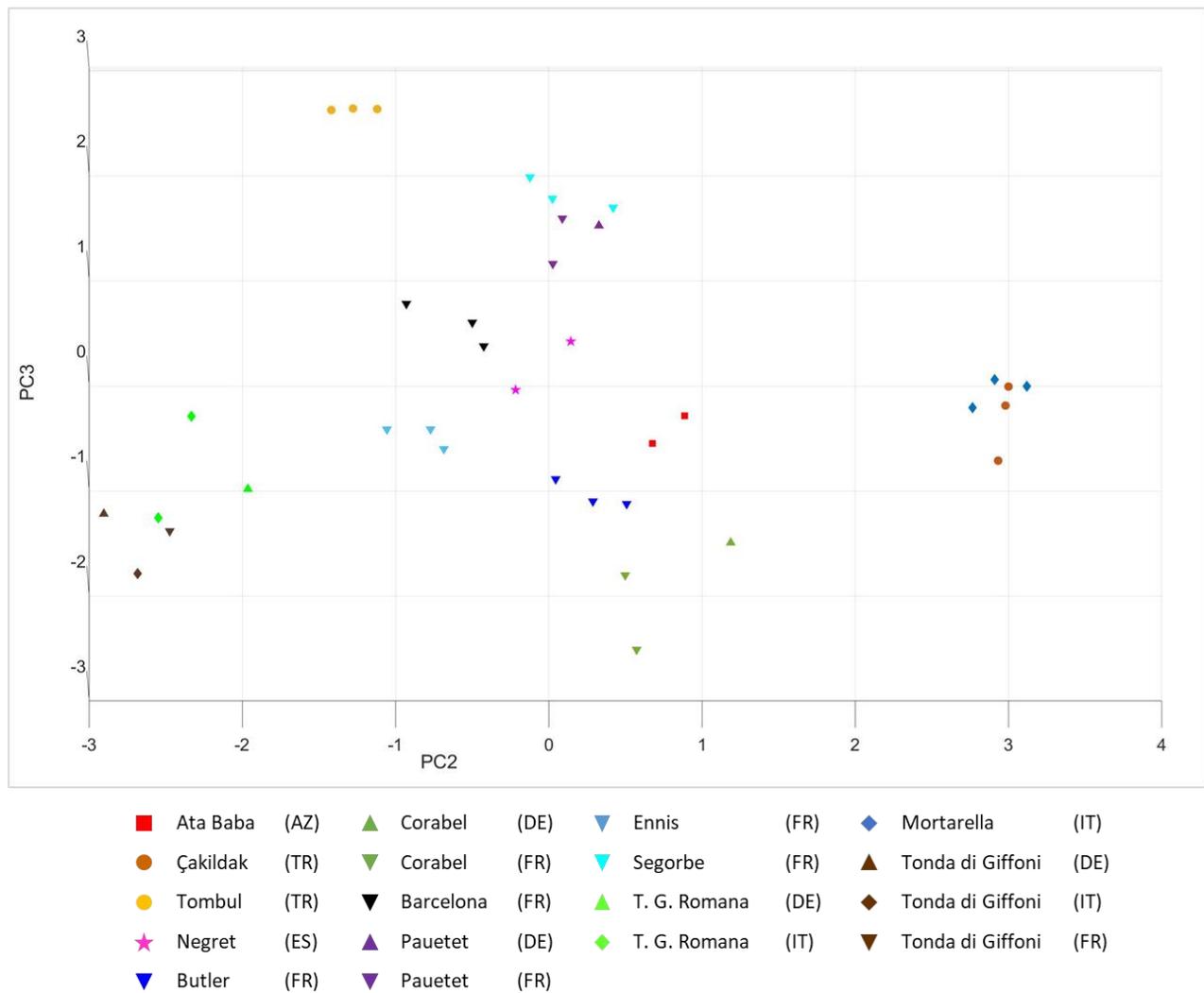


Abbildung 49: PCA - 2D-Plot der grafischen Darstellung der PC3 in Abhängigkeit von PC2.

Insgesamt basiert die Analyse auf einer sehr hohen Anzahl an Variablen im Verhältnis zu der Anzahl an Proben. Dennoch ist eine deutliche Trennung der Haselnussorten zu erkennen. Die Ergebnisse der PCA zeigen dieselben Gruppierungen wie sie auch die Verwandtschaftsanalyse mittels des *Neighbour-Joining*-Diagramms verdeutlicht. Damit konnten mit zwei unabhängigen Methoden vergleichbare Ergebnisse erzielt werden. Daher ist davon auszugehen, dass in dem vorliegenden Fall die hohe Anzahl an Variablen keinen negativen Einfluss auf das Ergebnis der PCA hat.

6.5.4 Zusammenfassung und Diskussion der Analyse von SCoT-Polymorphismen

Zusammenfassend ist zu sagen, dass basierend auf der Analyse von SCoT-Polymorphismen eine Sortendifferenzierung mit Hilfe multivariater Ansätze möglich ist. Ein Sortenbaum vergleichbar zu dem von Felbinger et al. (2020) oder Öztürk et al. (2018) wurde auf Basis der Bandenmuster nicht entwickelt, da die Varianz innerhalb der Proben einer Sorte als zu hoch eingestuft wurden. Zudem zeigte sich, dass auch die Reproduzierbarkeit der Ergebnis nur schwer zu gewährleisten ist. In Publikationen wird häufig die im Vergleich zur RAPD-PCR bessere Reproduzierbarkeit

beschrieben.^{[163],[239]} Dies kann darauf zurückgeführt werden, dass die PCR mit einer im Vergleich zu ca. 35 °C deutlich höheren Annealingtemperatur von 45-50 °C durchgeführt wird. Zudem sind die entwickelten Primer mit 18 bp fast doppelt so lang wie RAPD-Primer mit nur 10 bp. Diese beiden Parameter führen zu einer spezifischeren Bindung der Primer, was wiederum die Empfindlichkeit gegenüber Veränderungen der Reaktionsbedingungen verringert. Des Weiteren wurden die SCoT-Primer dahingehend entwickelt spezifisch im Bereich von Startcodons zu binden und nicht ungerichtet auf dem Genom.^{[163],[239]} Trotz der genannten Aspekte, die für eine erhöhte Reproduzierbarkeit sprechen, zeigten die vorliegenden Ergebnisse, dass sehr auf die Rahmenbedingungen geachtet werden muss, um eine ausreichende Reproduzierbarkeit zu gewährleisten. Bezüglich der Auswertung liegt der Vorteil eines Sortenbaums darin, dass eine reduzierte Anzahl an Reaktionsansätzen benötigt wird. Die Analyse mittels multivariater Methoden basiert hingegen auf einer statistischen Auswertung einer größeren Anzahl an Parametern wodurch Abweichungen einzelner Proben, bspw. aufgrund einer nicht vollständig gewährleisteten Reproduzierbarkeit, ausgeglichen werden können.

Insgesamt wurden von Collard und Mackill (2009) 36 Primer entwickelt.^[163] Im Rahmen der vorliegenden Analyse wurden neun davon eingesetzt. Andere Publikationen setzten mit 13-20 Primern mehr Primer für die Methodenentwicklung ein.^{[169],[172],[235]} Ein Ansatz zur Verbesserung der Ergebnisse stellt damit der Einsatz einer höheren Primeranzahl dar. Damit einhergehend nimmt jedoch auch der analytische und zeitliche Aufwand zu. Vorteil einer solchen *non targeted*-Analyse hingegen ist die einfache Durchführbarkeit und die Tatsache, dass keine spezifischen Sequenzinformationen vorliegen müssen. Ein Nachteil von *non targeted*-Ansätzen ist jedoch, dass Mischungen unterschiedlicher botanischer Herkunft nicht identifiziert werden können. Ebenso können sie dazu führen, dass diese, wenn sie unerkannt bleiben, statistische Modelle wie die PCA verfälschen. So zeigte sich bspw. im Zuge der Analyse des cpGenoms, dass es sich bei einer Probe (Nr. 120) der Sorte Tombul um eine Mischung handeln muss, da sie bei der Analyse bzgl. der *Indels* eine Doppelbande aufwies (siehe 6.3.1.1). Die beiden anderen Proben dieser Sorte wurden in Bezug auf das cpGenom nicht mehr analysiert, da im Verlauf der Arbeit anderes Probenmaterial vorlag, das in Bezug auf die Angaben der botanischen Herkunft als verlässlicher bewertet wurde. Daher können zu diesen beiden Proben keine Aussagen getroffen werden, ob es sich ebenfalls um Mischungen handelt. Es zeigte sich jedoch, dass anhand beider Auswertemethoden die drei Proben der Sorte Tombul dennoch in räumlicher Nähe zueinander lagen. Ein weiterer Nachteil dieser *non targeted*-Methoden ist, dass sie nicht auf verarbeitete Produkte anwendbar sind und daher sowohl für die verarbeitende Industrie als auch die Lebensmittelüberwachung häufig nicht einsetzbar sind. Die mechanische und thermische Beanspruchung führt dazu, dass die DNA degradiert. Daraus resultieren bzgl. ihrer Länge sehr variable DNA-Fragmente, die bei einer anschließenden Analyse zu keinen reproduzierbaren Ergebnissen führen.

Insgesamt handelt es sich bei der Analyse von SCoT-Polymorphismen um eine bereits vielfach angewandte Methode. Besonders geeignet ist sie im Bereich phylogenetischer Untersuchungen.^{[170],[235],[239]} Die Anwendungsmöglichkeit für die Authentizitätsbestimmung hingegen

ist komplex. In diesem Bereich überwiegen die Vorteile von *targeted*-Methoden für den spezifischen Nachweis von Sequenzunterschieden.

6.6 Entwicklung sortenspezifischer SCAR-Marker

SCAR-Marker (SCAR, engl. *sequence characterized polymorphic DNA*) wurden Anfang der 90er Jahre von I. Paran und R. W. Michelmore entwickelt.^[173] Es handelt sich hierbei um eine *targeted*-Methode bei der sequenzspezifische Primer basierend auf spezifischen Markern, die im Rahmen einer *non targeted*-Methode, wie der RAPD-PCR, identifiziert wurden, entwickelt werden, um diese anschließend gezielt amplifizieren und nachweisen zu können.^{[174],[175]} Der Grundgedanke dahinter ist, dass das Vorhandensein einer spezifischen Bande bei einer *fingerprinting*-Methode darauf hindeutet, dass es einen definierten DNA-Abschnitt, einen Marker, gibt, der nur bei den Sorten mit dieser Bande vorhanden ist und daher mittels eines entsprechenden spezifischen Primers amplifiziert werden kann. Solche Marker können schließlich als SCAR-Marker für den gezielten Nachweis herangezogen werden und bieten gegenüber der *non targeted*-Methode den Vorteil eines gezielten und spezifischen Nachweises. Zudem ist eine PCR mit einem spezifischen Primerpaar deutlich schneller durchzuführen und reproduzierbarer, da auch die Anfälligkeit gegenüber Änderungen der Reaktionsbedingungen geringer ist. Folgende Abbildung fasst das analytische Vorgehen zusammen.

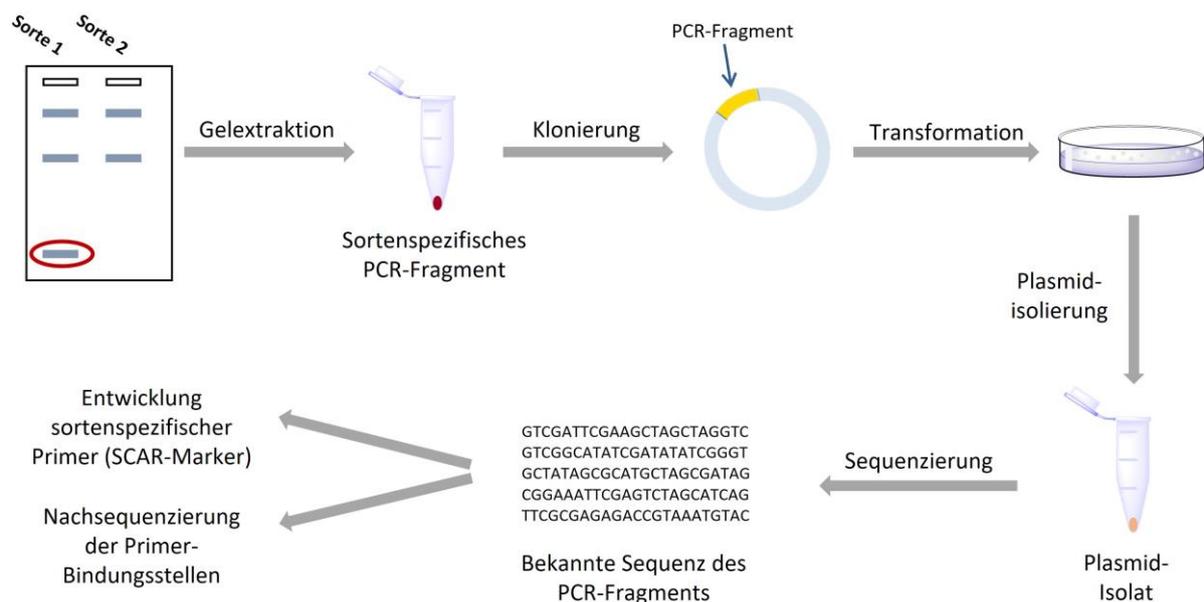


Abbildung 50: Analytisches Vorgehen für die Entwicklung eines *targeted*-Ansatzes auf Basis einer *non targeted*-Methode.

Als Grundlage für die Entwicklung von SCAR-Markern wurde die ebenfalls im Rahmen des FEI-Projekts 18751N entwickelte RAPD-PCR von Felbinger et al. (2020) herangezogen.^[234]

Zu Beginn wurde basierend auf den Ergebnissen von Felbinger et al. (2020) ein Primer ausgewählt, der spezifische Banden bei der RPAD-PCR aufweist. Die Wahl fiel dabei auf den Primer 17. Mit diesem resultiert eine Bande mit einer Länge von ca. 330 bp bei den Sorten Tombul und Ata Baba, die alle anderen untersuchten europäischen Haselnussproben nicht aufwiesen. Zunächst wurde geprüft, ob die Ergebnisse von Felbinger et al. (2020) reproduziert werden konnten. Das Ergebnis bzgl. der ausgewählten RAPD-Bande konnte bestätigt werden (siehe Abbildung 75). Für die Bestimmung der Sequenz der RAPD-Bande wurde die Sorte Tombul (Probe Nr. 109) eingesetzt. Es besteht hierbei die Möglichkeit der direkten oder indirekten Sequenzierung mit vorangehender Klonierung. Im vorliegenden Fall setzte sich die RAPD-Bande aus mehreren Fragmenten der gleichen Länge, jedoch unterschiedlichen Sequenzen zusammen, weshalb vor der Sequenzierung der Schritt der Klonierung notwendig war. Es wurden die Plasmidisolate von sechs Kolonien (K16, K17, K18, K20, K21, K22) mittels Sanger-Sequenzierung sequenziert und anschließend mit dem Online-Tool BLASTN[®] mit einer Nukleotidsequenzdatenbank abgeglichen. Dabei wurde eine hohe Übereinstimmung der Sequenzen mit den Accession-No. LR899423.1 und LR899424.1 identifiziert. Es handelt sich dabei um veröffentlichte Sequenzabschnitte des Kerngenoms der *Corylus avellana* (Sorte: Tombul). Da keine Übereinstimmungen mit veröffentlichten Gensequenzen identifiziert werden konnten, ist davon auszugehen, dass es sich um einen nicht kodierenden Bereich handelt. Die Kolonie K16 wurde für das folgende Primerdesign ausgewählt. Für das ursprüngliche Primerdesign, wie es bereits Paran und Michelmore (1993) entwickelten, werden die 10 bp-langen RAPD-Primer um 11-15 bp am 3'-Ende verlängert (Primerpaar SCAR_TomV_FW/RV).^{[173],[175]} Neben diesem Verfahren, wurde ein weiteres Primerpaar (SCAR_TomE_FW/RV) designt, das innerhalb des DNA-Fragments bindet und die ursprünglichen Primerbindungsstellen des RAPD-Primers nicht beinhaltet (siehe Abbildung 76). Die sequenzspezifischen Primer (siehe Tabelle 24) wurden anschließend auf vier Haselnussproben (Nr. 1, 8, 67 und 119) der Sorten T. G. Romana, Tonda di Giffoni, Segorbe und Tombul sowie vier Blattisolaten unterschiedlicher Haselnussarten (*C. avellana* (B21), *C. maxima* (B22, B23) und *C. colurna* (B24)) angewendet. Lediglich die Probe der Sorte Tombul sollte bei der durchgeführten PCR mit den entwickelten Primern ein Amplifikat aufweisen. Es zeigte sich jedoch bei beiden Primerpaaren eine klare und deutliche Bande der gewünschten Länge bei allen eingesetzten DNA-Isolaten (siehe Abbildung 77 und Abbildung 78).

Eine mögliche Ursache für das Auftreten eines Amplifikats bei allen Proben könnte der Sequenzunterschied selbst sein, der zu dem unterschiedlichen Bandenmuster bei der RPAD-PCR führte. Handelt es sich lediglich um einen sehr kleinen Unterschied, vielleicht sogar einen SNP, ist es möglich, dass dieser bei der RAPD-PCR aufgrund der sehr kurzen Primer dazu führte, dass kein PCR-Produkt in ausreichender Menge generiert und daher keine Bande detektiert wurde. Wird nun jedoch der Primer verlängert, erhöht sich der homologe Anteil, sodass der Primer trotz SNP binden und ein PCR-Produkt amplifiziert werden kann. Daher wurde zusätzlich der Ansatz verfolgt die Primerbindungsstellen auf Sequenzunterschiede hin zu analysieren. Hierfür ist es erforderlich Primer zu designen, die außerhalb des RAPD-Fragments binden und mit denen ein Amplifikat generiert werden kann, das bei der anschließenden Sequenzierung die ursprünglichen Primerbindungsstellen des RAPD-Primers beinhaltet. Für das Primerdesign muss jedoch die Sequenz um das ursprüngliche

Fragment herum bekannt sein. Daher wurden die Sequenzen mit den zu dem damaligen Zeitpunkt der Analyse veröffentlichten *Contigs* der ersten Genomsequenzierung einer *C. avellana* (Sorte: Jefferson) verglichen.^{[11],[240]} Das analytische Vorgehen für den beschriebenen Ansatz für die Sequenzierung der Primerbindungsstellen ist in Abbildung 51 zusammengefasst.



Abbildung 51: Schematische Darstellung des analytischen Vorgehens für die Sequenzierung eines RAPD-Fragments inklusive der Primerbindungsstellen.

Für die Sequenzen von vier Kolonien (K16, K17, K18 und K20) konnten Übereinstimmungen mit den veröffentlichten Sequenzen nachgewiesen werden. Es wurde jeweils ein Primerpaar (siehe Tabelle 24) designiert, das das Zielfragment umschließt. Mit dem Primerpaar Tom_K16_FW/RV war es trotz unterschiedlicher Optimierungen der Reaktionsbedingungen nicht möglich ein Amplifikat zu generieren. Eine mögliche Ursache hierfür könnte eine Poly-N-Region zwischen dem neuen Vorwärtsprimer und dem Zielfragment sein. Mit den Primerpaaren Tom_K17_FW/RV, Tom_K18_FW/RV und Tom_K20_FW/RV konnten zwar PCR-Produkte generiert werden, allerdings zeigte das Agarosegel eine Vielzahl unterschiedlicher Fragmente, die zudem nicht sauber voneinander getrennt werden konnten. Häufig wies auch das Zielfragment nicht die erwartete Länge auf, was ebenfalls mit den Poly-N-Regionen zu begründen ist. Lediglich für das Primerpaar Tom_K18_FW/RV wurde ein Fragment der erwarteten Länge amplifiziert. Es handelte sich aber auch um die einzige Sequenz ohne unbekannte Nukleotide. Auch der Versuch die jeweiligen Zielfragmente für eine Sequenzierung zu isolieren war nicht möglich. Für das Primerpaar Tom_K17_FW/RV wurde zudem eine *TouchDown*- und *Nested-PCR* angewandt. Diese führten jedoch ebenfalls zu keinem verbesserten Ergebnis. Zusammenfassend ist zu sagen, dass es sehr schwer war basierend auf den veröffentlichten *Contigs* spezifische Primer zu entwickeln, da die vorhandenen Sequenzabschnitte viele, zum Teil auch größere unbekannte Bereiche aufwiesen. Da die bisherigen Ergebnisse keine guten Resultate erzielten, wurden keine weiteren Alignments durchgeführt.

Bei der für die Analyse herangezogenen Probe der Sorte Tombul (Nr. 119) wurde bei der im weiteren Verlauf durchgeführten Analyse des Chloroplastengenoms festgestellt, dass es sich um eine Mischung unterschiedlicher Sorten handeln muss. Wie unter 6.3.1.1 geschildert, wies die Probe bei der Analyse der identifizierten Indels innerhalb des cpGenoms eine Doppelbande auf, was darauf hindeutet, dass es sich um keine sortenreine Probe handeln kann. Sie weist jedoch bei der RAPD-Analyse keine Abweichungen von den anderen beiden Tombul-Proben auf. Diese wurden jedoch bei der Analyse des Chloroplastengenoms nicht weiter analysiert, da neueres Probenmaterial zur Verfügung stand, bei dem die Vertrauenswürdigkeit auf Richtigkeit der Angaben zur biologischen Herkunft höher eingeschätzt wurde. Es zeigte sich jedoch insgesamt eine erhöhte Unsicherheit bzgl. der Authentizität in Bezug auf Handelsproben speziell aus den östlichen Anbaugebieten. Da die drei

Proben der Sorte Tombul jedoch vergleichbare Ergebnisse ohne Differenzen aufwiesen und auch eine Abgrenzung zu den anderen Sorten möglich war, scheint hier eine grundsätzliche Differenzierung bzgl. des genetischen Materials vorzuliegen. Zudem war auch in Bezug auf die gewählte Bande bei 350 bp (Primer 17) ein deutlicher Unterschied zu erkennen. Dennoch besteht die Möglichkeit, dass die Sequenz der ausgewählten Kolonie auf einem Sequenzabschnitt beruht, der in allen Proben vorhanden ist, bei den europäischen Sorten jedoch nicht in dem Umfang amplifiziert wurde, dass er als Bande detektiert wurde. Parallel wurden weitere RAPD-Banden auf Reproduzierbarkeit geprüft. Dabei ergab sich eine gewisse Unsicherheit, weshalb keine weiteren Klonierungen mit anderen Amplifikaten durchgeführt wurden.

Wie unterschiedliche Publikationen zeigen, eignen sich die SCAR-Marker sehr gut, wenn unterschiedliche botanische Gattungen miteinander verglichen und darauf basierend spezifische Primer entwickelt werden sollen. So wurden bspw. SCAR-Marker für den Nachweis von echtem Safran (*Crocus sativus* L.) und häufig eingesetzten Safran-Substituten oder für die Authentizitätskontrolle von Granatapfel entwickelt.^{[179],[184],[185]} Auch im Bereich der Kräuter und pflanzlichen Arzneimittel wurden Ergebnisse unterschiedlicher *non targeted*-Methoden in SCAR-Marker umgewandelt, v.a. aufgrund ihrer besseren Reproduzierbarkeit und der reduzierten Dauer und Komplexität der Auswertung.^{[174],[180]}

Es ist anzunehmen, dass die untersuchten Haselnussorten zu nah miteinander verwandt sind und daher keine variablen Sequenzabschnitte mittels des *non targeted*-Ansatzes identifiziert werden konnten, die einen so großen Unterschied darstellen, dass die Methodik der SCAR-Marker darauf anwendbar ist. Für kleine Differenzen wie bspw. SNPs müssen andere molekularbiologische Methoden herangezogen werden. Für diese ist jedoch der Sequenzunterschied eindeutig zu identifizieren und zu lokalisieren.

6.7 Zusammenfassung der Analyseergebnisse

Insgesamt zeigte die Analyse des Chloroplastengenoms, dass zwischen den Sorten ein sehr hoher Verwandtschaftsgrad vorliegt. Auf Basis der Sequenzierungs- und Resequenzierungsergebnisse konnten keine sortenspezifischen Sequenzunterschiede nachgewiesen werden. Es war jedoch möglich Gruppen mehrerer Haselnussorten voneinander zu trennen. In der Literatur sind die Ergebnisse in diesem Forschungsgebiet sehr unterschiedlich. Kane et al. (2012) und Hermann et al. (2014) konnten bspw. basierend auf Sequenzunterschieden innerhalb des cpGenoms zwischen unterschiedlichen Kakao-Sorten differenzieren.^{[43],[212]} Auch Khan et al. (2012, 2018) war die Identifizierung von Sequenzunterschieden innerhalb des cpGenoms zwischen verschiedenen Sorten der Dattelpalme (*Phoenix dactylifera* L.) möglich.^{[241],[242]} Andere Publikationen hingegen beschreiben auch den Nachweis vollkommen identischer Sequenzen wie bspw. Kahlau et al. (2006), die sich mit dem cpGenom der Tomate (*Solanum lycopersicum*) beschäftigen. So wiesen die Tomatensorten IPA-6 aus Brasilien und Ailsa Craig aus Europa keinerlei Sequenzunterschiede beim Vergleich der

cpGenomsequenzen auf.^[243] Auch Teske et al. (2020) konnten keinerlei Unterschiede beim Vergleich des cpGenoms zwölf unterschiedlicher Walnussorten identifizieren.^[244] Diese unterschiedlich stark ausgeprägten Differenzen können auf verschiedene Ursachen zurückgeführt werden. Zum einen spielt natürlich das Alter einer Sorte eine entscheidende Rolle, denn unter Berücksichtigung der Mutationsrate bedarf es eine gewisse Zeit bis sich Sequenzunterschiede ausgebildet haben.^[244] Es gibt jedoch auch grundlegende Unterschiede in der molekularen Entwicklung von Botaniken. Es wurden bereits Unterschiede zwischen einjährigen und mehrjährigen Pflanzen identifiziert, aber auch zwischen Bäumen und Sträuchern im Vergleich zu Kräutern. Diese beruhen u.a. auf Differenzen in Bezug auf die Größe der Pflanze, DNA-Reparaturmechanismen, Stoffwechselraten und Generationszeiten.^{[245]–[248]} Die Ausbildung genetischer Unterschiede ist damit ein äußerst komplexes Zusammenspiel unterschiedlichster Faktoren.^[245] Die Erfolgssausichten für die Identifizierung von Sequenzunterschieden sind nur schwer einzuschätzen und für jede Spezies individuell zu analysieren.

Grundsätzlich ist das cpGenom aufgrund seiner Länge von nur ca. 160.000 bp im Vergleich zu dem Kerngenom mit ca. 400 Mbp jedoch sehr gut als Ansatzpunkt geeignet, wenn keine Sequenzinformation vorhanden ist. Gleichzeitig weist es jedoch den Nachteil einer deutlich geringeren Mutationsrate im Vergleich zum Kerngenom auf. In der Literatur wird im Schnitt eine Mutationsrate von $1\text{--}3 \times 10^{-9}$ Basen pro Jahr angegeben.^[93] Diese Angaben sind sehr abhängig von der analysierten Spezies. So zeigten Untersuchungen von Willyard et al. (2006), dass die Gattung *Pinus* im Vergleich zu allgemeinen Angaben für die Gruppe der Angiospermen eine noch deutlich geringere Mutationsrate des cpGenoms von nur $0,2\text{--}0,4 \times 10^{-9}$ Basen pro Jahr aufweist.^[248] Für die Mutationsraten der drei unterschiedlichen Genome im Vergleich ergeben sich abhängig von der angewendeten Berechnungsmethode Angaben von $1 : 3 : 10$ bzw. $1 : 3 : 16$ (mtDNA : cpDNA : nDNA).^{[93],[249]} Das Mitochondriengenom weist damit die geringste Mutationsrate auf. Aufgrund dessen sind entsprechend lange Zeiträume notwendig, um Sequenzunterschiede ausbilden zu können, die letztendlich für die Sortendifferenzierung herangezogen werden können. Denn viele Sorten entstehen aus Kreuzungen bewährter Sorten, sodass dadurch bereits enge Verwandtschaften vorliegen. So zeigten beispielsweise Verwandtschaftsanalysen von Gökirmak et al. (2009) auf Basis von Mikrosatelliten, dass es sich bei den Haselnussorten Butler, Ennis und Corabel um Setzlinge der Sorte Barcelona handelt. Die ersten beiden Sorten sind Barcelona x Daviana Hybride und die Sorte Corabel ist ein Barcelona x Cosford Hybrid. Bei der Sorte Barcelona handelt es sich um die in den USA führende Sorte und bei der Sorte Daviana sowie der eng verwandten Sorte Cosford um die führenden Bestäuber.^[8]

Trotz der geringen Mutationsraten konnten Sequenzunterschiede innerhalb der LSC-Region des cpGenoms identifiziert werden. Auf Basis zwei 118 bzw. 115 bp langen Insertionen in den Bereichen B und C (siehe 6.3.1) sowie der identifizierten SNPs in Bereich D, speziell SNP 6 (siehe 6.3.2), konnten 19 untersuchte Haselnussorten aus sieben Ländern in drei Gruppen unterteilt werden (siehe Abbildung 52). Die Unterteilung zeigt dabei eine gewisse geografische Clusterung. Gruppe I bilden hierbei die europäischen Proben der Sorten Barcelona, Butler, Corabel, Ennis, Hallsche Riesen, Mortarella, Negret, Pautet, Segorbe, Tonda di Giffoni, T. G. Romana und T. G. Trilobata zusammen mit der türkischen Sorte Çakildak. Gruppe II besteht aus aserbaidchanischen Haselnüssen wie die

Sorte Ata Baba und den türkischen Sorten Sariyağlı und Yomra. Gruppe III bilden Haselnüsse georgischer Herkunft. Drei der untersuchten türkischen Sorten (Karayağlı, Palaz und Tombul) konnten anhand der vorliegenden Proben und der ambivalenten Ergebnisse (siehe 6.3.1) keiner Gruppe eindeutig zugeordnet werden.

Die Indels können mit einer schnell durchzuführenden PCR und anschließender Detektion mittels AGE analysiert werden. Für den Nachweis georgischer Haselnüsse ist im Anschluss noch ein enzymatischer Verdau unter Einsatz des Restriktionsenzym *MfeI* notwendig. Kombiniert mit der vorangehenden DNA-Isolierung mittels der optimierten *Rapid Extraktion* ist eine qualitative Aussage bzgl. einer möglichen Beimischung innerhalb weniger Stunden möglich. Auch über die entwickelte HRMA ist ein Nachweis der genannten Sequenzunterschiede und eine Unterteilung der drei genannten Gruppen mit nur einem PCR-Lauf mit anschließender hochauflösender Schmelzpunktanalyse möglich. Nachteil dieser Methode ist jedoch, die Notwendigkeit einer etwas spezielleren Laborausstattung und der im vorliegenden Fall genutzten Software *Precision Melt Analysis* der Firma Bio-Rad Laboratories Inc. (Hercules, USA) für die Erstellung eines *Difference plots*. Erst über diese Auswertung und Visualisierung der Ergebnisse können alle drei Gruppen eindeutig voneinander unterschieden werden. Des Weiteren ist authentisches Referenzmaterial, welches bei jeder Analyse parallel mit untersucht werden muss, für die Erstellung des *Difference plots* notwendig. Die anderen beiden vorgestellten qualitativen Analysemethoden benötigen neben einem Thermocycler kein besonderes Equipment oder Software.

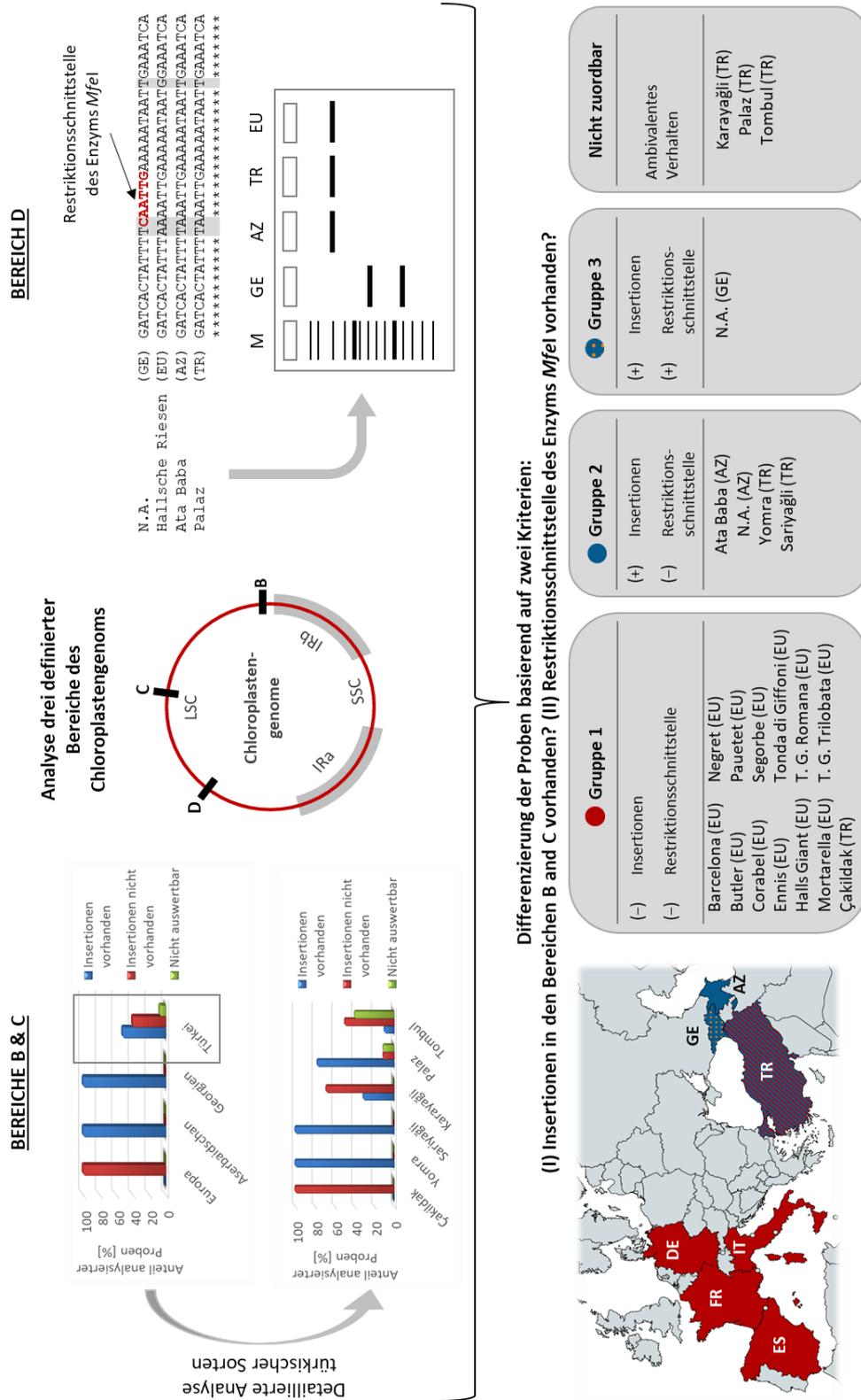


Abbildung 52: Überblick der Analysenergebnisse von drei identifizierten polymorphen Bereichen (Bereiche B-D) des cpGenoms. In den Bereichen B und C wurden zwei Indels identifiziert und in Bereich D ein SNP, welches spezifisch für georgische Haselnüsse ist und in einer Erkennungssequenz für das Restriktionsenzym *MfeI* resultiert. Basierend auf den genannten Sequenzunterschieden ist die dargestellte Gruppierung möglich. (LSC = *large single copy*, SSC = *small single copy*, IR = *inverted repeat*, AZ Aserbaidschan, DE = Deutschland, ES = Spanien, FR = Frankreich, GE = Georgien, IT = Italien, TR = Türkei, M = DNA-Leiter, N.A. nicht bekannt).

In der Literatur gibt es bereits verschiedene Publikationen, die sich ebenfalls mit der Methodenentwicklung für den Herkunftsnachweis von Haselnüssen beschäftigten. Diese beruhen dabei auf der Analyse des Metaboloms mittels NMR oder LC-MS.^{[250],[251]} Hierbei konnten auch Anbauggebiete innerhalb Europas (Deutschland, Frankreich und Italien) voneinander unterschieden werden.^{[250],[251]} Das Metabolom als Grundlage dieser Methoden weist jedoch den Nachteil auf, dass es durch anthropogene Faktoren beeinflusst wird. Dadurch ist die kontinuierliche Pflege einer Datenbank durch die jährliche Messung authentischer Referenzproben notwendig, um Differenzen zwischen verschiedenen Erntejahren aufgrund von bspw. veränderten Umwelteinflüssen zu berücksichtigen. Des Weiteren ist aufgrund der statistischen Auswertung der Ergebnisse auf einen ausreichend großen Probenpool an Referenzproben zu achten. Die im Rahmen dieser Arbeit entwickelten Methoden basierend auf dem Genom bieten hingegen den Vorteil, dass sie auf definierten, identifizierten Sequenzunterschieden beruhen. Ein weiterer Nachteil der Metabolombasierten Ansätze ist, dass nur Reinproben zugeordnet werden können. Verfälschungen durch Beimischungen minderwertiger oder günstigerer Ware können nicht bestimmt werden. Einen ersten Ansatz zur Identifizierung von Mischungen basierend auf der NMR-spektroskopischen Analyse des Metaboloms von Haselnüssen entwickelten Bachmann et al. (2019).^[252] Die kleinste eindeutig zu identifizierende Beimischung lag hierbei jedoch bei 50%.^[252] Mit den im Rahmen der vorliegenden Arbeit entwickelten Methoden hingegen können auch geringere Beimischungen identifiziert und semiquantitativ abgeschätzt werden.

Alle identifizierten Sequenzunterschiede sowie die darauf aufbauenden Methoden bieten hierfür die Möglichkeit. Für den quantitativen Ansatz basierend auf den Indels sowie die PCR-RFLP ist eine CGE als Detektionsmethode notwendig. Es konnte jedoch auch gezeigt werden, dass auch die AGE mit anschließender Graustufenanalyse eine kostengünstige Alternative darstellt, die dennoch verlässliche Ergebnisse liefert und lediglich der LOD und LOQ etwas höher liegen. Die erhöhten Schwankungen im Bereich niedriger Konzentrationen wiesen beide Detektionsmethoden auf. Für die quantitative Abschätzung mittels HRMA ist im Vergleich zur qualitativen Analyse kein weiteres Equipment notwendig, da die Auswertung ebenfalls über den zuvor bereits für die Identifizierung benötigten *Difference plot* erfolgt. Der ermittelte LOD und LOQ lagen jedoch etwas höher im Vergleich zu den anderen Ansätzen. Alle Methoden beruhen dabei auf einer externen Kalibrierung anhand einer Kalibriergerade und linearer Regression.

Nachteil der entwickelten, als auch vorgestellten, publizierten Methoden ist jedoch, dass lediglich Anbauggebiete voneinander unterschieden werden können, aber kein sortenspezifischer Nachweis möglich ist. Oddone et al. (2009) entwickelten eine Methode zur Differenzierung italienischer und türkischer Haselnussorten basierend auf der Analyse der Lanthanoide mittels ICP-MS.^[253] Mit Hilfe einer statistischen Auswertung zeigte sich auch eine Gruppierung innerhalb der italienischen Sorten und eine mögliche Differenzierung der Sorte T. G. Trilobata aus dem Piemont von zwei anderen italienischen Sorten.^[253] Auch Sciubba et al. (2014) gelang eine Differenzierung drei italienischer Haselnussorten auf Basis des Metaboloms nach Messung mittels NMR.^[254] Alasalvar et al. (2004) untersuchten hingegen flüchtige Komponenten mittels GC-MS und konnten ebenfalls basierend auf einer statistischen Auswertung der qualitativen und quantitativen Analyse dieser Substanzen fünf

türkische Haselnussorten unterscheiden.^[255] Es zeigt sich also, dass es bereits unterschiedlichste Ansätze auch für die Sortendifferenzierung gibt. Jedoch beschäftigen sich die Publikationen meist nur mit wenigen ausgewählten Sorten, häufig aus nur einem Anbaugebiet. Es fehlt jedoch an Publikationen, die solche sortenspezifischen Ansätze auf eine größere Anzahl an Sorten, v.a. auch aus unterschiedlichen Anbaugebieten anwendet, um hier eine Aussage treffen zu können, in welcher Methode das größte Potenzial liegt, um Sorten spezifisch und verlässlich voneinander zu unterscheiden.

Auch im Bereich der molekularbiologischen Methoden gibt es Ansätze für die Sortendifferenzierung. So wendeten Felbinger et al. (2020) den *non-targeted*-Ansatz der RAPD-PCR auf 13 unterschiedliche Haselnussorten an und konnten basierend auf den Ergebnissen einen Sortenbaum zur Unterscheidung dieser entwickeln.^[234] Auch Öztürk et al. (2018) war es möglich anhand von Mikrosatelliten einen Entscheidungsbaum zur Differenzierung von 19 türkischen Haselnussorten zu definieren.^[236] Ein zu Felbinger et al. (2020) vergleichbares Ergebnis zeigte die im Rahmen der vorliegenden Arbeit durchgeführte Analyse der SCoT-Polymorphismen (siehe 6.5). Von der Entwicklung eines Sortenbaums wurde aufgrund der hohen Variabilität der Bandenmuster abgesehen. Mittels einer Hauptkomponentenanalyse konnte jedoch eine deutliche Clusterung der Proben einer Sorte zueinander identifiziert werden. Grundsätzlich sind solche *non targeted*-Ansätze sehr sorgfältig zu bewerten, da bei diesen auch das Kerngenom erfasst wird. Im Fall der Haselnuss beinhaltet der Kern sowohl die Sequenzinformation der Mutterpflanze auch die des Bestäubers, bei dem es sich aufgrund der Selbstinkompatibilität nicht um die Sorte der Mutterpflanze handeln kann. Da Haselnüsse zusätzlich per Wind bestäubt werden, ist von einer hohen genetischen Variabilität auszugehen, die bereits in verschiedenen Publikationen bestätigt wurde.^{[256]–[259]} Das deckt sich mit der Tatsache, dass auf Basis des cpGenoms keine sortenspezifischen Sequenzunterschiede identifiziert werden konnten, mittels *non targeted*-Methoden, die das Kerngenom mit einschließen, eine Differenzierung aber möglich ist. Trotz der zu erwartenden höheren Variabilität war es nicht möglich innerhalb des Kerngenoms, speziell der ITS-Bereiche und der WRKY-Gene sequenzspezifische Unterschiede zu identifizieren.

In den meisten Fällen wird bei Haselnüssen, die in den Handel gebracht werden lediglich das Herkunftsland angegeben. Daher stellen die entwickelten Methoden einen wichtigen Schritt in Richtung einer analytischen Authentizitätsprüfung sowohl für die verarbeitende Industrie, als auch die amtliche Lebensmittelüberwachung dar. Mit verhältnismäßig einfachen und schnellen Methoden können relevante Beimischungen, wie die Verfälschung durch georgische oder aserbaidjanische Haselnüsse, die deutlich günstiger gehandelt werden, nachgewiesen und im Vergleich zu der reinen Authentizitätsprüfung anhand von Lieferpapieren, belastbarere Aussagen getroffen werden. Es können dabei nicht alle Varianten der Streckung, aber dennoch vor dem Hintergrund des finanziell motivierten *Food Frauds* sehr relevante Verfälschungen nachgewiesen werden. Ein mögliches analytisches Vorgehen ist anhand eines Fallbeispiels in Abbildung 53 dargestellt. Ausgangspunkt stellen beispielhaft Haselnüsse italienischen Ursprungs dar. Nach einer DNA-Isolierung mittels der *Rapid Extraction*-Methode erfolgt parallel eine PCR auf den in dieser Arbeit definierten Bereichen B und D des cpGenoms. Mit dem generierten Amplifikat des Bereichs D wird anschließend noch ein

enzymatischer Verdau mit dem Restriktionsenzym *MfeI* durchgeführt. Zuletzt werden die PCR-Fragmente von Bereich B und die des enzymatischen Verdau mittels AGE getrennt. Anhand der Bandenmuster kann anschließend direkt eine Aussage bzgl. der Authentizität getroffen werden. Zeigt bspw. das Agarosegel von Bereich B eine Doppelbande, würde bereits feststehen, dass eine Mischung unterschiedlicher Sorten vorliegt. Damit wäre nachgewiesen, dass die Deklaration falsch ist. Könnten zusätzlich bei dem Gel von Abschnitt D drei Banden detektiert werden, kann zusätzlich die Aussage getroffen werden, dass die untersuchte Probe mit Haselnüssen georgischen Ursprungs verfälscht wurde. Damit lässt sich die Authentizität innerhalb von 3-5 h zu einem deutlich höheren Grad verifizieren, als es nur anhand von Lieferantendokumenten möglich wäre.

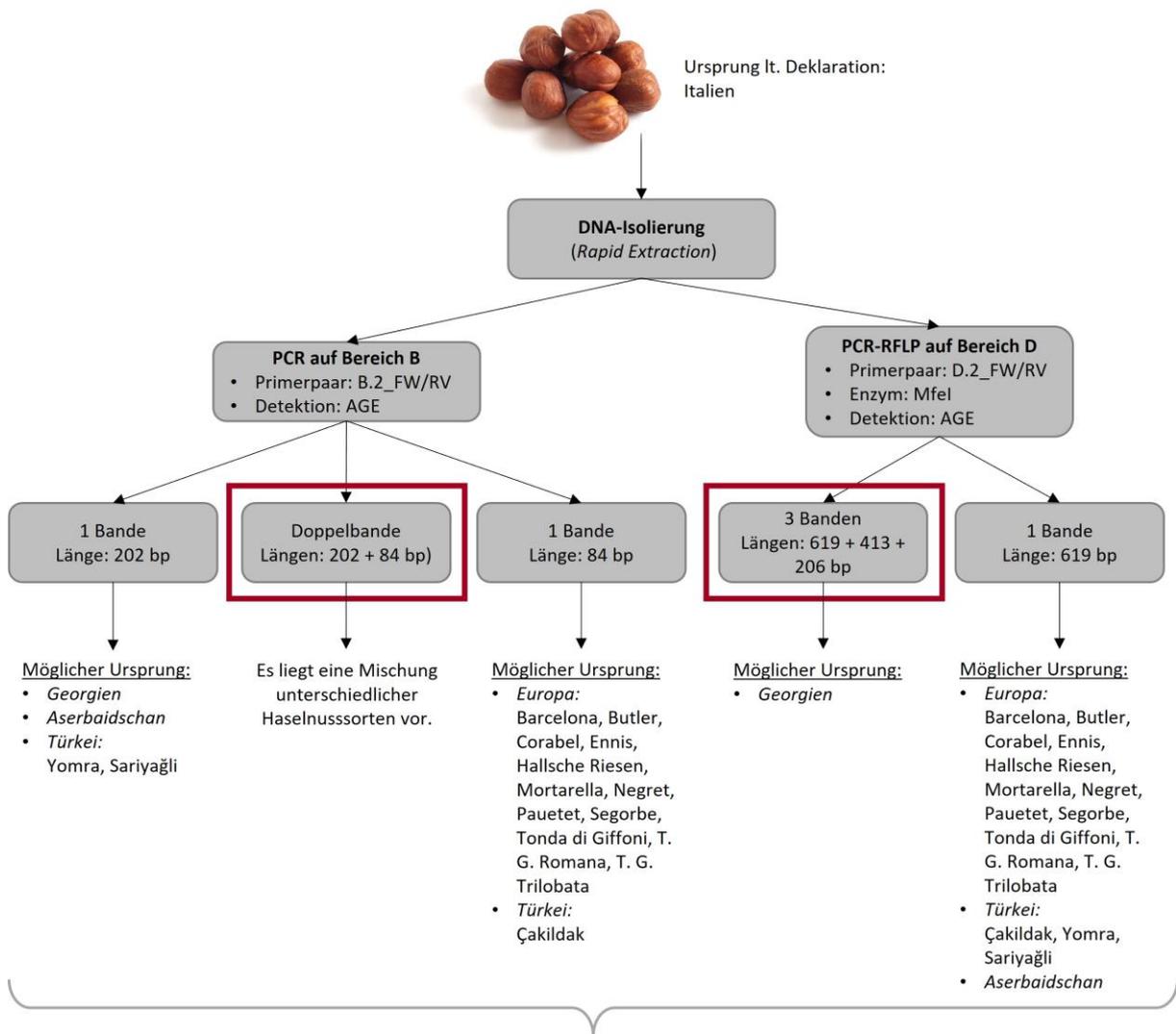


Abbildung 53: Fallbeispiel für eine mögliche Anwendung der entwickelten *targeted*-Methoden für die Authentizitätsüberprüfung von Haselnüssen.

7 Material und Methoden

7.1 Probenmaterial

Im Rahmen dieser Arbeit wurden sowohl unverarbeitete als auch prozessierte Haselnussproben sowie Haselnussblattproben verwendet. Bei der Wahl des Probenmaterials wurde auf eine hohe Authentizität geachtet. Im Folgenden wird detailliert auf die verschiedenen Probengruppen eingegangen.

7.1.1 Haselnusskerne (Rohware)

Der Großteil der verwendeten Haselnusskerne wurde unverarbeitet, das heißt nach der Ernte lediglich getrocknet, mit oder ohne Schale von verschiedenen Industriepartnern bezogen. Eine Übersicht der Proben inkl. der Lieferanten ist der Tabelle 27 im Anhang unter 9.6 zu entnehmen. Ein Teil der Proben wurde unter trockenen Bedingungen bis zur weiteren Aufarbeitung gelagert. Der andere Teil hingegen wurde direkt wie folgt zur Analyse aufgearbeitet:

1. Nüsse mit Schale aufbrechen und die Schale manuell entfernen
2. Nüsse in Gefrierbeutel verpacken und in flüssigem Stickstoff frieren
3. Gefrorene Nüsse im Mahlbehälter der Messermühle GM300 mit Trockeneis (1:1) vermengen und in zwei Intervallen mahlen:
 - 1x 20 s bei 1000 rpm
 - 4x 30 s bei 4000 rpm

Anschließend wurde das Homogenisat bei -20 °C bis zur Verwendung gelagert.

7.1.2 Prozessierte Haselnussproben

Neben den unverarbeiteten Haselnüssen wurden auch prozessierte Haselnusskerne für die Analytik herangezogen. Der wichtigste Verarbeitungsschritt liegt hierbei bei der Röstung der Haselnüsse. Neben kommerziell erworbenen gerösteten Haselnüssen (Nr. V2, V3) wurden zwei Proben (Nr. V4 und V5) selbst geröstet. Dies erfolgte mit geknackten ganzen Haselnüssen in einem Trockenschrank bei 180 °C für 12 min wobei die Kerne alle 2 min geschwenkt wurden. Anschließend wurde die Samenschale weitestgehend entfernt. Des Weiteren wurden Kekse mit gemahlene Haselnüssen untersucht, welche nach folgendem Rezept hergestellt wurden.

- | | |
|------------------------------|--|
| • 300 g Butter | Alle Zutaten zu einem festen Teig verkneten. |
| • 150 g Zucker | Anschließend Kugeln formen und auf ein mit |
| • 300 g Mehl | Backpapier ausgelegtes Backblech legen. |
| • 300 g gemahlene Haselnüsse | Bei 175 °C für 10 min backen. |

Die prozessierten Haselnussprodukte wurden nach demselben Vorgehen wie bei den unverarbeiteten Produkten wie unter 7.1.1 beschrieben für die weitere Analyse vorbereitet.

7.1.3 Haselnussblätter

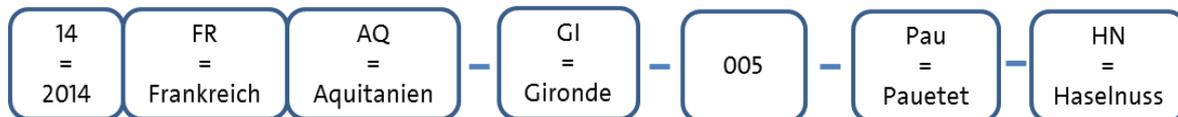
Die Haselnussblätter wurden zum einen von Mitarbeitern der *Hamburg School of Food Science* im Loki-Schmidt-Garten (Botanischer Garten der Universität Hamburg) in Hamburg als auch der Bayerischen Landesanstalt für Wein- und Gartenbau (Fürth) gepflückt sowie von Industriepartnern zur Verfügung gestellt. Eine Übersicht aller verwendeten Blattproben ist der Tabelle 29 im Anhang unter 9.6 zu entnehmen. Die Blätter wurden nach Erhalt mit destilliertem Wasser von Verunreinigungen wie Sand und Erde befreit und anschließend bei -20 °C bis zur weiteren Verwendung gelagert.

7.1.4 Kennzeichnung des Probenmaterials

Jede Probe wurde mit einem spezifischen Barcode versehen. Dieser gibt Auskunft über die wichtigsten Hintergrundinformationen zu der Probe. Er beinhaltet Angaben bzgl. des Erntejahres, der Herkunft, der Sorte sowie der Probenart (Nuss oder Blatt). Mit Hilfe dieser Kennzeichnung kann jede Probe eindeutig identifiziert werden. Im Folgenden ist der Aufbau der Barcodes schematisch dargestellt und anhand eines Beispiels nachzuvollziehen.



Beispiel:



Alle für die Barcodes verwendeten Abkürzungen sind unter 9.7 in Tabelle 30 und Tabelle 31 erläutert.

7.2 Gesamt-DNA-Isolierung

Die Isolierung der DNA wurde jeweils der entsprechenden Matrix angepasst. Die DNA-Isolierung aus Haselnussblättern erfolgte mit Hilfe des peqGOLD Plant DNA Mini Kits der Firma Peqlab (VWR). Für die Gewinnung der DNA aus Haselnusskernen und verarbeiteten Produkten wurde zum einen eine CTAB-Methode nach Brüning et al. (2011) und zum anderen eine neu entwickelte zeit- und kosteneffizientere Methode in Anlehnung an Cenis (1992) verwendet. Die DNA-Isolate wurden jeweils wie unter 7.3.1 beschrieben quantifiziert und gemäß 7.3.2 qualitativ bewertet.

7.2.1 Isolierung der Gesamt-DNA aus Haselnussblättern

Die Gesamt-DNA aus Haselnussblättern wurde mit dem peqGOLD Plant DNA Mini Kit der Firma Peqlab (VWR) isoliert. Hierfür wurden die Blätter zunächst aufgetaut und erneut gründlich gewaschen, um Mikroorganismen vollständig zu entfernen. Folgende Lösungen wurden dafür in genannter Reihenfolge angewendet: (1) verdünnte Tween®-Lösung, (2) 0,1 M Natronlauge und (3) 70%-iger Ethanol. Anschließend wurden die Blätter unter Zugabe von flüssigem Stickstoff gemörsert, um ein feines Pulver zu erhalten. Die Durchführung der DNA-Isolierung erfolgte schließlich gemäß der Herstellerangabe (Arbeitsanleitung, Version V0815) nach Abschnitt B für frische und gefrorene Proben. Die Einwaage betrug jeweils 100 mg und die Elution wurde mit 2x 25 µL Reinstwasser durchgeführt.

7.2.2 Isolierung der Gesamt-DNA aus Haselnusskernen

7.2.2.1 CTAB-Methode

Die Isolierung der Gesamt-DNA mittels CTAB-Methode erfolgte in Anlehnung an die Methode *Precipitation with Isopropyl Alcohol and Subsequent Silica Adsorption (Methode Ip + S)* nach Brüning et al. (2011), die im Rahmen des IGF-Projektes AiF 15304 N etabliert wurde.^[42] Die Zentrifugationsschritte wurden alle bei 12.000 rpm durchgeführt.

Zunächst wurden 150 mg gemahlene Haselnusskerne/-produkte in ein 2 mL Reaktionsgefäß eingewogen, mit 1 mL Extraktionspuffer (55 mM CTAB, 1,4 M NaCl, 0,1 M Tris-HCl, 20 mM EDTA, HCl q.s. pH 8,0) versetzt und bei 65 °C für 60 min in einem Thermoshaker inkubiert. Anschließend wurden 700 µL Chloroform zugegeben, sorgfältig gemischt und 10 min zentrifugiert. 800 µL des Überstandes wurden in ein neues Reaktionsgefäß überführt und erneut mit 700 µL Chloroform versetzt. Nach erneuter Zentrifugation wurden 700 µL des Überstandes in einem neuen Reaktionsgefäß mit 700 µL eisgekühltem Isopropanol versetzt und für 1 h bei 4 °C inkubiert. Nach 10-minütiger Zentrifugation wurde der Überstand abdekantiert, der Rückstand in 650 µL Bindepuffer (5,5 M Guanidiniumhydrochlorid, 20 mM Tris-HCl, HCl q.s. pH 6,6) aufgenommen und vollständig auf eine EconoSpin® Säule überführt. Im Anschluss erfolgte ein Zentrifugationsschritt von 2 min. Nun wurde die Säule zunächst mit 600 µL Waschpuffer (20 mM Tris-HCl, 1 mM EDTA, 50 mM NaCl in 50 %-igem EtOH, HCl q.s. pH 7,4) und anschließend mit 500 µL 70 %-igem Ethanol (-20 °C) gewaschen. Hierbei wurde jeweils für 2 min zentrifugiert. Im Anschluss wurde die Säule durch einen erneuten Zentrifugationsschritt von 2 min getrocknet. Anschließend wurde die Säule in ein neues Reaktionsgefäß überführt und die DNA in zwei Schritten mittels Reinstwasser eluiert. Hierfür wurde jeweils 10 µL Wasser auf die Säule pipettiert, diese 10 min bei RT inkubiert und 1 min zentrifugiert. Die DNA-Isolate wurden bei -20 °C bis zur weiteren Verwendung gelagert.

7.2.2.2 Rapid Extraction

Bei der DNA-Isolierung in Anlehnung an Cenis (1992) handelt es sich um eine im Vergleich zu der zuvor beschriebenen Methode nach Brüning et al. (2011) deutlich kosten- und zeiteffizientere Methode.^{[42],[260]} Die Zentrifugationsschritte erfolgten jeweils bei 16.000 g.

Zunächst wurden 30 mg gemahlene oder mit einem Skalpell zerkleinerte Haselnüsse mit 500 µL auf 65 °C vortemperiertem Extraktionspuffer (200 mM Tris, 250 mM NaCl, 25 mM EDTA, 0,5 % SDS, HCl q.s. pH 8,5) versetzt und die Suspension mit einem Pistill homogenisiert. Nach Zugabe von 300 µL Natriumacetatlösung (3 M, pH 5,2) erfolgte eine 10-minütige Inkubation bei -20 °C. Anschließend wurde 5 min zentrifugiert, der Überstand mit 500 µL Isopropanol (-20 °C) versetzt und 5 min bei RT inkubiert. Nach einem Zentrifugationsschritt von 15 min wurde das Pellet mit 500 µL 70 %igem Ethanol (-20 °C) gewaschen. Nach erneuter Zentrifugation (5 min) wurde das DNA-Pellet bei 37 °C für 10 min in einem Heizblock getrocknet und zuletzt in 50 µL Reinstwasser aufgenommen. Die Lagerung der DNA-Isolate bis zur weiteren Verwendung erfolgte bei -20 °C.

7.3 Qualitätsbestimmung von DNA-Isolaten

7.3.1 Fluorimetrische Konzentrationsbestimmung

Für die fluorimetrische Konzentrationsbestimmung wurde das Quantus™ Fluorometer der Firma Promega (Mannheim, Deutschland) in Verbindung mit dem QuantiFluor® dsDNA System verwendet. Das Kit beinhaltet einen 20-fach konzentrierten TE-Puffer, einen Interkalationsfarbstoff (QuantiFluor® dsDNA Dye) sowie einen dsDNA-Standard (Lambda DNA Standard, c = 100 µg/mL). Zu Beginn muss der TE-Puffer mit Reinstwasser 1:20 verdünnt werden. Anschließend wird durch eine 1:400-Verdünnung des *dsDNA Dye* mit TE-Puffer (1x) die *dsDNA Dye Working Solution* hergestellt. Die Kalibrierung erfolgt nach Herstellerangabe über eine Einpunktkalibrierung mit Hilfe des dsDNA-Standards. Die Kalibrierung wurde in regelmäßigen Abständen durch Messung des Standards überprüft. Bei jeder Probenmessung wurde zudem ein Blindwert mitgeführt, um alle Chemikalien auf eine mögliche Kontamination mit DNA zu überprüfen. Ein Pipettierschema für Standard, Blindwert und Probe ist in Tabelle 15 dargestellt.

Tabelle 15: Pipettierschema für die fluorimetrische Konzentrationsbestimmung.

	Standard	Blindwert	Probe
<i>dsDNA Dye working solution</i> [µL]	200	200	200
dsDNA-Standard [µL]	2	---	---
1x TE-Puffer [µL]	---	2	---
Probe [µL]	---	---	2

Vor der Messung ist zu prüfen, dass bei dem Messgerät das richtige Target, in diesem Fall dsDNA, sowie das eingesetzte Probenvolumen und die gewünschte Einheit [ng/µL] des Ergebnisses

ausgewählt sind. Die Reaktionsansätze werden anschließend in dem Gerät platziert und der Deckel geschlossen. Die Messung startet daraufhin automatisch. Die DNA-Konzentration wird in der angegebenen Einheit angezeigt.

7.3.2 Photometrische Reinheitsbestimmung

Die photometrische Reinheitsbestimmung der DNA-Isolate erfolgte mittels Mikrovolumenmessung mit dem DeNovix DS-11 + Spektrophotometer. Das Gerät verfügt über einen unteren fixierten Messpunkt, auf den die zu untersuchende Probe pipettiert wird, und einen oberen flexiblen Messpunkt, welcher auf einem Hebel angebracht ist. Durch Schließen des Hebels entsteht eine Probensäule, welche anschließend photometrisch vermessen wird. Gemessen wird die Absorption bei 230 nm, 260 nm und 280 nm. Zu Beginn erfolgt jeweils ein Nullabgleich mit 1 μ L Reinstwasser. Anschließend wird je 1 μ L Probe aufgetragen. Zwischen den Messungen werden die beiden Messpunkte mit einem Tuch gereinigt. Als Ergebnis werden die DNA-Konzentration in ng/ μ L sowie die Quotienten A_{260}/A_{280} und A_{260}/A_{230} ausgegeben.

7.4 Qualitative Polymerasekettenreaktion (PCR)

Mit Hilfe der PCR kann ein definierter Sequenzabschnitt gezielt vervielfältigt werden. Die klassische qualitative PCR erfolgt mit einem spezifischen Primerpaar, das an definierten DNA-Abschnitten bindet. Es resultiert ein DNA-Fragment bekannter Länge. Die Detektion erfolgt mittels AGE oder CGE. Mit Hilfe entsprechender Marker können die Amplifikate auf ihre korrekte Länge überprüft bzw. anhand derer zugeordnet werden.

Die Pipettierarbeiten werden unter einer DNA-Workstation durchgeführt. Diese wird zuvor mit 70 %igem Ethanol gereinigt und anschließend 15 min mit UV-Licht bestrahlt, um Kontaminationen zu vermeiden. Um eine lange Haltbarkeit der Chemikalien zu gewährleisten und ein vorzeitiges Starten der Reaktion zu verhindern, erfolgen die Pipettierarbeiten unter Kühlung. Ein Standard-PCR-Ansatz enthält folgende Bestandteile: Wasser, polymerase-spezifischer Puffer, dNTP-Mix, Primer bzw. Primerpaar, DNA Polymerase und DNA-Isolat. Optional kann auch BSA zugegeben werden. Die gewünschte Konzentration der DNA-Isolate wird ggf. durch Verdünnen mit Reinstwasser eingestellt. Als Negativkontrolle wird parallel zu den Proben ein Blindwert, der Reinstwasser anstatt des Templates enthält, mitgeführt. Die analysenspezifischen Zusammensetzungen der Reaktionsansätze sind den jeweiligen Abschnitten im Folgenden zu entnehmen. Die PCR wird in einem programmierbaren *Thermocycler* durchgeführt. Tabelle 16 zeigt beispielhaft ein Temperaturprogramm einer PCR mit einer *Taq* DNA Polymerase. Bei Verwendung anderer Polymerasen, können die benötigten Programme abweichen. Alle verwendeten Temperaturprogramme sind im Abschnitt 9.8 aufgelistet.

Tabelle 16: Allgemeines Temperaturprogramm einer PCR mit einer *Taq* DNA Polymerase.

Prozess	Temperatur / Zyklus	Zeit
Deckelheizung	95 °C	Dauerhaft
Denaturierung	95 °C	5 min
Beginn Zyklus (x35)		
Denaturierung	95 °C	20-40 s*
Annealing	variabel	20-40 s*
Elongation	72 °C	20-40 s*
Ende Zyklus		
Finale Elongation	72 °C	5 min
Kühlung	4 °C	Bis Entnahme

* variiert je nach Amplifikatlänge

Eine Übersicht aller verwendeten Primer inklusive der jeweiligen Annealingtemperaturen ist unter 9.3 in Tabelle 24 dargestellt. Die Dauer der einzelnen Schritte eines PCR-Zyklus ist abhängig von der gewünschten Amplifikatlänge. Bei längeren Amplifikaten wird die Zeit erhöht und bei kurzen PCR-Produkten verringert.

Im Anschluss an die PCR erfolgt die Detektion und Visualisierung entweder mittels AGE und anschließender Färbung mit *GelRed* (siehe 7.5.1) oder mittels CGE (siehe 7.5.2).

7.5 Detektion von PCR-Produkten mittels Elektrophorese

7.5.1 Agarosegelelektrophorese (AGE)

Für die Trennung von DNA-Molekülen mittels AGE wurden Agarosegele in einem Konzentrationsbereich von 1-3 % verwendet. Hierfür wurden je nach gewünschter Konzentration 0,5-1,5 g Agarose in 50 mL Tris-Acetat-EDTA-Puffer (TAE-Puffer: 40 mM Tris, 2 mM EDTA, Eisessig *q.s.* pH 8,2) suspendiert und anschließend etwa 2 min bei 460 W in der Mikrowelle erhitzt bis die Agarose vollständig gelöst war. Nach kurzem Abkühlen wurde die Agarose in Gelschlitten gegossen und Kämme für die Probenaschen eingesetzt. Nachdem das Gel vollständig ausgehärtet war, wurde der gesamte Gelschlitten in eine Elektrophoresekammer überführt, mit TAE-Puffer überschichtet und anschließend die Kämme entfernt. Die Proben wurden für die AGE vorbereitet, indem 1-2 µL *Loading Dye* (49,9975 % Glycerin, 49,9975 % TAE-Puffer, 0,005 % Xylencyanol) zugegeben wurden. Dies dient zum einen zur besseren Visualisierung der Proben aufgrund des enthaltenen Xylencyanol und zum anderen zur Beschwerung der Proben durch das enthaltene Glycerin. Dadurch kann die Probe nicht aus der Geltasche in den Puffer diffundieren. Neben den Proben wurde auch ein Molekulargewichtsmarker (peqGOLD DNA Ladder Mix, 100-100.000 bp) in eine Probenasche pipettiert, damit schließlich die Fragmentlängen der DNA-Moleküle bestimmt werden konnten. Die Auftrennung erfolgte für 20 min bei einer Spannung von $U = 170$ V und einer Leistung von $P = 150$ W. Im Anschluss daran wurde das Gel in ein Färbebad, das den Interkalationsfarbstoff *GelRed*® (1:3333

verdünnt in VE-Wasser) beinhaltet, überführt. Die Färbedauer beträgt ca. 5-20 min, abhängig von der Anzahl an Gelen, die bereits gefärbt wurden. Nachdem das Gel kurz in VE-Wasser gewaschen wurde, erfolgt die Visualisierung der Banden unter einer Geldokumentationseinheit bei einer Anregungswellenlänge von 312 nm und einer Emissionswellenlänge von 605 nm. Mit Hilfe der Steuerungssoftware Argus X1 (Version 5.1.11) kann das Gel für die weitere Auswertung fotografisch dokumentiert werden. Eine quantitative Auswertung der entsprechenden Banden über eine Graustufenanalyse erfolgte mit dem Bildbearbeitungsprogramm ImageJ (Version 1.52a). Hierfür wurden die Gelbilder zunächst in Graustufen umgewandelt und anschließend farblich (schwarz/weiß) invertiert, sodass die Banden als dunkle Striche auf hellem Hintergrund dargestellt wurden. Anschließend wurden die Bahnen der Proben definiert und die Banden grafisch als Peak, vergleichbar zu einem Elektropherogramm oder Chromatogramm, dargestellt. Die Peaks konnten schließlich integriert und die daraus resultierende Fläche für die folgende Berechnung ausgewertet werden.

7.5.2 Kapillargelelektrophorese (CGE)

Für die kapillargelelektrophoretische Auftrennung von DNA-Molekülen wurde der *Agilent 2100 Bioanalyzer* der Firma Agilent Technologies Inc. (Santa Clara, USA) je nach Anwendung in Verbindung mit dem *Agilent High Sensitivity DNA Kit* oder dem *Agilent DNA 7500 Reagent Kit* verwendet. Die Durchführung erfolgte gemäß der Herstellerangaben (*Agilent High Sensitivity DNA Kit Quick Start Guide*, Edition 01/2017 bzw. *Agilent DNA 7500 und DNA 12000 Kit Quick Start Guide*, Edition 01/2018).

Um die Menge an benötigten Chips zu reduzieren und damit Kosten einsparen zu können, wurden diese zum Teil nach Nguyen et al. (2014) gewaschen und wiederverwendet.^[261] Hierfür wurden nach einem Lauf zunächst alle Lösungen aus den *wells* entfernt. Anschließend wurden alle *wells* mit 15 µL TAE-Puffer befüllt, der Chip in dem dafür vorgesehenen Vortexmischer geschüttelt und anschließend die Lösungen wieder entfernt. Für die Lagerung wurden die *wells* wie zuvor beschrieben erneut befüllt und der gesamte Chip gründlich mit Parafilm abgedeckt, um ein Austrocknen der Gelmatrix zu verhindern, und bis zur weiteren Verwendung bei 4 °C gelagert. Für die Wiederbenutzung wurden zunächst alle Lösungen aus dem Chip entfernt. Anschließend wurde die neue Gelmatrix, wie in der Gebrauchsanleitung beschrieben, in die Kapillaren eingebracht. Nun wurde die verbliebene frische Gelmatrix aus dem entsprechenden *well* herauspipettiert und in der Pipettenspitze im Dunklen aufbewahrt. Die alte Gelmatrix wurde mit 15 µL TAE-Puffer pro *well* aus diesen entfernt. Anschließend wurde die neue Gelmatrix wieder in das vorgesehene *well* eingebracht und weiter nach Herstellerangabe vorgegangen.^[261] Das beschriebene Vorgehen für die Wiederverwendung von Chips eignet sich allerdings nicht für die Chips des *Agilent High Sensitivity DNA Kit*, da die Kapillaren zu dünn sind und die alte Gelmatrix daher nicht entfernt werden kann.

7.6 Aufreinigung von PCR-Produkten

Zur Aufreinigung von PCR-Produkten wurden je nach Ausgangssituation zwei unterschiedliche kommerziell erhältliche Kits verwendet. Das *Monarch® PCR & DNA Cleanup Kit* der Firma New England Biolabs GmbH (NEB; Frankfurt am Main, Deutschland) wurde für die Aufreinigung von Amplifikaten für eine folgende Sequenzierung ohne unspezifische Banden und zur Aufreinigung von Ansätzen zur Vorbereitung für die CGE angewendet. Das *peqGOLD Gel Extraction Kit* der Firma Peqlab (VWR International GmbH; Darmstadt, Deutschland) hingegen wurde verwendet, wenn ein PCR-Ansatz unspezifische Nebenprodukte aufwies und zunächst eine Auftrennung mittels AGE erfolgen musste. Mit Hilfe des Kits wurde anschließend das gewünschte Amplifikat aus der Agarose isoliert und aufgereinigt.

7.6.1 *Monarch® PCR & DNA Cleanup Kit*

Die Aufreinigung mittels *Monarch® PCR & DNA Cleanup Kit* der Firma NEB erfolgte gemäß den Herstellerangaben (*Instruction Manual*, Version: 2.0, 4/17).

7.6.2 *peqGOLD Gel Extraction Kit*

Die Aufreinigung mit Hilfe des *peqGOLD Gel Extraction Kits* der Firma Peqlab (VWR) wurde nach Herstellerangabe (Arbeitsanleitung, Version: V0815) durchgeführt.

7.7 Klonierung

Bei der Klonierung handelt es sich um eine molekularbiologische Methode mit der u.a. mehrere DNA-Fragmente gleicher Länge getrennt werden können. Während der Klonierung werden die entsprechenden Amplifikate in einen Vektor eingebracht, welcher mittels Elektroporation in elektrokompente *E. coli*-Zellen überführt und in diesen repliziert wird. Hierfür wurde *das Topo™ TA Cloning™ Kit for Sequencing* der Firma Invitrogen (Thermo Fisher Scientific) verwendet. Zunächst wurden die zur Klonierung benötigten Nährböden und -medien hergestellt (siehe Tabelle 22). Die Arbeitsschritte der Klonierung erfolgten unter einer Sicherheitswerkbank, um mögliche Kontaminationen zu vermeiden.

7.7.1 Herstellung elektrokompenter *E. coli*-Zellen

Für die Herstellung elektrokompenter XL-1 *E. coli*-Zellen wurden zunächst einige Zellen aus einer *Stock-Solution* auf LB-Agar ausgestrichen und über Nacht bei 37 °C inkubiert. Am nächsten Tag wurden 50 mL LB-Medium mit einer Kolonie beimpft und erneut über Nacht bei 37 °C im Schüttelinkubator inkubiert. Für die Hauptkultur wurden schließlich 5 mL dieser Vorkultur in 500 mL LB-Medium überführt. Diese wurde so lange bei 37 °C inkubiert, bis der OD-Wert bei 0,4-0,7 lag. Anschließend wurde die Lösung für 10 min bei 4 °C und 5000 g zentrifugiert. Der Überstand wurde dekantiert und das Zellpellet dreimal mit eisgekühlter 10 %iger Glycerin-Lösung gewaschen. Hierbei

wurden folgende Volumina eingesetzt: (1) 400 mL, (2) 200 mL und (3) 15-20 mL. Zuletzt wurde das Zellpellet in 1 mL derselben Glycerin-Lösung aufgenommen, aliquotiert (je 55 μL) und bis zur weiteren Verwendung in flüssigem Stickstoff gelagert.

7.7.2 Herstellung und Aufreinigung des Ligationsansatzes

Der Ligationsansatz wurde, wie in Tabelle 17 dargestellt, gemäß der Herstellerangaben für das *Topo™ TA Cloning™ Kit for Sequencing* der Firma Invitrogen (Life Technologies GmbH, Darmstadt, Deutschland) hergestellt (User Guide TOPO TA Cloning Kit for Sequencing, Pub. No.: MAN0000109, Version: A.0).

Tabelle 17: Herstellung des Ligationsansatzes gemäß der Herstellerangaben des *Topo™ TA Cloning™ Kit for Sequencing* der Firma Invitrogen.

Reagenz	Volumen pro Reaktion
Probe (ca. 50 ng)	1-4 μL *
Salzlösung	1 μL
Vektor (10 ng/ μL)	1 μL
Wasser (reinst)	0-3 μL
	Σ 6 μL

* Eingesetztes Volumen variiert je nach Konzentration der Probe

Zunächst wurden die vier Komponenten vorsichtig in einem 1,5 mL Reaktionsgefäß vermischt. Anschließend wurde die Lösung für 30 min bei RT und darauffolgend 5 min auf Eis inkubiert. Vor der Elektroporation wurde der Ligationsansatz mit dem *Monarch® PCR & DNA Cleanup Kit* (siehe 7.6.1) aufgereinigt (Elutionsvolumen: 10 μL), um Salze zu entfernen, die bei einer Transformation mittels Elektroporation nicht enthalten sein dürfen.

7.7.3 Elektroporation und Transformation

Für die Elektroporation und Transformation wurden zunächst die LBamp-Agar-Platten und das LBamp-Medium auf RT erwärmt, 1 mL SOC-Medium und eine Elektroporationsküvette auf Eis bereitgestellt. Ein Aliquot an elektrokompenten *E. coli*-Zellen wurde auf Eis langsam aufgetaut. Es wurden schließlich 50 μL der Zellsuspension mit 5 μL des aufgereinigten Ligationsansatzes vorsichtig gemischt und in die Küvette überführt. Die Küvette wurde von außen sorgfältig getrocknet, in dem Elektroporationsgerät platziert und eine Spannung von $U = 1700 \text{ V}$ angelegt. Nach Ertönen des Signaltons wurden sofort 1 mL des eisgekühlten SOC-Mediums zugegeben, die Zellsuspension in ein neues 2 mL Reaktionsgefäß überführt und für 1 h bei 37 °C inkubiert. Die Zellsuspension wurde 7 min bei 6000 g zentrifugiert. 100 μL des Überstandes wurden abgenommen, der restliche Überstand dekantiert und das Zellpellet in den 100 μL resuspendiert. Von dieser Lösung wurden 95 μL auf einer LBamp-Agar-Platte ausgestrichen. Die verbliebene Lösung wurde mit 95 μL frischem LB-Medium versetzt, 90 μL der Suspension ausgestrichen und erneut 90 μL LB-Medium zugegeben. Die letzten beiden Schritte wurden ein weiteres Mal wiederholt. Zuletzt wurde die gesamte Suspension ausgestrichen und alle LBamp-Agar-Platten über Nacht bei 37 °C inkubiert.

Um sicherzustellen, dass die für die Elektroporation verwendeten *E. coli*-Zellen nicht verunreinigt sind und kein Plasmid enthalten, wurde jeweils als Kontrolle auch ein Teil dieser Zellen ausgestrichen. Aufgrund des fehlenden Vektors, der ein Resistenzgen gegen Ampicillin enthält, dürfen diese Zellen kein Wachstum zeigen.

7.7.4 Picking & Screening

Die gewachsenen Kolonien wurden ausgezählt und im Zuge des *Picking & Screening* mittels einer PCR und anschließender Detektion über AGE hinsichtlich einer erfolgreichen Transformation und Replikation des integrierten DNA-Fragments überprüft.

Hierfür wurden zunächst einzelne deutlich abzugrenzende Kolonien ausgewählt. Je Kolonie wurden ca. 50 µL TE-Puffer in einem 1,5 mL Reaktionsgefäß vorgelegt. Die Kolonien wurden mit einem autoklavierten Zahnstocher abgenommen und in den TE-Puffer überführt. Mit demselben Zahnstocher wurde anschließend eine LBamp-Agar-Platte beimpft. Hierbei musste darauf geachtet werden, genügend Abstand zwischen den Kolonien einzuhalten, damit diese nicht ineinander überwachsen. Die Platte wurde über Nacht bei 37 °C inkubiert. Die TE-Lösung mit den Zellen wurde für 5 min bei 95 °C erhitzt, um die Zellen aufzuschließend, und anschließend 1 min bei 16.000 g zentrifugiert. Der Überstand wurde für die folgende *Colony*-PCR als Template eingesetzt.

Das Screening der Kolonien, also die Überprüfung, ob ein Vektor mit ligiertem DNA-Fragment aufgenommen wurde, erfolgte mittels *Colony*-PCR. Hierfür wird ein 50 µL-PCR-Ansatz hergestellt, der 1x *Taq Reaction Buffer*, 0,16 mM dNTPs, 5,0 µg BSA, 1,0 µM je Primer, 2,5 U *Taq* DNA Polymerase, 1 µL des Überstandes der Zellsuspension als Template und DNase-/RNase-freies Wasser enthält. Detaillierte Informationen zu dem eingesetzten Primerpaar M13_FW/_RV sind der Tabelle 24 und das angewendete Temperaturprogramm der Tabelle 32, wobei die Dauer pro Zyklusschritt auf 40 s festgelegt wurde, zu entnehmen.

Die Detektion und Visualisierung der Amplifikate erfolgte mittels AGE und Färbung mit *GelRed* (siehe 7.5.1). Bei einem positiven Ergebnis des Screenings wurden am nächsten Tag je 20 mL LB-Medium mit den entsprechenden Kolonien beimpft und über Nacht bei 37 °C inkubiert.

7.7.5 Plasmidextraktion

Für die Plasmidextraktion wurde das *Monarch® Plasmid Miniprep Kit* von New England Biolabs® verwendet. Die Durchführung erfolgte gemäß der Herstellerangabe (*Instruction Manual*, Version 2.0, 11/17). Die erhaltenen Plasmid-Isolate wurden fluorimetrisch quantifiziert (siehe 7.3.1).

7.7.6 Finale Überprüfung und Sequenzierung

Um sicherzustellen, dass die Plasmidisolate nicht überwiegend aus chromosomaler DNA der Bakterien bestehen, wurde eine AGE mit einem 1%igem Agarosegel durchgeführt. Nach erfolgreicher Überprüfung wurden die Plasmide mit Reinstwasser auf eine Konzentration von 30-100 ng/µL verdünnt und über die Firma GATC Biotech (heute: Eurofins Genomics) sequenziert.

Hierfür wurde der bei der Firma vorhandene Universal-Primer M13-FP genutzt, welcher in räumlicher Nähe zu dem eingebauten DNA-Fragment bindet.

7.8 Erstellung von *Alignments*

Bei einem Alignment handelt es sich um einen Vergleich verschiedener Sequenzen miteinander. Diese wurden entweder mit dem Online-Tool *Clustal Omega Alignment Tool* des EMBL-EBI (*European Molecular Biology Laboratory - European Bioinformatics Institute*) oder der Software *Genome Compiler* von Genome Compiler Corporation (Los Altos, Kalifornien, USA) durchgeführt.

7.9 Sequenzierung des cpGenoms mittels NGS

Im Rahmen der vorliegenden Arbeit wurden insgesamt 13 Haselnussproben, davon 12 unterschiedliche Sorten, mittels NGS sequenziert. Die Sequenzierung erfolgte zum einen an dem an der *Hamburg School of Food Science* vorhandenen MiSeq™ von Illumina Inc. (San Diego, USA) und zum anderen bei der Firma GATC Biotech (heute: Eurofins Genomics). In Tabelle 18 sind alle sequenzierten Proben mit Angabe der Herkunft der Haselnussblätter, der Institution, welche die Sequenzierung durchgeführt hat, sowie des verwendeten Kits für die *Library preparation* aufgeführt. Detaillierte Informationen bzgl. der Herkunft sind unter 9.6 in Tabelle 29 aufgelistet.

Tabelle 18: Übersicht der sequenzierten Haselnussblattproben inkl. der Angabe bzgl. der Herkunft, der Institution, die die Sequenzierung durchgeführt hat, sowie des eingesetzten Kits für die *Library preparation*. (Für die *Library preparation* verwendete Kits der Firma Illumina Inc. (San Diego, USA): ¹ = Nextera XT DNA Library Prep Kit, ² = Nextera™ DNA Flex Library Prep Kit)

Probe Nr.	Sorte	Herkunft	HSFS	GATC
B1	Ata Baba	Türkei		X
B3	Barcelona	Frankreich	X ²	
B4	Butler	Frankreich	X ²	
B5	Corabel	Frankreich	X ²	
B6	Corabel	Schweiz	X ²	
B7	Hallsche Riesen	Deutschland		X
B8	Negret	Deutschland	X ¹	
B9	Palaz	Türkei	X ¹	
B10	Pauetet	Deutschland	X ¹	
B13	Segorbe	Deutschland		X
B16	Tonda di Giffoni	Deutschland		X
B17	Tonda di Giffoni	Frankreich		X
B18	Tonda di Giffoni	Frankreich		X
B19	Tonda Gentile Romana	Deutschland	X ¹	
B20	Tonda Gentile Trilobata	Deutschland	X ¹	
B21	C. avellana (Sorte unbekannt)	Deutschland	X ¹	X

Für die Sequenzierung wurde ausschließlich Blattmaterial eingesetzt. Die Gesamt-DNA-Isolierung wurde wie unter 7.2.1 beschrieben durchgeführt. Die DNA-Isolate wurden fluorimetrisch quantifiziert (siehe 7.3.1) und die Reinheit mittels Nanodrop (7.3.2) überprüft. Bei ausreichender Qualität wurden die DNA-Isolate direkt für die Sequenzierung eingesetzt.

Für die Probenvorbereitung (*Library preparation*) wurden die Kits Nextera XT DNA Library Prep Kit und Nextera™ DNA Flex Library Prep Kit der Firma Illumina Inc. (San Diego, USA) eingesetzt. Die Durchführung erfolgte gemäß der Herstellerangaben (Nextera XT DNA Library Prep Kit, Reference Guide: Document #15031942 v03, Februar 2018; Nextera™ DNA Flex Library Prep Kit, Reference Guide: Document #1000000025416 v02, Juni 2018). Die gewählten Parameter der variablen Angaben sind der Tabelle 19 zu entnehmen.

Tabelle 19: Übersicht über die gewählten Parameter bei variablen Angaben in den Anleitungen.

Nextera XT DNA Library Prep Kit	
Eingesetzte DNA-Menge	1 ng
PCR-Zyklen (<i>Amplify Tagmented Mix</i>)	12 PCR-Zyklen
Ladekonzentration	10 nM
Anteil PhiX	1 %
Nextera™ DNA Flex Library Prep Kit	
Eingesetzte DNA-Menge	100 ng
PCR-Zyklen (<i>Amplify Tagmented Mix</i>)	5 PCR-Zyklen
Ladekonzentration	7,5 pM
Anteil PhiX	1 %

Die Auswertung der Rohdaten erfolgte an der Universität Tübingen am Lehrstuhl für Angewandte Bioinformatik wie in der zugehörigen Publikation (Lang et al. 2021) beschrieben.^[210]

7.10 Resequenzierung vier ausgewählter Bereiche des cpGenoms

Anhand der erhaltenen Sequenzen wurden vier Bereiche A-D mit möglichen Sequenzunterschieden auf dem cpGenom definiert, die anschließend mittels Sanger-Sequenzierung verifiziert werden sollten. Die Sequenzen der vier Bereiche inklusive der für die Amplifizierung entwickelten Primer sind unter 9.10.2 der Abbildung 55 bis Abbildung 58 zu entnehmen. Es wurden von den mittels NGS analysierten Blattproben jeweils zwei weitere DNA-Isolate, die aus je nur einem Haselnussblatt hergestellt wurden, untersucht. Das eingesetzte Probenmaterial ist der Tabelle 29 zu entnehmen. Zusätzlich wurde eine Probe georgischer Haselnüsse dem Probenpool hinzugefügt und ebenfalls mittels Sanger-Sequenzierung analysiert. Da hier jedoch ausschließlich Nuss-Material vorhanden war, wurde dieses eingesetzt. Diese Probe ist der Tabelle 27 zu entnehmen.

Die PCR wurde jeweils in einem 100 µL-Ansatz mit 1x HF-Buffer, 5 µM dNTPs, 100 µM je Primer, 2 U *S7 Fusion High-Fidelity* DNA Polymerase, 2 ng DNA und DNase-/RNase-freies Wasser durchgeführt. Die Primer (Bereich A: A_FW/RV, Bereich B: B.1_FW/RV, Bereich C: C_FW/RV, Bereich D: D.1_FW/RV) sowie die zugehörigen Annealingtemperaturen sind in Tabelle 24 zusammengefasst. Das Temperaturprogramm ist in Tabelle 33 dargestellt. Im Anschluss an die PCR wurden die Amplifikate bis zur weiteren Verwendung bei 4 °C gelagert. Vor der Sequenzierung erfolgte die Aufreinigung mittels *Monarch® PCR & DNA Cleanup Kit* (siehe 7.6.1) bzw. *peqGOLD Gel Extraction Kit* mit vorangegangener AGE (siehe 7.6.2). Die Sanger-Sequenzierung wurde durch die Firma GATC Biotech (heute: Eurofins Genomics) durchgeführt. Das anschließende *Alignment* und der Vergleich der Sequenzen wurde mit den unter 7.8 erläuterten Tools durchgeführt.

7.11 Entwicklung von *targeted*-Methoden auf Basis polymorpher Bereiche des cpGenoms

7.11.1 Nachweis der Indels 1 und 3 in den Bereichen B und C

7.11.1.1 Qualitativer Nachweis der Indels 1 und 3

Das für den qualitativen Nachweis der Indels 1 und 3 in den Bereichen B und C verwendete Probenmaterial ist der Tabelle 27 zu entnehmen. Die PCR erfolgte ausschließlich unter Verwendung von Gesamt-DNA-Isolaten von Haselnüssen. Die Isolierung hierfür wurde sowohl mit Hilfe der CTAB-Methode (siehe 7.2.2.1) als auch der *Rapid Extraction* (siehe 7.2.2.2) durchgeführt.

Die PCR für die qualitative Analyse der Bereiche B und C wurde mit einem 10 µL-Reaktionsansatz mit 1x *Taq Reaction Buffer*, 0,3 mM dNTPs, 0,7 µM je Primer, 0,5 U *Taq* DNA Polymerase, 1 ng Gesamt-DNA und DNase-/RNase-freiem Wasser durchgeführt. Die benötigten Primerpaare B.1_FW/RV und C_FW/RV sowie deren spezifische Annealingtemperatur sind der Tabelle 24 zu entnehmen. Das angewendete Temperaturprogramm ist in Tabelle 32 dargestellt. Die Dauer eines Zyklusschrittes betrug jeweils 30 s. Nach abgeschlossener PCR wurden die Reaktionslösungen bis zur weiteren Verwendung auf 4 °C gekühlt. Die Detektion erfolgte mittels AGE (siehe 7.5.1).

Der qualitative Nachweis der Indels wurde auch auf verarbeitete Haselnüsse angewendet. Hierfür wurden sowohl zwei kommerziell erworbene Proben gerösteter Haselnüsse aus der Türkei (Nr. V2) und Italien (Nr. V3), als auch selbst gebackene Kekse mit französischen Haselnüssen (Nr. V1) untersucht (siehe Tabelle 28). Die Proben wurden wie unter 7.1.1 beschrieben für die DNA-Isolierung vorbereitet und diese anschließend nach 7.2.2.1 durchgeführt. Die PCR erfolgte wie zuvor beschrieben.

7.11.1.2 Quantifizierung von Haselnussmischungen auf Basis der identifizierten Indels 1 und 3

Die validierten Indels dienten als Grundlage für die Entwicklung einer *targeted*-Methode für den quantitativen Nachweis von Mischungen von Haselnüssen der beiden Gruppen (+ Insert und - Insert).

Die Quantifizierung erfolgte anhand einer externen Kalibrierung. Für die Kalibrierstandards wurden gemahlene georgische Haselnüsse (Nr. S4) und italienische Haselnüsse der Sorte T. G. Trilobata (Nr. 9) in Verhältnissen von 2,5-80% (bezogen auf den Anteil georgischer Nüsse) gemischt und mittels CTAB-Methode (siehe 7.2.2.1) isoliert. Um die Reproduzierbarkeit zu überprüfen, wurden drei identische Kalibrierreihen hergestellt. Um eine Aussage bzgl. der quantitativen Bestimmung anhand der erhaltenen Kalibrierung treffen zu können, wurde zusätzlich ein Test-Set erstellt. Hierfür wurden fünf Mischungen mit 5%, 10%, 30%, 50% und 70% georgischen Haselnüssen jeweils in Dreifachbestimmung nach dem gleichen Vorgehen und mit denselben Haselnussproben hergestellt. Die PCR erfolgte mit dem Primerpaar B.2_FW/RV (siehe Tabelle 24) unter Verwendung eines 10 µL-Reaktionsansatzes mit 1x *Taq Reaction Buffer*, 0,3 mM dNTPs, 50 nM je Primer, 0,5 U *Taq* DNA Polymerase (*DreamTaq* DNA Polymerase), 1 ng Gesamt-DNA und DNase-/RNase-freiem Wasser. Das angewendete Temperaturprogramm ist der Tabelle 32 zu entnehmen. Die Reaktionszeit pro Zyklusschritt betrug jeweils 15 s und die Anzahl an Zyklen wurde von 35 auf 45 erhöht. Die Detektion wurde sowohl mittels CGE (siehe 7.5.2), als auch AGE mit anschließender Graustufenanalyse (siehe 7.5.1) durchgeführt. Die Aufreinigung vor der Messung mittels CGE erfolgte mit Hilfe des *Monarch® PCR & DNA Cleanup Kits* (siehe 7.6.1).

Bestimmung des LOD und LOQ:

Für die Bestimmung der Nachweisgrenze (LOD = *limit of detection*) und der Bestimmungsgrenze (LOQ = *limit of quantitation*) wurden drei DNA-Isolate bestehend aus 100% T. G. Trilobata (Nr. 9) wie zuvor beschrieben analysiert und sowohl mittels CGE, als auch AGE gemessen. Es wurde das Hintergrundrauschen der Matrix an der Position, an der sonst das Signal bzw. die Bande georgischer Haselnüsse detektiert wurde, bestimmt. Anhand des Mittelwertes der Dreifachbestimmung ($\bar{x}_{0\%}$) und der zugehörigen Standardabweichung ($s_{0\%}$) wurden anschließend der LOD und LOQ wie folgt berechnet.

$$\text{LOD} = \bar{x}_{0\%} + (3 \times s_{0\%}) \quad \text{Gleichung 1}$$

$$\text{LOQ} = \bar{x}_{0\%} + (10 \times s_{0\%}) \quad \text{Gleichung 2}$$

LOD	Nachweisgrenze relative Area [%]
LOQ	Bestimmungsgrenze relative Area [%]
$\bar{x}_{0\%}$	Mittelwert der relativen Area [%]
$s_{0\%}$	Standardabweichung

Basierend auf diesen Formeln wurden der LOD und LOQ als prozentuale relative Area berechnet. Um eine Aussage darüber treffen zu können in welchem Bereich die Grenzen bezogen auf den Anteil georgischer Haselnüsse liegen, wurden die erhaltenen Werte mit den relativen Areas der Kalibriergeraden verglichen. Eine Umrechnung anhand der Regressionsgeraden erfolgte nicht, da diese im Bereich kleiner prozentualer Beimischungen keine ausreichende Genauigkeit aufwiesen. In Abschnitt 6.3.1.3 wird darauf näher eingegangen. Zuletzt wurde die visuelle Bewertung der

Elektropherogramme und Agarosegelbilder der Kalibriergeraden berücksichtigt und ggf. die Werte entsprechend nochmals angepasst.

7.11.2 PCR-RFLP für den Nachweis georgischer Haselnüsse auf Basis eines SNPs in Bereich D

7.11.2.1 Qualitativer Nachweis georgischer Haselnüsse mittels PCR-RFLP

Georgische Haselnüsse weisen im Bereich D einen spezifischen SNP (SNP 6) auf, der mittels PCR-RFLP nachgewiesen werden kann. Das für den qualitativen Nachweis verwendete Probenmaterial ist der Tabelle 27 zu entnehmen. Die PCR erfolgte ausschließlich mit Gesamt-DNA-Isolaten von Haselnüssen. Die Isolierung hierfür wurde sowohl mit Hilfe der CTAB-Methode (siehe 7.2.2.1) als auch der *Rapid Extraction* (7.2.2.2) durchgeführt.

Die PCR für die Amplifizierung des Bereichs D erfolgte in einem 10 µL-Reaktionsansatz mit 1x *Taq Reaction Buffer*, 0,5 mM dNTPs, 0,8 µM je Primer (D.2_FW/RV, siehe Tabelle 24), 0,5 U *Taq* DNA Polymerase, 1 ng Gesamt-DNA sowie DNase-/RNase-freiem Wasser. Das verwendete Temperaturprogramm ist der Tabelle 32 zu entnehmen. Die Dauer eines Zyklusschritts betrug jeweils 20 s. Anschließend wurden die Proben bis zur weiteren Verwendung auf 4 °C gekühlt. Der enzymatische Verdau wurde in einem 10 µL-Reaktionsansatz mit 1x *CutSmart Buffer*, 1 U *MfeI*, 2 µL des PCR-Ansatzes und DNase-/RNase-freiem Wasser durchgeführt. Die Inkubation erfolgte in einem Wasserbad bei 37 °C für eine Stunde. Anschließend wurden die Proben bis zur weiteren Verwendung bei 4 °C gelagert. Das Enzym *MfeI* benötigt keine weitere Deaktivierung. Die Detektion wurde mittels AGE durchgeführt (siehe 7.5.1).

Für die Überprüfung der Anwendbarkeit der PCR-RFLP auf verarbeitete Haselnüsse wurden zwei Proben georgischer und aserbaidischischer Haselnüsse (Nr. V4 und V5) sowie Haselnüsse aus einer Schokolade (Nr. V6) (siehe Tabelle 28) analysiert. Die Proben wurden wie unter 7.1.2 und 7.1.1 erläutert für die DNA-Isolierung vorbereitet und diese anschließend nach der CTAB-Methode (siehe 7.2.2.1) durchgeführt. Die anschließende qualitative Untersuchung erfolgte wie zuvor beschrieben.

7.11.2.2 PCR-RFLP für den semi-quantitativen Nachweis georgischer Haselnüsse

Für die quantitative Bestimmung von Haselnussmischungen, bei denen der Anteil georgischer Haselnüsse mittels RFLP-Analyse bestimmt werden soll, wurden dieselben Haselnussmischungen bestehend aus georgischen und italienischen Haselnüssen, wie sie unter 7.11.1.2 beschrieben wurden, eingesetzt. Nach Optimierung wurde das Primerpaar D.2_FW/RV (siehe Tabelle 24) für die PCR verwendet. Diese wurde in einem vergrößerten Reaktionsansatz mit einem Volumen von 20 µL mit 1x *Taq Reaction Buffer*, 0,5 mM dNTPs, 0,8 µM je Primer, 1 U *Taq* DNA Polymerase, 15 ng Gesamt-DNA und DNase-/RNase-freiem Wasser durchgeführt. Der anschließende enzymatische Verdau erfolgte mit einem optimierten Reaktionsansatz mit 45 µL Reaktionsvolumen und 1x *CutSmart Buffer*, 9 U *MfeI*, 15 µL des PCR-Ansatzes und DNase-/RNase-freiem Wasser. Der Reaktionsansatz wurde für eine Stunde bei 37 °C im Wasserbad inkubiert und anschließend bei 4 °C bis zur weiteren Verwendung gelagert.

Für die Detektion mittels AGE (siehe 7.5.1) wurden 15 µL des Reaktionsansatzes direkt eingesetzt. Für die Messung mittels CGE wurden die verbliebenen 30 µL mit Hilfe des *Monarch® PCR & DNA Cleanup Kits* aufgereinigt (siehe 7.6.1) und die Fragmente anschließend kapillargelelektrophoretisch aufgetrennt (siehe 7.5.2). War nur eine der beiden Detektionsmethoden erwünscht, wurden die Reaktionsansätze entsprechend angepasst.

Bestimmung des LOD und LOQ:

Die Bestimmung des LOD und LOQ erfolgte nach demselben Prinzip wie unter 7.11.1.2 beschrieben. Das Grundrauschen wurde dabei in dem Bereich ermittelt, in dem sonst bei georgischen Haselnüssen das längere der beiden Fragmente (413 bp) nach dem enzymatischen Verdau detektiert wurde.

7.11.3 HRMA für die Differenzierung georgischer und aserbaidischer Haselnüsse

7.11.3.1 Berechnung von Schmelzkurven

Die Berechnung von Schmelzkurven erfolgte mit dem Online-Tool *uMelt^{Quartz}* (DNA-UTAH.ORG). Folgende Tabelle listet alle Parameter auf, die für die Berechnung individuell angepasst werden können.

Tabelle 20: Parameter, die für die Berechnung von Schmelzkurven mit Hilfe des Online-Tools *uMelt^{Quartz}* angegeben werden müssen.

Parameter	Bedeutung	Angabe (z.B.)
Thermodynamics	Thermodynamische Annahmen	Blake & Delcourt (1998) ^[262]
Free [Mg ⁺⁺]	Konzentration freier Mg ²⁺ -Ionen	0,8 mM
[Mono+]	Konzentration freie monovalenter Kationen	10 mM
DMSO	Konzentration DMSO	0%
Resolution	Abschnitt der Berechnungspunkte	High – 0,25 °C

Mit Hilfe des Tools können Sequenzen mit einer Länge von bis zu 1000 bp analysiert werden. Es kann jede Basenabfolge individuell angegeben werden. Die Daten für die berechneten Schmelzkurven können anschließend als csv-Datei gespeichert und mittels Microsoft Excel bearbeitet werden.

7.11.3.2 Methodenentwicklung und -validierung

Die HRMA wurde auf Basis identifizierter Sequenzunterschiede innerhalb des cpGenoms georgischer Haselnüsse im Bereich D entwickelt. Das verwendete Probenmaterial ist der Tabelle 27 zu entnehmen. Die HRMA wurde mit dem *AccuMeltTM HRM SuperMix* der Firma QuantaBio (Beverly, USA), der 2-fach konzentriert ist und einen geeigneten Puffer, dNTPs sowie die *AccuStart Taq DNA Polymerase* und den Interkalationsfarbstoff *SytoTM 9 Green* enthält, durchgeführt. Es wurde ein 20 µL-PCR-Ansatz mit 1x *AccuMeltTM HRM SuperMix*, 0,5 µM je Primer, 2 ng Gesamt-DNA und DNase-/RNase-freiem Wasser hergestellt. Es wurde das Primerpaar D.HRMA_FW/RV (siehe Tabelle 24) eingesetzt. Das angewendete Temperaturprogramm ist Tabelle 34 zu entnehmen. Die Auswertung der erhaltenen Daten erfolgte mit der *CFX ManagerTM* Software von Bio-Rad Laboratories, Inc.

(Hercules, USA) und zusätzlich mit der *Precision Melt AnalysisTM* Software, ebenfalls der Firma Bio-Rad Laboratories, Inc..

Nach erfolgreicher Optimierung der Reaktionsbedingungen wurden diese mittels Dreifachbestimmung eines definierten Probenpools bestehend aus vier türkischen (Nr. 128, 131, 134, 142), fünf georgischen (Nr. S4, S5, S6, 151, 152), fünf aserbaidischen (Nr. S2, 102, 121, 122, 123) und neun europäischen (Nr. F10, 7, 8, 9, 31, 32, 107, 108, 109) Haselnussproben validiert. Des Weiteren erfolgte die Überprüfung der Anwendbarkeit der HRMA auf prozessierte Produkte. Hierfür wurde die Proben Nr. V3-V6 mit der zuvor validierten Methode untersucht.

7.11.3.3 Quantitative Bestimmung von Haselnussmischungen

Für die Entwicklung einer Methode zur Quantifizierung von Mischungen wurden zunächst DNA-Isolate georgischer (Nr. S4) und aserbaidische (Nr. 123) Haselnüsse in den Verhältnissen 0, 2,5, 5, 10, 20, 40, 60 und 100%, bezogen auf den Anteil des Isolats georgischer Nüsse, hergestellt. Die DNA-Isolate waren zuvor mittels der CTAB-Methode (siehe 7.2.2.1) hergestellt worden. Zusätzlich wurden im zweiten Schritt Matrixmischungen untersucht. Hierfür wurden gemahlene georgische (Nr. S4) und aserbaidische (Nr. 102) Haselnüsse in den Verhältnissen 2,5, 5, 10, 30, 50 und 70% (bezogen auf den Anteil georgischer Nüsse) gemischt und anschließend mit Hilfe der CTAB-Methode (siehe 7.2.2.1) isoliert. Die Analyse der Isolate mittels HRMA erfolgte mit den unter 7.11.3.2 beschriebenen Reaktionsbedingungen.

7.12 Analyse definierter Bereiche des Karyoms

7.12.1 Analyse der ITS-Bereiche

Die Analyse der ITS-Bereiche erfolgte in Anlehnung an die Untersuchungen von Erdogan und Mehlenbacher (2000).^[3] Das eingesetzte Probenmaterial ist im Anhang unter 9.6 in Tabelle 27 aufgelistet. Aus Gründen der Vergleichbarkeit wurde das veröffentlichte Primerpaar (ITS-5 (FW), 26S-25R (RV) siehe Tabelle 24) für die Analyse eingesetzt. Die PCR wurde in einem 20 µL-Ansatz mit 1x *DreamTaq Reaction Buffer*, 150 µM dNTPs, 2 µg BSA, 1 µM je Primer, 1,0 U *DreamTaq* DNA Polymerase, 8 ng Gesamt-DNA und DNase-/RNase-freiem Wasser durchgeführt. Für die Amplifikation wurde das Standard-Temperaturprogramm (siehe Tabelle 32) mit der primerspezifischen Annealingtemperatur von 50,1 °C und einer Dauer von 20 s pro Zyklusschritt herangezogen.

Die PCR-Ansätze wurden mit Hilfe des *Monarch[®] PCR & DNA Cleanup Kits* von NEB aufgereinigt (siehe 7.6.1). Die Konzentration der Amplifikate wurde fluorimetrisch quantifiziert (siehe 7.3.1). Bei einer ausreichenden DNA-Konzentration von mind. 10 ng/µL wurden die Isolate mittels AGE (siehe 7.5.1) auf das Vorhandensein von Nebenprodukten überprüft. Nach Bestätigung der Reinheit erfolgte die Sequenzierung der PCR-Fragmente über die Firma GATC Biotech (heute: Eurofins Genomics). Die

erhaltenen Sequenzen wurden mit den unter 7.8 beschriebenen Tools miteinander verglichen und hinsichtlich Sequenzunterschieden untersucht.

7.12.2 Analyse der WRKY-Gene

Die für die Analyse der WRKY-Gene herangezogenen Proben sind im Anhang der Tabelle 27 und Tabelle 29 zu entnehmen. Im ersten Schritt sollten Sequenzen der WRKY-Gene der *C. avellana* identifiziert werden. Hierfür wurde zunächst mit den von Borrone et al. (2004) veröffentlichten degenerierten Primern WRKY_FP1 und WRKY_RP2 (siehe Tabelle 24) eine PCR durchgeführt.^[69] Der Reaktionsansatz mit 25 µL enthält 1x *DreamTaq Reaction Buffer*, 0,2 mM dNTPs, 0,4 µM je Primer, 0,25 U *DreamTaq* DNA Polymerase, 5 ng Gesamt-DNA (Proben Nr. 31 und 33) und DNase-/RNase-freies Wasser. Das zugehörige Temperaturprogramm ist der Tabelle 35 zu entnehmen. Die Auftrennung und Visualisierung der Fragmente erfolgte mittels AGE (siehe 7.5.1).

Für die Sequenzierung der erhaltenen Fragmente wurden diese kloniert, da nicht ausgeschlossen werden konnte, dass eine Bande unterschiedliche Amplifikate enthält. Der Fokus lag dabei auf den Banden mit Längen von 280 bp, 320 bp, 800 bp und 950 bp. Zunächst wurde die PCR mit einem vergrößerten Volumen von 100 µL wiederholt. Die Auftrennung der Fragmente erfolgte mittels AGE. Die einzelnen Banden wurden mit Hilfe eines Skalpellens aus dem Gel ausgeschnitten und die DNA wie unter 7.6.2 beschrieben aus der Agarose isoliert und aufgereinigt. Die Konzentrationsbestimmung erfolgte fluorimetrisch (siehe 7.3.1). Die Klonierung wurde wie unter 7.7 erläutert durchgeführt. Auf Basis der anschließend erhaltenen Sequenzen wurden vier sequenzspezifische Primerpaare WRKY-280bp_K1_FW/RV, WRKY-280K2_FW/RV, WRKY-320K3_FW/RV und WRKY-950K16_FW/RV (siehe Tabelle 24) entwickelt.

Für die anschließend durchgeführten Amplifikationen wurden Blattisolate der Sorten Ata Baba (Nr. B1 und B2), Palaz (Nr. B26 und B27) und Pautet (Nr. B5 und B7) eingesetzt. Die PCR wurde jeweils in einem 50 µL-Reaktionsansatz mit 1x *Taq Reaction Buffer*, 0,16 mM dNTPs, 5,0 µg BSA, 1,0 µM je Primer, 2,5 U *Taq* DNA Polymerase, 1 ng DNA-Isolat (Blatt) und DNase-/RNase-freiem Wasser und dem Temperaturprogramm gemäß Tabelle 35 durchgeführt. Die primerspezifischen Annealingtemperaturen sind der Tabelle 24 zu entnehmen. Die Amplifikate wurden über die AGE aufgetrennt, die gewünschten Fragmente isoliert und aufgereinigt und anschließend mittels Sanger-Sequenzierung über die Firma GATC Biotech (heute: Eurofins Genomics) sequenziert. Die erhaltenen Sequenzen wurden schließlich mit Hilfe des *Clustal Omega Alignment Tools* des EMBL-EBI (*European Bioinformatics Institute*) miteinander verglichen, um Sequenzunterschiede zu identifizieren.

7.13 Analyse von SCoT Polymorphismen

7.13.1 Probenmaterial

Für die Verwandtschaftsanalyse und die Identifizierung geeigneter SCoT-Marker auf Basis von *Start Codon Targeted* Polymorphismen zur Sortenunterscheidung von Haselnüssen wurden Gesamt-DNA-Isolate von 13 Haselnusssorten herangezogen. Hierbei wurde darauf geachtet, dass soweit vorhanden von jeder Sorte drei Proben unterschiedlicher Herkunft und Erntejahr für die Analyse verwendet wurden. Die verwendeten Proben sind der Tabelle 27 im Anhang unter 9.6 zu entnehmen.

7.13.2 Durchführung der PCR

Die PCR erfolgte mit einer Auswahl der von Collard und Mackill (2009) veröffentlichten Primer (siehe Tabelle 24).^[163] Nach erfolgreicher Optimierung wurden die Analysen mit einem 10 µL-Reaktionsansatz mit 1x *DreamTaq Reaction Buffer*, 0,3 mM dNTPs, 1 µg BSA, 0,5 µM je Primer, 0,15 U *DreamTaq* DNA Polymerase, 1,0 ng Gesamt-DNA (Primer SCoT 7 und 17: 3 ng) und DNase-/RNase-freiem Wasser durchgeführt. Bei der Wahl und Optimierung des Temperaturprogramms zeigte das von Collard und Mackill (2009) veröffentlichte die besten Ergebnisse (siehe Abschnitt 9.8 Tabelle 36).^[163] Lediglich für einen Primer (SCoT 21) musste die Annealingtemperatur auf 46,4 °C reduziert werden.

7.13.3 Auswertung

Die erhaltenen Gelbilder wurden zunächst visuell hinsichtlich der vorhandenen Banden ausgewertet. Darauf basierend wurde eine 1/0-Matrix erstellt. 1 steht dabei für „Bande vorhanden“ und 0 für „Bande nicht vorhanden“. Die erhaltene Matrix diente anschließend zum einen für die Erstellung eines *Neighbour-Joining*-Diagramms und zum anderen für eine Hauptkomponentenanalyse (PCA = *principle component analysis*). Das *Neighbour-Joining*-Diagramm wurde mit Hilfe des Programms PAST erstellt.^[263] Für die Berechnung wurde der *Similarity Index* nach Jaccard und ein *Bootstrap* von 10.000 verwendet. Die PCA wurde mit dem Programm MATLAB erstellt.^[264] Der Berechnung liegt ein SVD-Algorithmus zugrunde.

7.14 Entwicklung sortenspezifischer SCAR-Marker

Die Entwicklung sortenspezifischer SCAR-Marker sollte auf Basis einer im Vorfeld dieser Arbeit von Felbinger et al. (2020) entwickelten RAPD-PCR durchgeführt werden. Ziel war es spezifische Banden zu identifizieren und diese anschließend zu klonieren, um die genaue Sequenz zu bestimmen. Darauf basierend sollten schließlich individuelle und sortenspezifische Primer entwickelt werden.

7.14.1 RAPD-PCR

Die von Felbinger et al. (2020) durchgeführte RAPD-PCR basierte auf 13 unterschiedlichen Haselnussorten.^[234] Anhand der vorliegenden Ergebnisse wurde der RAPD-Primer 17 (siehe Tabelle 24) und die Sorten Tombul (Nr. 119), Ennis (Nr. 77), Butler (Nr. 69), Segorbe (Nr. 67) und Pauetet (Nr. 72) für die weitere Methodenentwicklung ausgewählt. Um die Reproduzierbarkeit der RAPD-PCR sicherzustellen, wurde diese zunächst mit dem genannten Primer und den aufgeführten Proben nach Felbinger et al. (2020) wiederholt. Die PCR erfolgte mit einem 10 µL-Reaktionsansatz mit 1x *DreamTaq Reaction Buffer*, 0,8 mM dNTPs, 1 µg BSA, 5 µM Primer, 0,5 U *DreamTaq* DNA Polymerase, 2 ng Gesamt-DNA und DNase-/RNase-freiem Wasser. Das angewendete Temperaturprogramm ist der Tabelle 32 und die Annealingtemperatur Tabelle 24 zu entnehmen. Die Dauer pro Zyklusschritt betrug 40 s. Die Auftrennung der Amplifikate erfolgte mittels AGE.

Nach bestätigter Reproduzierbarkeit wurden die für die weitere Entwicklung ausgewählte Bande mit einem Isolat der Probe Nr. 119 (Sorte: Tombul) in ausreichender Konzentration hergestellt, indem der PCR-Ansatz auf 100 µL vergrößert wurde. Anschließend wurde das gewünschte Amplifikat nach erfolgter AGE mit einem Skalpell aus dem Agarosegel ausgeschnitten und mit dem *peqGOLD Gel Extraction Kit* (siehe 7.6.2) aufgereinigt. Die Konzentration des erhaltenen Amplifikats wurde fluorimetrisch bestimmt (siehe 7.3.1).

7.14.2 Klonierung

Aufgrund möglicher Überlagerungen mehrerer Fragmente mit der gleichen Anzahl an Basenpaaren, erfolgte vor der Sequenzierung eine Trennung mittels Klonierung. Diese wurde wie unter 7.7 beschrieben durchgeführt.

7.14.3 Entwicklung der SCAR-Primer

Anhand der erhaltenen Sequenzen wurden die SCAR-Primer entwickelt. Dies ist detailliert im Abschnitt „Ergebnisse und Diskussion“ unter 6.6 erläutert. Die designten Primer sind der Tabelle 24 zu entnehmen. Ihre Spezifität wurde anschließend mittels PCR überprüft. Hierfür wurde ein 10µL-Reaktionsansatz mit 1x *Taq Reaction Buffer*, 0,8 mM dNTPs, 1,0 µg BSA, 1,0 µM je Primer, 0,5 U *Taq* DNA Polymerase, 1,0 ng Gesamt-DNA und DNase-/RNase-freiem Wasser gewählt. Das angewendete Temperaturprogramm ist der Tabelle 32 und die primerspezifischen Annealingtemperaturen Tabelle 24 zu entnehmen. Die Dauer pro Zyklusschritt betrug 30 s.

8 Literaturverzeichnis

- (1) Johnson, R. *Food Fraud and "Economically Motivated Adulteration" of Food and Food Ingredients*, **January 10, 2014**.
- (2) Janick, J.; Paull, R. E., Eds. *The Encyclopedia of Fruit & Nuts*; CABI: Oxfordshire, UK, **2008**.
- (3) Erdogan, V.; Mehlenbacher, S. A., Phylogenetic Relationships of *Corylus* Species (Betulaceae) Based on Nuclear Ribosomal DNA ITS Region and Chloroplast *matK* Gene Sequences. *Syst. Bot.* **2000**, *25*, 4, 727–737.
- (4) Mehlenbacher, S. A., HAZELNUTS (CORYLUS). *Acta Hort.* **1991**, 290, 791–838.
- (5) Lieberei, R.; Reisdorff, C. *Nutzpflanzen*, 8th edn.; Georg Thieme Verlag: Stuttgart, **2012**.
- (6) Boccacci, P.; Akkak, A.; Botta, R., DNA typing and genetic relations among European hazelnut (*Corylus avellana* L.) cultivars using microsatellite markers. *Genome* **2006**, *49*, 6, 598–611.
- (7) Kadereit, J. W.; Körner, C.; Kost, B.; Sonnewald, U. *Lehrbuch der Pflanzenwissenschaften*, 37th edn.; Springer Spektrum: Berlin, Heidelberg, **2014**.
- (8) Gökirmak, T.; Mehlenbacher, S. A.; Bassil, N. V., Characterization of European hazelnut (*Corylus avellana*) cultivars using SSR markers. *Genet. Resour. Crop. Evol.* **2009**, *56*, 2, 147–172.
- (9) Rovira, M., Genetic variability among hazelnut (*Corylus avellana* L.) cultivars. *Acta Hort.* **1997**, 445, 45–50.
- (10) Özenç, N.; Özenç, D. B., Nut traits and nutritional composition of hazelnut (*Corylus avellana* L.) as influenced by zinc fertilization. *Journal of the science of food and agriculture* **2015**, *95*, 9, 1956–1962.
- (11) Rowley, E. R.; VanBuren, R.; Bryant, D. W.; Priest, H. D.; Mehlenbacher, S. A.; Mockler, T. C., A Draft Genome and High-Density Genetic Map of European Hazelnut (*Corylus avellana* L.) **2018**.
- (12) Özdemir, M.; Devres, O., Turkish hazelnuts: Properties and effect of microbiological and chemical changes on quality. *Food Rev. Int.* **1999**, *15*, 3, 309–333.
- (13) İslam, A., Hazelnut cultivation in Turkey. *Akademik Ziraat Dergisi* **2018**, *7*, 2, 259–266.
- (14) Petriccione, M.; Ciarmiello, L. F.; Boccacci, P.; Luca, A. de; Piccirillo, P., Evaluation of 'Tonda di Giffoni' hazelnut (*Corylus avellana* L.) clones. *Sci. Hort.* **2010**, *124*, 2, 153–158.
- (15) European Commission. *eAmbrosia*, <https://ec.europa.eu/info/food-farming-fisheries/food-safety-and-quality/certification/quality-labels/geographical-indications-register/>, Zuletzt geprüft am: 15.09.2021.
- (16) Burdack-Freitag, A.; Schieberle, P., Characterization of the key odorants in raw Italian hazelnuts (*Corylus avellana* L. var. Tonda Romana) and roasted hazelnut paste by means of molecular sensory science. *J. Agr. Food Chem.* **2012**, *60*, 20, 5057–5064.
- (17) Kiefl, J.; Pollner, G.; Schieberle, P., Sensomics analysis of key hazelnut odorants (*Corylus avellana* L. 'Tonda Gentile') using comprehensive two-dimensional gas chromatography in combination with time-of-flight mass spectrometry (GC×GC-TOF-MS). *J. Agr. Food Chem.* **2013**, *61*, 22, 5226–5235.

- (18) Caligiani, A.; Coisson, J. D.; Travaglia, F.; Acquotti, D.; Palla, G.; Palla, L.; Arlorio, M., Application of ^1H NMR for the characterisation and authentication of "Tonda Gentile Trilobata" hazelnuts from Piedmont (Italy). *Food Chem.* **2014**, *148*, 77–85.
- (19) Bonvehí, J. S.; Coll, F. V., Oil content, stability and fatty acid composition of the main varieties of Catalanian Hazelnuts (*Corylus avellana* L.). *Food Chem.* **1993**, *48*, 3, 237–241.
- (20) Brufau, G.; Boatella, J.; Rafecas, M., Nuts: source of energy and macronutrients. *The British journal of nutrition* **2006**, *96 Suppl 2*, S24-8.
- (21) Andersen, G.; Souci, S.; Fachmann, W.; Kraut, H. *Lebensmitteltabelle für die Praxis: der kleine Souci-Fachmann-Kraut*, 5th edn.; Wiss. Verl.-Ges: Stuttgart, **2011**.
- (22) Parcerisa, J.; Boatella, J.; Codony, R.; Farrà, A.; Garcia, J.; Lopez, A.; Rafecas, M.; Romero, A., Influence of variety and geographical origin on the lipid fraction of hazelnuts (*Corylus avellana* L.) from Spain: I. Fatty acid composition. *Food Chem.* **1993**, *48*, 4, 411–414.
- (23) Parcerisa, J.; Richardson, D. G.; Rafecas, M.; Codony, R.; Boatella, J., Fatty acid, tocopherol and sterol content of some hazelnut varieties (*Corylus avellana* L.) harvested in Oregon (USA). *J. Chromatogr. A* **1998**, *805*, 1, 259–268.
- (24) Sabate, J.; Fraser, G. E.; Burke, K.; Knutsen, S. F.; Bennett, H.; Lindsted, K. D., Effects of Walnuts on Serum Lipid Levels and Blood Pressure in Normal Men. *N. Engl. J. Med.* **1993**, *328*, 9, 603–607.
- (25) Horrobin, D. F.; Manku, M. S., How do polyunsaturated fatty acids lower plasma cholesterol levels? *Lipids* **1983**, *18*, 8, 558–562.
- (26) Kris-Etherton, P. M., Monounsaturated fatty acids and risk of cardiovascular disease. *Circulation* **1999**, *100*, 11, 1253–1258.
- (27) Özdemir, M.; Açıkturk, F.; Kaplan, M.; Yıldız, M.; Löker, M.; Gürcan, T.; Biringen, G.; Okay, A.; Seyhan, F. G., Evaluation of new Turkish hybrid hazelnut (*Corylus avellana* L.) varieties: fatty acid composition, α -tocopherol content, mineral composition and stability. *Food Chem.* **2001**, *73*, 4, 411–415.
- (28) Boccacci, P.; Botta, R., Investigating the origin of hazelnut (*Corylus avellana* L.) cultivars using chloroplast microsatellites. *Genet. Resour. Crop. Evol.* **2009**, *56*, 6, 851–859.
- (29) FAO, <http://www.fao.org/faostat/en/#data>, Zuletzt geprüft am: 07.06.2019.
- (30) Statista GmbH. *Export- und Importmengen von Haselnüssen in Deutschland in den Jahren 2009 bis 2021*, <https://de.statista.com/statistik/daten/studie/1105040/umfrage/export-und-importmenge-haselnuessen/#statisticContainer>, Zuletzt geprüft am: 14.08.2022.
- (31) European Commission. *eAmbrosia*, <https://ec.europa.eu/info/food-farming-fisheries/food-safety-and-quality/certification/quality-labels/geographical-indications-register/>, Zuletzt geprüft am: 04.12.2022.
- (32) Federal Office of Food Safety and Consumer Protection. *OPSON Operationen*, https://www.bvl.bund.de/DE/01_Lebensmittel/03_Verbraucher/16_Food_Fraud/06_OPSON_Operationen/OPSON_Operationen_node.html, Zuletzt geprüft am: 10.10.2019.
- (33) BVL. *OPSON VI (2016/2017) - Haselnusserzeugnisse aus der Türkei, Georgien und Italien*, https://www.bvl.bund.de/DE/01_Lebensmittel/03_Verbraucher/16_Food_Fraud/06_OPSON_Operationen/OpsonVI/OPSON_Operationen_node.html, Zuletzt geprüft am: 09.06.2019.

- (34) Dettmer, K.; Aronov, P. A.; Hammock, B. D., Mass spectrometry-based metabolomics. *Mass Spectrom. Rev.* **2007**, *26*, 1, 51–78.
- (35) Felbinger, C.; Schelm, S.; Fischer, M., Food Fraud – Hindernisse und Lösungswege bei der Authentizitätsbestimmung von Lebensmitteln. *J. Verbr. Lebensm.* **2015**, *10*, S1, 25–30.
- (36) Schmidt, O. *Genetik und Molekularbiologie*; Springer Spektrum: Berlin, Heidelberg, **2017**.
- (37) Fischer, M.; Creydt, M.; Felbinger, C.; Fischer, C.; Klockmann, S.; Werner, P.; Klare, J.; Hünninger, T.; Hackl, T., Food Science Dialog 2014 @ HAW Hamburg: Lebensmittelsicherheit & Verpackung. *J. Verbr. Lebensm.* **2014**, *9*, 4, 391–420.
- (38) Esslinger, S.; Riedl, J.; Fauhl-Hassek, C., Potential and limitations of non-targeted fingerprinting for authentication of food in official control. *Food Res. Int.* **2014**, *60*, 189–204.
- (39) Neubauer, U., Interview: Foodomics. *analytica pro - Das offizielle Messemagazin 2022* **2022**, 54–55.
- (40) Pressestelle LGL - Bayerisches Landesamt für Gesundheit und Lebensmittelsicherheit. *LGL weist in sechs Proben nicht deklariertes Pferdefleisch im Spurenbereich nach*, **15.02.2013**.
- (41) Dr. Claus Schlicht. *Verfälschung oder Verunreinigung von Haselnussprodukten und Haselnusseis - Untersuchungsergebnisse 2017*,
https://www.lgl.bayern.de/lebensmittel/warengruppen/wc_23_huelsenfruechte_schalenobst/ue_2017_haselnuesse_verfaelschungen.htm, Zuletzt geprüft am: 08.05.2022.
- (42) Brüning, P.; Haase, I.; Matissek, R.; Fischer, M., Marzipan: polymerase chain reaction-driven methods for authenticity control. *J. Agric. Food Chem.* **2011**, *59*, 22, 11910–11917.
- (43) Herrmann, L.; Haase, I.; Blauhut, M.; Barz, N.; Fischer, M., DNA-based differentiation of the Ecuadorian cocoa types CCN-51 and Arriba based on sequence differences in the chloroplast genome. *J. Agric. Food Chem.* **2014**, *62*, 50, 12118–12127.
- (44) Wolf, C.; Burgener, M.; Hübner, P.; Lüthy, J., PCR-RFLP Analysis of Mitochondrial DNA: Differentiation of Fish Species. *LWT - Food Sci. Technol.* **2000**, *33*, 2, 144–150.
- (45) Kappel, K.; Eschbach, E.; Fischer, M.; Fritsche, J., Design of a user-friendly and rapid DNA microarray assay for the authentication of ten important food fish species. *Food Chem.* **2020**, *311*, 125884.
- (46) Reineke, A. *Gentechnik*, 1st edn.; Ulmer; UTB: Stuttgart, **2004**.
- (47) Weiler, E. W.; Nover, L. *Allgemeine und molekulare Botanik*; Georg Thieme Verlag: Stuttgart, **2008**.
- (48) Dolezel, J.; Bartos, J., Plant DNA flow cytometry and estimation of nuclear genome size. *Ann. Bot.* **2005**, *95*, 1, 99–110.
- (49) Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A., et al., The sequence of the human genome. *Science* **2001**, *291*, 5507, 1304–1351.
- (50) Martínez-García, P. J.; Crepeau, M. W.; Puiu, D.; Gonzalez-Ibeas, D.; Whalen, J.; Stevens, K. A.; Paul, R.; Butterfield, T. S.; Britton, M. T.; Reagan, R. L., et al., The walnut (*Juglans regia*) genome sequence reveals diversity in genes coding for the biosynthesis of non-structural polyphenols. *Plant J.* **2016**, *87*, 5, 507–532.

- (51) Verde, I.; Abbott, A. G.; Scalabrin, S.; Jung, S.; Shu, S.; Marroni, F.; Zhebentyayeva, T.; Dettori, M. T.; Grimwood, J.; Cattonaro, F., et al., The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat. Genet.* **2013**, *45*, 5, 487–494.
- (52) Storch, V.; Welsch, U.; Wink, M. *Evolutionsbiologie*, 3rd edn.; Springer Spektrum: Berlin, **2013**.
- (53) Wallace, H.; Birnstiel, M. L., Ribosomal cistrons and the nucleolar organizer. *Biochim. Biophys. Acta - Nucleic Acids Protein Synth.* **1966**, *114*, 2, 296–310.
- (54) Graw, J. *Genetik*, 6th edn.; Springer Spektrum: Berlin, **2015**.
- (55) Prokopowich, C. D.; Gregory, T. R.; Crease, T. J., The correlation between rDNA copy number and genome size in eukaryotes. *Genome* **2003**, *46*, 1, 48–50.
- (56) Rogers, S. O.; Bendich, A. J., Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. *Plant Mol. Biol.* **1987**, *9*, 509–520.
- (57) Rogers, S. O.; Bendich, A. J., Heritability and Variability in Ribosomal RNA Genes of *Vicia faba*. *Genetics* **1987**, *117* (2), 285–295.
- (58) Baldwin, B. G.; Sanderson, M. J.; Porter, J. M.; Wojciechowski, M. F.; Campbell, C. S.; Donoghue, M. J., The ITS Region of Nuclear Ribosomal DNA: A Valuable Source of Evidence on Angiosperm Phylogeny. *Ann. Mo. Bot. Gard.* **1995**, *82*, 2, 247.
- (59) Hillis, D. M.; Dixon, M. T., Ribosomal DNA: Molecular Evolution and Phylogenetic Inference. *Q. Rev. Biol.* **1991**, *66* (4), 411–453.
- (60) Álvarez, I.; Wendel, J. F., Ribosomal ITS sequences and plant phylogenetic inference. *Mol. Phylogenet. Evol.* **2003**, *29*, 3, 417–434.
- (61) Schoch, C. L.; Seifert, K. A.; Huhndorf, S.; Robert, V.; Spouge, J. L.; Levesque, C. A.; Chen, W., Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 16, 6241–6246.
- (62) Chu, K. H.; Li, C. P.; Ho, H. Y., The first internal transcribed spacer (ITS-1) of ribosomal DNA as a molecular marker for phylogenetic and population analyses in crustacea. *Mar. Biotechnol.* **2001**, *3*, 4, 355–361.
- (63) Leaw, S. N.; Chang, H. C.; Sun, H. F.; Barton, R.; Bouchara, J.-P.; Chang, T. C., Identification of medically important yeast species by sequence analysis of the internal transcribed spacer regions. *J. Clin. Microbiol.* **2006**, *44*, 3, 693–699.
- (64) El Karkouri, K.; Murat, C.; Zampieri, E.; Bonfante, P., Identification of internal transcribed spacer sequence motifs in truffles. *Appl. Environ. Microbiol.* **2007**, *73*, 16, 5320–5330.
- (65) Kim, O. T.; Bang, K. H.; In, D. S.; Lee, J. W.; Kim, Y. C.; Shin, Y. S.; Hyun, D. Y.; Lee, S. S.; Cha, S. W.; Seong, N. S., Molecular authentication of ginseng cultivars by comparison of internal transcribed spacer and 5.8S rDNA sequences. *Plant Biotechnol. Rep.* **2007**, *1*, 3, 163–167.
- (66) Zhang, Y.; Wang, L., The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evol. Biol.* **2005**, *5*, 1.
- (67) Eulgem, T.; Rushton, P. J.; Robatzek, S.; Somssich, I. E., The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* **2000**, *5*, 5, 199–206.
- (68) Piasecka, A.; Kachlicki, P.; Stobiecki, M., Analytical Methods for Detection of Plant Metabolomes Changes in Response to Biotic and Abiotic Stresses. *Int. J. Mol. Sci.* **2019**, *20*, 2.

- (69) Borrone, J. W.; Kuhn, D. N.; Schnell, R. J., Isolation, characterization, and development of WRKY genes as useful genetic markers in *Theobroma cacao*. *Theor. Appl. Genet.* **2004**, *109*, 3, 495–507.
- (70) Ishiguro, S.; Nakamura, K., Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and beta-amylase from sweet potato. *Mol. Gen. Genet.* **1994**, *244*, 6, 563–571.
- (71) Rushton, P. J.; Torres, J. T.; Parniske, M.; Wernert, P.; Hahlbrock, K.; Somssich, I. E., Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes. *EMBO J.* **1996**, *15*, 20, 5690–5700.
- (72) Pater, S. de; Greco, V.; Pham, K.; Memelink, J.; Kijne, J., Characterization of a zinc-dependent transcriptional activator from *Arabidopsis*. *Nucleic Acids Res.* **1996**, *24*, 23, 4624–4631.
- (73) Rushton, P. J.; Macdonald, H.; Huttly, A. K.; Lazarus, C. M.; Hooley, R., Members of a new family of DNA-binding proteins bind to a conserved cis-element in the promoters of alpha-Amy2 genes. *Plant Mol. Biol.* **1995**, *29*, 4, 691–702.
- (74) Rushton, P. J.; Somssich, I. E.; Ringler, P.; Shen, Q. J., WRKY transcription factors. *Trends Plant Sci.* **2010**, *15*, 5, 247–258.
- (75) Yamasaki, K.; Kigawa, T.; Inoue, M.; Tateno, M.; Yamasaki, T.; Yabuki, T.; Aoki, M.; Seki, E.; Matsuda, T.; Tomo, Y., et al., Solution structure of an *Arabidopsis* WRKY DNA binding domain. *Plant Cell* **2005**, *17*, 3, 944–956.
- (76) Duan, M.-R.; Nan, J.; Liang, Y.-H.; Mao, P.; Lu, L.; Li, L.; Wei, C.; Lai, L.; Li, Y.; Su, X.-D., DNA binding mechanism revealed by high resolution crystal structure of *Arabidopsis thaliana* WRKY1 protein. *Nucleic Acids Res.* **2007**, *35*, 4, 1145–1154.
- (77) Dong, J.; Chen, C.; Chen, Z., Expression profiles of the *Arabidopsis* WRKY gene superfamily during plant defense response. *Plant Mol. Biol.* **2003**, *51*, 1, 21–37.
- (78) Rushton, P. J.; Somssich, I. E., Transcriptional control of plant genes responsive to pathogens. *Curr. Opin. Plant Biol.* **1998**, *1*, 4, 311–315.
- (79) Li, W.; Ruf, S.; Bock, R., Constancy of organellar genome copy numbers during leaf development and senescence in higher plants. *Mol. Genet. Genomics* **2006**, *275*, 2, 185–192.
- (80) Wang, W.; Lanfear, R., Long-Reads Reveal That the Chloroplast Genome Exists in Two Distinct Versions in Most Plants. *Genome Biol. Evol.* **2019**, *11*, 12, 3372–3381.
- (81) Powikrowska, M.; Oetke, S.; Jensen, P. E.; Krupinska, K., Dynamic composition, shaping and organization of plastid nucleoids. *Front. Plant Sci.* **2014**, *5*, 424.
- (82) Rauwolf, U.; Golczyk, H.; Greiner, S.; Herrmann, R. G., Variable amounts of DNA related to the size of chloroplasts III. Biochemical determinations of DNA amounts per organelle. *Mol. Genet. Genomics* **2010**, *283*, 1, 35–47.
- (83) Liere, K.; Börner T., Development-Dependent Changes in the Amount and Structural Organization of Plastid DNA. *Plastid Development in Leaves during Growth and Senescence* **2013**, 36.
- (84) Boffey, S. A.; Leech, R. M., Chloroplast DNA Levels and the Control of Chloroplast Division in Light-Grown Wheat Leaves. *Plant Physiol.* **1982**, *69*, 1387–1391.
- (85) Baumgartner, B. J.; Rapp, J. C.; Mullet, J. E., Plastid Transcription Activity and DNA Copy Number Increase Early in Barley Chloroplast Development. *Plant Physiol.* **1989**, 89.

- (86) Kuroiwa, T.; Suzuki, T.; Ogawa, K.; Kawano, S., The Chloroplast Nucleus: Distribution, Number, Size, and Shape, and a Model for the Multiplication of the Chloroplast Genome during Chloroplast Development. *Plant Cell Physiol.* **1981**, *22*, 3, 381–396.
- (87) Oldenburg, D. J.; Bendich, A. J., Changes in the structure of DNA molecules and the amount of DNA per plastid during chloroplast development in maize. *J. Mol. Biol.* **2004**, *344*, 5, 1311–1330.
- (88) Golczyk, H.; Greiner, S.; Wanner, G.; Weihe, A.; Bock, R.; Börner, T.; Herrmann, R. G., Chloroplast DNA in mature and senescing leaves: a reappraisal. *Plant Cell* **2014**, *26*, 3, 847–854.
- (89) Birky, C. W., Uniparental inheritance of organelle genes. *Curr. Biol.* **2008**, *18*, 16, R692-5.
- (90) Weihe, A.; Apitz, J.; Pohlheim, F.; Salinas-Hartwig, A.; Börner, T., Biparental inheritance of plastidial and mitochondrial DNA and hybrid variegation in Pelargonium. *Mol. Genet. Genomics* **2009**, *282*, 6, 587–593.
- (91) Vendramin, G. G.; Ziegenhagen, B., Characterisation and inheritance of polymorphic plastid microsatellites in Abies. *Genome* **1997**, *40*, 6, 857–864.
- (92) Dumolin, S.; Demesure, B.; Petit, R. J., Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. *Theoretical and Applied Genetics* **1995**, *91*, 8, 1253–1256.
- (93) Wolfe, K. H.; Li, W. H.; Sharp, P. M., Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 24, 9054–9058.
- (94) Hu, G.; Cheng, L.; Lan, Y.; Cao, Q.; Huang, W., The complete chloroplast genome sequence of *Corylus chinensis* Franch. *Conservation Genet. Resour.* **2016**.
- (95) Kahraman, K.; Lucas, S. J., Comparison of different annotation tools for characterization of the complete chloroplast genome of *Corylus avellana* cv Tombul. *BMC Genomics* **2019**, *20*, 1, 874.
- (96) Saiki, R. K.; Scharf, S.; Faloona, F.; Mullis, K. B.; Horn, G. T.; Erlich, H. A.; Arnheim, N., Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **1985**, *230*, 4732, 1350–1354.
- (97) The Nobel Foundation. *The Nobel Prize in Chemistry 1993*, <https://www.nobelprize.org/prizes/chemistry/1993/mullis/facts/>, Zuletzt geprüft am: 17.04.2022.
- (98) Mullis, K.; Faloona, F.; Scharf, S.; Saiki, R.; Horn, G.; Erlich, H., Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor symposia on quantitative biology* **1986**, *51 Pt 1*, 263–273.
- (99) Baumforth, K. R.; Nelson, P. N.; Digby, J. E.; O'Neil, J. D.; Murray, P. G., Demystified ... the polymerase chain reaction. *Mol. Pathol.* **1999**, *52*, 1, 1–10.
- (100) Müller, H.-J. *PCR - Polymerase-Kettenreaktion*, 2nd edn.; Springer Spektrum: Berlin, Heidelberg, **2016**.
- (101) Ahsen, N. von; Wittwer, C. T.; Schütz, E., Oligonucleotide Melting Temperatures under PCR Conditions: Nearest-Neighbor Corrections for Mg²⁺, Deoxynucleotide Triphosphate, and Dimethyl Sulfoxide Concentrations with Comparison to Alternative Empirical Formulas. *Clin. Chem.* **2001**, *47*, 11, 1956–1961.
- (102) Matissek, R.; Fischer, M. *Lebensmittelanalytik*, 7th edn.; Springer Spektrum: Berlin, Germany, **2021**.

- (103) Mullis, K. B., The unusual origin of the polymerase chain reaction. *Sci. Am.* **1990**, *262*, 4, 56-61, 64-5.
- (104) Saiki, R. K.; Gelfand, D. H.; Stoffel, S.; Scharf, S. J.; Higuchi, R.; Horn, G. T.; Mullis, K. B.; Erlich, H. A., Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **1988**, *239*, 4839, 487–491.
- (105) Brock, T. D.; Freeze, H., *Thermus aquaticus* gen. n. and sp. n., a nonsporulating extreme thermophile. *J. Bacteriol.* **1969**, *98*, 1, 289–297.
- (106) Tiselius, A., A new apparatus for electrophoretic analysis of colloidal mixtures. *Trans. Faraday Soc.* **1937**, *33*, 524.
- (107) Gey, M. *Instrumentelle Analytik und Bioanalytik*, 3rd edn.; Springer Spektrum: Berlin, **2015**.
- (108) Berg, J. M.; Tymoczko, J. L.; Gatto jr., G. J.; Stryer, L. *Stryer Biochemie*, 8th edn.; Springer Spektrum: Berlin, **2018**.
- (109) Hjertén, S., Agarose as an anticonvection agent in zone electrophoresis. *Biochim. Biophys. Acta* **1961**, *53*, 3, 514–517.
- (110) Araki, C., Structure of the Agarose Constituent of Agar-agar. *Bull. Chem. Soc. Jpn.* **1956**, *29*, 4, 543–544.
- (111) Mao, F.; Leung, W.-Y. *Methods of using dyes in association with nucleic acid staining or detection and associated technology*, **11.08.2010**.
- (112) Aaij, C.; Borst, P., The gel electrophoresis of DNA. *Biochim. Biophys. Acta - Nucleic Acids Protein Synth.* **1972**, *269*, 2, 192–200.
- (113) Sharp, P. A.; Sugden, B.; Sambrook, J., Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose-ethidium bromide electrophoresis. *Biochemistry* **1973**, *12*, 16, 3055–3063.
- (114) Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung (IFA). *GESTIS Stoffdatenbank*, www.dguv.de/ifa/stoffdatenbank.
- (115) Guzaev, M.; Li, X.; Park, C.; Leung, W.-Y.; Roberts, L. *Comparison of Nucleic Acid Gel Stains*, <https://biotium.com/wp-content/uploads/2017/02/Gel-Stains-Comparison.pdf>, Zuletzt geprüft am: 29.09.2020.
- (116) Biotium, Inc. *Safety Report for GelRed® and GelGreen®*, <https://biotium.com/wp-content/uploads/2013/07/GelRed-and-GelGreen-Safety-Report.pdf>, Zuletzt geprüft am: 29.09.2020.
- (117) Woolley, A. T.; Mathies, R. A., Ultra-high-speed DNA fragment separations using microfabricated capillary array electrophoresis chips. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 24, 11348–11352.
- (118) Nachamkin, I.; Panaro, N. J.; Li, M.; Ung, H.; Yuen, P. K.; Kricka, L. J.; Wilding, P., Agilent 2100 bioanalyzer for restriction fragment length polymorphism analysis of the *Campylobacter jejuni* flagellin gene. *J. Clin. Microbiol.* **2001**, *39*, 2, 754–757.
- (119) Chen, L.; Ren, J., High-throughput DNA analysis by microchip electrophoresis. *Comb. Chem. High Throughput Screen.* **2004**, *7*, 1, 29–43.
- (120) Agilent Technologies, Inc. 2009-2016, 2017, Agilent High Sensitivity DNA Kit Quick Start Guide **2017**.

- (121) Agilent Technologies, Inc. 2002-2018, Agilent DNA 7500 and DNA 12000 Kit Quick Start Guide **2018**.
- (122) Mueller, O.; Hahnenberger, K.; Dittmann, M.; Yee, H.; Dubrow, R.; Nagle, R.; Ilsley, D., A microfluidic system for high-speed reproducible DNA sizing and quantitation. *Electrophoresis* **2000**, *21*, 1, 128–134.
- (123) Agilent Technologies, Inc., Agilent 2100 Bioanalyzer System, Agilent 4200 TapeStation System **2018**.
- (124) Sanger, F.; Coulson, A. R., A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* **1975**, *94*, 3, 441–448.
- (125) Sanger, F.; Air, G. M.; Barrell, B. G.; Brown, N. L.; Coulson, A. R.; Fiddes, C. A.; Hutchison, C. A.; Slocombe, P. M.; Smith, M., Nucleotide sequence of bacteriophage phi X174 DNA. *Nature* **1977**, *265*, 5596, 687–695.
- (126) Sanger, F.; Nicklen, S.; Coulson, A. R., DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 12, 5463–5467.
- (127) Atkinson, M. R.; Deutscher, M. P.; Kornberg, A.; Russell, A. F.; Moffatt, J. G., Enzymatic synthesis of deoxyribonucleic acid. XXXIV. Termination of chain growth by a 2',3'-dideoxyribonucleotide. *Biochemistry* **1969**, *8*, 12, 4897–4904.
- (128) Valencia, C. A.; Pervaiz, M. A.; Husami, A.; Qian, Y.; Zhang, K., Sanger Sequencing Principles, History, and Landmarks. In *Next Generation Sequencing Technologies in Medical Genetics*; Valencia, C. Alexander; Pervaiz, M. Ali; Husami, Ammar; Qian, Yaping; Zhang, Kejian, Eds.; Springer New York: New York, NY, **2013**.
- (129) Valencia, C. A.; Pervaiz, M. A.; Husami, A.; Qian, Y.; Zhang, K., Eds. *Next Generation Sequencing Technologies in Medical Genetics*; Springer New York: New York, NY, **2013**.
- (130) Ju, J.; Glazer, A. N.; Mathies, R. A., Energy transfer primers: a new fluorescence labeling paradigm for DNA sequencing and analysis. *Nat. Med.* **1996**, *2*, 2, 246–249.
- (131) Schuster, S. C., Next-generation sequencing transforms today's biology. *Nat. Methods* **2008**, *5*, 1, 16–18.
- (132) Eurofins Genomics Germany GmbH. *SupremeRun Tube - Supreme Sanger Sequencing*, <https://eurofinsgenomics.eu/en/custom-dna-sequencing/gatc-services/supremerun-tube/>, Zuletzt geprüft am: 16.04.2022.
- (133) Metzker, M. L., Sequencing technologies - the next generation. *Nat Rev Genet* **2010**, *11*, 1, 31–46.
- (134) McCombie, W. R.; McPherson, J. D.; Mardis, E. R., Next-Generation Sequencing Technologies. *Cold Spring Harbor perspectives in medicine* **2019**, *9*, 11.
- (135) Anderson, S., Shotgun DNA sequencing using cloned DNase I-generated fragments. *Nucleic Acids Res.* **1981**, *9*, 13, 3015–3027.
- (136) Niedringhaus, T. P.; Milanova, D.; Kerby, M. B.; Snyder, M. P.; Barron, A. E., Landscape of next-generation sequencing technologies. *Anal. Chem.* **2011**, *83*, 12, 4327–4341.
- (137) Bentley, D. R.; Balasubramanian, S.; Swerdlow, H. P.; Smith, G. P.; Milton, J.; Brown, C. G.; Hall, K. P.; Evers, D. J.; Barnes, C. L.; Bignell, H. R., et al., Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **2008**, *456*, 7218, 53–59.

- (138) Stranneheim, H.; Lundeberg, J., Stepping stones in DNA sequencing. *Biotechnology journal* **2012**, *7*, 9, 1063–1073.
- (139) Buermans, H. P. J.; Dunnen, J. T. den, Next generation sequencing technology: Advances and applications. *Biochim. Biophys. Acta* **2014**, *1842*, 10, 1932–1941.
- (140) Goodwin, S.; McPherson, J. D.; McCombie, W. R., Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* **2016**, *17*, 6, 333–351.
- (141) Rothberg, J. M.; Hinz, W.; Rearick, T. M.; Schultz, J.; Mileski, W.; Davey, M.; Leamon, J. H.; Johnson, K.; Milgrew, M. J.; Edwards, M., et al., An integrated semiconductor device enabling non-optical genome sequencing. *Nature* **2011**, *475*, 7356, 348–352.
- (142) Eid, J.; Fehr, A.; Gray, J.; Luong, K.; Lyle, J.; Otto, G.; Peluso, P.; Rank, D.; Baybayan, P.; Bettman, B., et al., Real-time DNA sequencing from single polymerase molecules. *Science* **2009**, *323*, 5910, 133–138.
- (143) Levene, M. J.; Korlach, J.; Turner, S. W.; Foquet, M.; Craighead, H. G.; Webb, W. W., Zero-mode waveguides for single-molecule analysis at high concentrations. *Science* **2003**, *299*, 5607, 682–686.
- (144) Branton, D.; Deamer, D. W.; Marziali, A.; Bayley, H.; Benner, S. A.; Butler, T.; Di Ventra, M.; Garaj, S.; Hibbs, A.; Huang, X., et al., The potential and challenges of nanopore sequencing. *Nature biotechnology* **2008**, *26*, 10, 1146–1153.
- (145) Ku, C.-S.; Roukos, D. H., From next-generation sequencing to nanopore sequencing technology: paving the way to personalized genomic medicine. *Expert review of medical devices* **2013**, *10*, 1, 1–6.
- (146) Jain, M.; Olsen, H. E.; Paten, B.; Akeson, M., The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome biology* **2016**, *17*, 1, 239.
- (147) Oxford Nanopore Technologies. *MinION*, <https://nanoporetech.com/sites/default/files/s3/literature/MinION-Mk1C-brochure.pdf>, Zuletzt geprüft am: 17.11.2022.
- (148) Donoso, A.; Valenzuela, S., In-field molecular diagnosis of plant pathogens: recent trends and future perspectives. *Plant Pathol* **2018**, *67*, 7, 1451–1461.
- (149) Escalona, M.; Rocha, S.; Posada, D., A comparison of tools for the simulation of genomic next-generation sequencing data. *Nat Rev Genet* **2016**, *17*, 8, 459–469.
- (150) Pingoud, A.; Jeltsch, A., Structure and function of type II restriction endonucleases. *Nucleic Acids Res.* **2001**, *29*, 18, 3705–3727.
- (151) Holton, T. A.; Graham, M. W., A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. *Nucleic Acids Res.* **1991**, *19*, 5, 1156.
- (152) Clark, J. M., Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* **1988**, *16*, 20, 9677–9686.
- (153) Marchuk, D.; Drumm, M.; Saulino, A.; Collins, F. S., Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res.* **1991**, *19*, 5, 1154.
- (154) Shuman, S.; Prescott, J., Specific DNA cleavage and binding by vaccinia virus DNA topoisomerase I. *J. Biol. Chem.* **1990**, *265*, 29, 17826–17836.

- (155) Aune, T. E. V.; Aachmann, F. L., Methodologies to increase the transformation efficiencies and the range of bacteria that can be transformed. *Appl. Microbiol. Biotechnol.* **2010**, *85*, 5, 1301–1313.
- (156) Mandel, M.; Higa, A., Calcium-dependent bacteriophage DNA infection. *J. Mol. Biol.* **1970**, *53*, 1, 159–162.
- (157) Cohen, S. N.; Chang, A. C.; Hsu, L., Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc. Natl. Acad. Sci. USA* **1972**, *69*, 8, 2110–2114.
- (158) Cohen, S. N.; Chang, A. C.; Boyer, H. W.; Helling, R. B., Construction of biologically functional bacterial plasmids in vitro. *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 11, 3240–3244.
- (159) Pope, B.; Kent, H. M., High efficiency 5 min transformation of *Escherichia coli*. *Nucleic Acids Res.* **1996**, *24*, 3, 536–537.
- (160) Neumann, E.; Schaefer-Ridder, M.; Wang, Y.; Hofschneider, P. H., Gene transfer into mouse lyoma cells by electroporation in high electric fields. *EMBO J.* **1982**, *1*, 7, 841–845.
- (161) Shi, J.; Ma, Y.; Zhu, J.; Chen, Y.; Sun, Y.; Yao, Y.; Yang, Z.; Xie, J., A Review on Electroporation-Based Intracellular Delivery. *Molecules* **2018**, *23*, 11.
- (162) Dallas-Yang, Q.; Jiang, G.; Sladek, F. M., Avoiding false positives in colony PCR. *BioTechniques* **1998**, *24*, 4, 580–582.
- (163) Collard, B. C. Y.; Mackill, D. J., Start Codon Targeted (SCoT) Polymorphism: A Simple, Novel DNA Marker Technique for Generating Gene-Targeted Markers in Plants. *Plant Mol. Biol. Rep.* **2009**, *27*, 1, 86–93.
- (164) Boujard, D.; Anselme, B.; Cullin, C.; Raguénès-Nicol, C. *Zell- und Molekularbiologie im Überblick*; Springer Spektrum: Berlin, **2014**.
- (165) Kozak, M., An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **1987**, *15*, 20, 8125–8148.
- (166) Joshi, C. P.; Zhou, H.; Huang, X.; Chiang, V. L., Context sequences of translation initiation codon in plants. *Plant Mol. Biol.* **1997**, *35*, 6, 993–1001.
- (167) Sawant, S. V.; Singh, P. K.; Gupta, S. K.; Madnala, R.; Tuli, R., Conserved nucleotide sequences in highly expressed genes in plants. *J. Genet.* **1999**, *78*, 2, 123–131.
- (168) Gillings, M.; Holley, M., Amplification of anonymous DNA fragments using pairs of long primers generates reproducible DNA fingerprints that are sensitive to genetic variation. *Electrophoresis* **1997**, *18*, 9, 1512–1518.
- (169) QiaoQiao, J.; GuiYou, L.; WuWen, L.; ZiNiu, D., Identification of genetic variation in *Citrus sinensis* from Hunan based on Start Codon Targeted Polymorphism. *Hunan Agric. Sci. Technol.* **2011**, *12* (11), 1594–1599.
- (170) Luo, C.; He, X.; Chen, H.; Ou, S.; Gao, M.; Brown, J. S.; Tondo, C. T.; Schnell, R. J., Genetic diversity of mango cultivars estimated using SCoT and ISSR markers. *Biochem. Syst. Ecol.* **2011**, *39*, 4, 676–684.
- (171) Gorji, A. M.; Poczai, P.; Polgar, Z.; Taller, J., Efficiency of Arbitrarily Amplified Dominant Markers (SCoT, ISSR and RAPD) for Diagnostic Fingerprinting in Tetraploid Potato. *Am. J. Pot. Res.* **2011**, *88*, 3, 226–237.

- (172) Mahjbi, A.; Baraket, G.; Oueslati, A.; Salhi-Hannachi, A., Start Codon Targeted (SCoT) markers provide new insights into the genetic diversity analysis and characterization of Tunisian Citrus species. *Biochem. Syst. Ecol.* **2015**, *61*, 390–398.
- (173) Paran, I.; Michelmore, R. W., Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theor. Appl. Genet.* **1993**, *85*, 8, 985–993.
- (174) Kiran, U.; Khan, S.; Mirza, K. J.; Ram, M.; Abdin, M. Z., SCAR markers: a potential tool for authentication of herbal drugs. *Fitoterapia* **2010**, *81*, 8, 969–976.
- (175) Bhagyawant, S. S., RAPD-SCAR Markers: An Interface Tool for Authentication of Traits. *JBM* **2016**, *04*, 01, 1–9.
- (176) Williams, J. G.; Kubelik, A. R.; Livak, K. J.; Rafalski, J. A.; Tingey, S. V., DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* **1990**, *18*, 22, 6531–6535.
- (177) Welsh, J.; McClelland, M., Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* **1990**, *18*, 24, 7213–7218.
- (178) Parasnis, A. S.; Gupta, V. S.; Tamhankar, S. A.; Ranjekar, P. K., A highly reliable sex diagnostic PCR assay for mass screening of papaya seedlings. *Mol. Breed.* **2000**, *6*, 3, 337–344.
- (179) Marieschi, M.; Torelli, A.; Bruni, R., Quality control of saffron (*Crocus sativus* L.): development of SCAR markers for the detection of plant adulterants used as bulking agents. *J. Agric. Food Chem.* **2012**, 10998–11004.
- (180) Choi, Y.-E.; Ahn, C. H.; Kim, B.-B.; Yoon, E.-S., Development of species specific AFLP-derived SCAR marker for authentication of *Panax japonicus* C. A. MEYER. *Biol. Pharm. Bull.* **2008**, *31*, 1, 135–138.
- (181) Julio, E.; Verrier, J.-L.; Dorlhac de Borne, F., Development of SCAR markers linked to three disease resistances based on AFLP within *Nicotiana tabacum* L. *Theor. Appl. Genet.* **2006**, *112*, 2, 335–346.
- (182) Lahogue, F.; This, P.; Bouquet, A., Identification of a codominant scar marker linked to the seedlessness character in grapevine. *Theoretical and Applied Genetics* **1998**, *97*, 5-6, 950–959.
- (183) Ohmori, T.; Murata, M.; Motoyoshi, F., Molecular characterization of RAPD and SCAR markers linked to the *Tm-1* locus in tomato. *Theor. Appl. Genet.* **1996**, *92*, 2, 151–156.
- (184) Torelli, A.; Marieschi, M.; Bruni, R., Authentication of saffron (*Crocus sativus* L.) in different processed, retail products by means of SCAR markers. *Food Control* **2014**, *36*, 1, 126–131.
- (185) Marieschi, M.; Torelli, A.; Beghé, D.; Bruni, R., Authentication of *Punica granatum* L.: Development of SCAR markers for the detection of 10 fruits potentially used in economically motivated adulteration. *Food Chem.* **2016**, 438–444.
- (186) Wu, Y.; Zhang, Z.; Chen, Y.; Wang, B.; Yang, G.; Yang, W., Authentication of Thailand jasmine rice using RAPD and SCAR methods. *Eur. Food Res. Technol.* **2009**, *229*, 3, 515–521.
- (187) Arber, W., Promotion and limitation of genetic exchange. *Science* **1979**, *205*, 4404, 361–365.
- (188) Williams, R. J., Restriction Endonucleases: Classification, Properties, and Applications. *Mol. Biotechnol.* **2003**, *23*, 3, 225–244.
- (189) Smith, H. O.; Nathans, D., A Suggested Nomenclature for Bacterial Host Modification and Restriction Systems and their Enzymes. *J. Mol. Biol.* **1973**, *81* (3).

- (190) Jeltsch, A.; Pingoud, A., Kinetic characterization of linear diffusion of the restriction endonuclease EcoRV on DNA. *Biochemistry* **1998**, *37*, 8, 2160–2169.
- (191) Halden, N. F.; Wolf, J. B.; Leonard, W. J., Identification of a novel site specific endonuclease produced by *Mycoplasma fermentans*: discovery while characterizing DNA binding proteins in T lymphocyte cell lines. *Nucleic Acids Res.* **1989**, *17*, 9, 3491–3499.
- (192) Schelm, S.; Siemt, M.; Pfeiffer, J.; Lang, C.; Tichy, H.-V.; Fischer, M., Food Authentication: Identification and Quantitation of Different Tuber Species via Capillary Gel Electrophoresis and Real-Time PCR. *Foods* **2020**, *9*, 4.
- (193) Spaniolas, S.; May, S. T.; Bennett, M. J.; Tucker, G. A., Authentication of coffee by means of PCR-RFLP analysis and lab-on-a-chip capillary electrophoresis. *J. Agric. Food Chem.* **2006**, *54*, 20, 7466–7470.
- (194) Haider, N.; Nabulsi, I.; Al-Safadi, B., Identification of meat species by PCR-RFLP of the mitochondrial COI gene. *Meat Sci.* **2012**, *90*, 2, 490–493.
- (195) Ririe, K. M.; Rasmussen, R. P.; Wittwer, C. T., Product differentiation by analysis of DNA Melting Curves during the Polymerase Chain Reaction. *Anal. Biochem.* **1997**, *245*, 154–160.
- (196) Lay, M. J.; Wittwer, C. T., Real-time fluorescence genotyping of factor V Leiden during rapid-cycle PCR. *Clin. Chem.* **1997**, *43*, 12, 2262–2267.
- (197) Simko, I., High-Resolution DNA Melting Analysis in Plant Research. *Trends Plant Sci.* **2016**, *21*, 6, 528–537.
- (198) Ganopoulos, I.; Argiriou, A.; Tsaftaris, A., Adulterations in Basmati rice detected quantitatively by combined use of microsatellite and fragrance typing with High Resolution Melting (HRM) analysis. *Food Chem.* **2011**, *129*, 2, 652–659.
- (199) Vossen, R. H. A. M.; Aten, E.; Roos, A.; Dunnen, J. T. den, High-resolution melting analysis (HRMA): more than just sequence variant screening. *Hum. Mutat.* **2009**, *30*, 6, 860–866.
- (200) Druml, B.; Cichna-Markl, M., High resolution melting (HRM) analysis of DNA--its role and potential in food analysis. *Food Chem.* **2014**, *158*, 245–254.
- (201) Wittwer, C. T., High-resolution DNA melting analysis: advancements and limitations. *Hum. Mutat.* **2009**, *30*, 6, 857–859.
- (202) Wittwer, C. T.; Reed, G. H.; Gundry, C. N.; Vandersteen, J. G.; Pryor, R. J., High-resolution genotyping by amplicon melting analysis using LCGreen. *Clin. Chem.* **2003**, *49*, 6 Pt 1, 853–860.
- (203) Reed, G. H.; Wittwer, C. T., Sensitivity and specificity of single-nucleotide polymorphism scanning by high-resolution melting analysis. *Clin. Chem.* **2004**, *50*, 10, 1748–1754.
- (204) Mackay, J. F.; Wright, C. D.; Bonfiglioli, R. G., A new approach to varietal identification in plants by microsatellite high resolution melting analysis: application to the verification of grapevine and olive cultivars. *Plant Methods* **2008**, *4*, 8.
- (205) Distefano, G.; La Malfa, S.; Gentile, A.; Wu, S.-B., EST-SNP genotyping of citrus species using high-resolution melting curve analysis. *Tree Genet. Genomes* **2013**, *9*, 5, 1271–1281.
- (206) Wu, S.-B.; Wirthensohn, M. G.; Hunt, P.; Gibson, J. P.; Sedgley, M., High resolution melting analysis of almond SNPs derived from ESTs. *Theor. Appl. Genet.* **2008**, *118*, 1, 1–14.

- (207) Garritano, S.; Gemignani, F.; Voegelé, C.; Nguyen-Dumont, T.; Le Calvez-Kelm, F.; Silva, D. de; Lesueur, F.; Landi, S.; Tavtigian, S. V., Determining the effectiveness of High Resolution Melting analysis for SNP genotyping and mutation scanning at the TP53 locus. *BMC Genet.* **2009**, *10*, 5.
- (208) Erali, M.; Voelkerding, K. V.; Wittwer, C. T., High resolution melting applications for clinical laboratory medicine. *Exp. Mol. Pathol.* **2008**, *85*, 1, 50–58.
- (209) Montgomery, J. L.; Sanford, L. N.; Wittwer, C. T., High-resolution DNA melting analysis in clinical research and diagnostics. *Expert Rev. Mol. Diagn.* **2010**, *10*, 2, 219–240.
- (210) Lang, C.; Weber, N.; Möller, M.; Schramm, L.; Schelm, S.; Kohlbacher, O.; Fischer, M., Genetic authentication: Differentiation of hazelnut cultivars using polymorphic sites of the chloroplast genome. *Food Control* **2021**, *130*, 108344.
- (211) Illumina, Inc. *Coverage depth recommendations*, <https://www.illumina.com/science/technology/next-generation-sequencing/plan-experiments/coverage.html>, Zuletzt geprüft am: 09.10.2022.
- (212) Kane, N.; Sveinsson, S.; Dempewolf, H.; Yang, J. Y.; Zhang, D.; Engels, J. M. M.; Cronk, Q., Ultra-barcoding in cacao (*Theobroma* spp.; Malvaceae) using whole chloroplast genomes and nuclear ribosomal DNA. *Am. J. Bot.* **2012**, *99*, 2, 320–329.
- (213) Nock, C. J.; Waters, D. L. E.; Edwards, M. A.; Bowen, S. G.; Rice, N.; Cordeiro, G. M.; Henry, R. J., Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnol. J.* **2011**, *9*, 3, 328–333.
- (214) Meyers, S. C.; Liston, A., Characterizing the genome of wild relatives of *Limnanthes alba* (Meadowfoam) using massively parallel sequencing. *Acta Hort.* **2010**, 859, 309–314.
- (215) Yang, J.-B.; Li, D.-Z.; Li, H.-T., Highly effective sequencing whole chloroplast genomes of angiosperms by nine novel universal primer pairs. *Mol. Ecol. Resour.* **2014**, *14*, 5, 1024–1031.
- (216) Li, Y.; Bi, H.; Liu, B.; Guo, X.; Hao, G.; He, Q.; Ma, T., The complete chloroplast genome of *Ostrya rehderiana*. *Mitochondrial DNA* **2015**, 1–2.
- (217) Hu, G.; Cheng, L.; Lan, Y.; Cao, Q.; Wang, X.; Huang, W., The complete chloroplast genome sequence of the endangered Chinese endemic tree *Corylus fargesii*. *Conservation Genet. Resour.* **2016**.
- (218) Gürçan, K.; Mehlenbacher, S. A.; Botta, R.; Boccacci, P., Development, characterization, segregation, and mapping of microsatellite markers for European hazelnut (*Corylus avellana* L.) from enriched genomic libraries and usefulness in genetic diversity studies. *Tree Genet. Genomes* **2010**, *6*, 4, 513–531.
- (219) Jack, R. S.; Eggert, H., Restriction enzymes have limited access to DNA sequences in *Drosophila* chromosomes. *EMBO J.* **1990**, *9* (8), 2603–2609.
- (220) Ram, J. L.; Ram, M. L.; Baidoun, F. F., Authentication of Canned Tuna and Bonito by Sequence and Restriction Site Analysis of Polymerase Chain Reaction Products of Mitochondrial DNA. *J. Agric. Food Chem.* **1996**, *44*, 8, 2460–2467.
- (221) Meyer, R.; Höfelein, C.; Lüthy, J.; Candrian, U., Polymerase Chain Reaction–Restriction Fragment Length Polymorphism Analysis: A Simple Method for Species Identification in Food. *J. AOAC Int.* **1995**, *78*, 6, 1542–1551.

- (222) Céspedes, A. N.; García, T.; Carrera, E.; Gonzáles, I.; Sanz, B.; Hernáz, P. E.; Martín, R., Identification of Flatfish Species Using Polymerase Chain Reaction (PCR) Amplification and Restriction Analysis of the Cytochrome b Gene. *J. Food Sci.* **1998**, *63*, 2, 206–209.
- (223) Fernandez, A.; Garcia, T.; Asensio, L.; Rodriguez, M. A.; Gonzalez, I.; Hernandez, P. E.; Martin., R., PCR-RFLP Analysis of the Internal Transcribed Spacer (ITS) Region for Identification of 3 Clam Species. *J. Food Sci.* **2001**, *66*, 5, 657–661.
- (224) Dwight, Z.; Palais, R.; Wittwer, C. T., uMELT: prediction of high-resolution melting curves and dynamic melting profiles of PCR products in a rich web application. *Bioinformatics* **2011**, *27*, 7, 1019–1020.
- (225) Dove, W. F.; Davidson, N., Cation effects on the denaturation of DNA. *J. Mol. Biol.* **1962**, *5*, 5, 467–478.
- (226) Nakano, S.; Fujimoto, M.; Hara, H.; Sugimoto, N., Nucleic acid duplex stability: influence of base composition on cation effects. *Nucleic Acids Res.* **1999**, *27*, 14, 2957–2965.
- (227) Escara, J. F.; Hutton, J. R., Thermal stability and renaturation of DNA in dimethyl sulfoxide solutions: acceleration of the renaturation rate. *Biopolymers* **1980**, *19*, 7, 1315–1327.
- (228) QuantaBio. *AccuMelt™ HRM SuperMix*, **2018**.
- (229) Sakaridis, I.; Ganopoulos, I.; Argiriou, A.; Tsaftaris, A., High resolution melting analysis for quantitative detection of bovine milk in pure water buffalo mozzarella and other buffalo dairy products. *Int. Dairy J.* **2013**, *28*, 1, 32–35.
- (230) Bosmali, I.; Ganopoulos, I.; Madesis, P.; Tsaftaris, A., Microsatellite and DNA-barcode regions typing combined with High Resolution Melting (HRM) analysis for food forensic uses: A case study on lentils (*Lens culinaris*). *Food Res. Int.* **2012**, *46*, 1, 141–147.
- (231) Reed, G. H.; Kent, J. O.; Wittwer, C. T., High-resolution DNA melting analysis for simple and efficient molecular diagnostics. *Pharmacogenomics* **2007**, *8*, 6, 597–608.
- (232) Munns, A. R.; Tollin, P., The crystal and molecular structure of inosine. *Acta Crystallogr. B Struct. Crystallogr. Cryst. Chem.* **1970**, *26*, 8, 1101–1113.
- (233) Case-Green, S. C.; Southern, E. M., Studies on the base pairing properties of deoxyinosine by solid phase hybridisation to oligonucleotides. *Nucleic Acids Res.* **1994**, *22*, 2, 131–136.
- (234) Felbinger, C.; Kutzsche, F.; Mönkediek, S.; Fischer, M., Genetic profiling: Differentiation and identification of hazelnut cultivars (*Corylus avellana* L.) using RAPD-PCR. *Food Control* **2020**, *107*, 106791.
- (235) Satya, P.; Karan, M.; Jana, S.; Mitra, S.; Sharma, A.; Karmakar, P. G.; Ray, D. P., Start codon targeted (SCoT) polymorphism reveals genetic diversity in wild and domesticated populations of ramie (*Boehmeria nivea* L. Gaudich.), a premium textile fiber producing species. *Meta Gene* **2015**, *3*, 62–70.
- (236) Öztürk, S. C.; Göktay, M.; Allmer, J.; Doğanlar, S.; Frary, A., Development of Simple Sequence Repeat Markers in Hazelnut (*Corylus avellana* L.) by Next-Generation Sequencing and Discrimination of Turkish Hazelnut Cultivars. *Plant Mol Biol Rep* **2018**, *36*, 5-6, 800–811.
- (237) *PAST: Paleontological statistics software package for education and data analysis*, **2001**.
- (238) Boccacci, P.; Botta, R.; Rovira, M., Genetic Diversity of Hazelnut (*Corylus avellana* L.) Germplasm in Northeastern Spain. *HortScience* **2008**, *43*, 3, 667–672.

- (239) Hajibarat, Z.; Saidi, A.; Hajibarat, Z.; Talebi, R., Characterization of genetic diversity in chickpea using SSR markers, Start Codon Targeted Polymorphism (SCoT) and Conserved DNA-Derived Polymorphism (CDDP). *Physiol. Mol. Biol. Plants* **2015**, *21*, 3, 365–373.
- (240) *European Hazelnut (Corylus avellana L.)*, <https://www.cavellanagenomeportal.com/>, Zuletzt geprüft am: 22.06.2022.
- (241) Khan, A.; Khan, I. A.; Heinze, B.; Azim, M. K., The Chloroplast Genome Sequence of Date Palm (Phoenix dactylifera L. cv. 'Aseel'). *Plant Mol Biol Rep* **2012**, *30*, 3, 666–678.
- (242) Khan, A. L.; Asaf, S.; Lee, I.-J.; Al-Harrasi, A.; Al-Rawahi, A., First chloroplast genomics study of Phoenix dactylifera (var. Naghal and Khanezi): A comparative analysis. *PLoS one* **2018**, *13*, 7, e0200104.
- (243) Kahlau, S.; Aspinall, S.; Gray, J. C.; Bock, R., Sequence of the tomato chloroplast DNA and evolutionary comparison of solanaceous plastid genomes. *Journal of molecular evolution* **2006**, *63*, 2, 194–207.
- (244) Teske, D.; Peters, A.; Möllers, A.; Fischer, M., Genomic Profiling: The Strengths and Limitations of Chloroplast Genome-Based Plant Variety Authentication. *J. Agr. Food Chem.* **2020**, *68*, 49, 14323–14333.
- (245) Smith, S. A.; Donoghue, M. J., Rates of molecular evolution are linked to life history in flowering plants. *Science* **2008**, *322*, 5898, 86–89.
- (246) Petit, R. J.; Hampe, A., Some Evolutionary Consequences of Being a Tree. *Annu. Rev. Ecol. Evol. Syst.* **2006**, *37*, 1, 187–214.
- (247) Matsuoka, Y.; Yamazaki, Y.; Ogihara, Y.; Tsunewaki, K., Whole chloroplast genome comparison of rice, maize, and wheat: implications for chloroplast gene diversification and phylogeny of cereals. *Mol. Biol. Evol.* **2002**, *19*, 12, 2084–2091.
- (248) Willyard, A.; Syring, J.; Gernandt, D. S.; Liston, A.; Cronn, R., Fossil Calibration of Molecular Divergence Infers a Moderate Mutation Rate and Recent Radiations for Pinus. *Mol. Biol. Evol.* **2006**, *24*, 2, 620.
- (249) Drouin, G.; Daoud, H.; Xia, J., Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. *Mol. Phylogenet. Evol.* **2008**, *49*, 3, 827–831.
- (250) Bachmann, R.; Klockmann, S.; Haerdter, J.; Fischer, M.; Hackl, T., ¹H NMR Spectroscopy for Determination of the Geographical Origin of Hazelnuts. *J. Agr. Food Chem.* **2018**, *66*, 44, 11873–11879.
- (251) Klockmann, S.; Reiner, E.; Bachmann, R.; Hackl, T.; Fischer, M., Food Fingerprinting: Metabolomic Approaches for Geographical Origin Discrimination of Hazelnuts (Corylus avellana) by UPLC-QTOF-MS. *J. Agr. Food Chem.* **2016**, *64*, 48, 9253–9262.
- (252) Bachmann, R.; Shakiba, N.; Fischer, M.; Hackl, T., Assessment of Mixtures by Spectral Superposition. An Approach in the Field of Metabolomics. *J. Proteome Res.* **2019**, *18*, 6, 2458–2466.
- (253) Oddone, M.; Aceto, M.; Baldizzone, M.; Musso, D.; Osella, D., Authentication and traceability study of hazelnuts from Piedmont, Italy. *J. Agr. Food Chem.* **2009**, *57*, 9, 3404–3408.

- (254) Sciubba, F.; Di Cocco, M. E.; Gianferri, R.; Impellizzeri, D.; Mannina, L.; Salvador, F. R. de; Venditti, A.; Delfini, M., Metabolic profile of different Italian cultivars of hazelnut (*Corylus avellana*) by nuclear magnetic resonance spectroscopy. *Nat. Prod. Res.* **2014**, *28*, 14, 1075–1081.
- (255) Alasalvar, C.; Odabasi, A. Z.; Demir, N.; Balaban, M. Ö.; Shahidi, F.; Cadwallader, K. R., Volatiles and Flavor of Five Turkish Hazelnut Varieties as Evaluated by Descriptive Sensory Analysis, Electronic Nose, and Dynamic Headspace Analysis/Gas Chromatography-Mass Spectrometry. *J. Food Sci.* **2004**, *69*, 3, SNQ99-SNQ106.
- (256) Ozturk, S. C.; Ozturk, S. E.; Celik, I.; Stampar, F.; Veberic, R.; Doganlar, S.; Solar, A.; Frary, A., Molecular genetic diversity and association mapping of nut and kernel traits in Slovenian hazelnut (*Corylus avellana*) germplasm. *Tree Genet. Genomes* **2017**, *13*, 1.
- (257) Öztürk, S. C.; Balık, H. İ.; Balık, S. K.; Kızılcı, G.; Duyar, Ö.; Doğanlar, S.; Frary, A., Molecular genetic diversity of the Turkish national hazelnut collection and selection of a core set. *Tree Genet. Genomes* **2017**, *13*, 5.
- (258) Campa, A.; Trabanco, N.; Pérez-Vega, E.; Rovira, M.; Ferreira, J. J., Genetic relationship between cultivated and wild hazelnuts (*Corylus avellana* L.) collected in northern Spain. *Plant Breeding* **2011**, *130*, 3, 360–366.
- (259) Martins, S.; Simões, F.; Mendonça, D.; Matos, J.; Silva, A. P.; Carnide, V., Western European Wild and Landraces Hazelnuts Evaluated by SSR Markers. *Plant Mol Biol Rep* **2015**, *33*, 6, 1712–1720.
- (260) Cenis, J. L., Rapid extraction of fungal DNA for PCR amplification. *Nucleic Acids Res.* **1992**, *20*, 9, 2380.
- (261) Nguyen, T.; Kwak, S.; Karpowicz, S. J., Re-use of commercial microfluidics chips for DNA, RNA, and protein electrophoresis. *BioTechniques* **2014**, *57*, 5, 267–271.
- (262) Blake, R. D.; Delcourt, S. G., Thermal stability of DNA. *Nucleic Acids Res.* **1998**, *26*, 14, 3323–3332.
- (263) Hammer, Ø.; Harper, D. A.; Ryan, P. D., PAST: Paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* **2001**, *4* (1).
- (264) *MATLAB*; MathWorks, Inc.: Natick, USA, **2021**.
- (265) Zhang, Z.; Schwartz, S.; Wagner, L.; Miller, W., A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* **2000**, *7*, 1-2, 203–214.

9 Anhang

9.1 Gefahrstoffverzeichnis

Substanz	Gefahrensymbol	Hazard Statement	Precaution Statement	Entsorgung ^{a)}
Cetyltrimethylammoniumbromid (CTAB)		302, 315, 318, 335, 373, 410	273, 280, 302+352, 305+351+338, 314	(3)
Chloroform		302, 331, 315, 319, 351, 361d, 336, 372	261, 281, 305+351+338, 311	(3)
Ethanol		225, 319	210, 240, 305+351+338, 403+233	(2)
Ethidiumbromid		302, 330, 341	201, 202, 260, 284, 301+312, 304+340+310	(3)
Ethylendiamintetraessigsäure (EDTA)		319, 332, 373	280, 304+340, 312, 305+351+338, 308+310	(3)
Guanidinhydrochlorid		302+332, 315, 319	260, 280, 302+352, 305+351+338	(1)
Isopropanol		225, 319, 336	210, 233, 240, 305+351+338, 403+235	(2)
n-Hexan		225, 304, 361f, 373, 315, 336, 411	210, 240, 273, 301+310, 331, 302+352, 403, 235	(3)
Natriumdodecylsulfat (SDS)		228, 302+332, 315, 318, 335, 412	210, 261, 280, 301+312+330, 305+351+338, 308+310, 370+378	(3)

Substanz	Gefahrensymbol	Hazard Statement	Precaution Statement	Entsorgung ^{a)}
Salzsäure		290, 314, 335	280, 301+330+331, 305+351+338, 312, 403+233	(1)
Xylencyanol		315, 319, 335	261, 305+351+338	(2)

a) Entsorgungsschlüssel

- (1) ggf. lösen, pH-Wert einstellen und in den Sammelbehälter für andere Säuren (HCl- und H₂SO₄-haltig)
- (2) ggf. lösen und in den Sammelbehälter für halogenfreie organische Lösungsmittel
- (3) ggf. lösen und in den Sammelbehälter für halogenhaltige, giftige sowie KMR-haltige organische Lösungsmittel

9.2 Verwendete Verbrauchsmaterialien

Die für die Durchführung der praktischen Arbeiten verwendeten Verbrauchsmaterialien sind der Tabelle 21 zu entnehmen.

Tabelle 21: Verwendete Verbrauchsmaterialien mit Angabe des Herstellers und der Spezifikation.

Material	Hersteller	Spezifikation
96 PCR-Platte (qPCR)	Sarstedt AG & Co. KG, Nümbrecht, Deutschland	76x119,9 mm, transparent, randlos, Max.-Volumen: 200 µL, Material: Polypropylen
Biosphere® Filter Tips 0,1-20 µL	Sarstedt AG & Co. KG, Nümbrecht, Deutschland	Länge: 46 mm, PP, Filter: PE, steril, DNA-, DNase-/RNase-, PCR Inhib.-, ATP-frei, pyrogenfrei/endotoxinfrei
Biosphere® Filter Tips 100-1000 µL	Sarstedt AG & Co. KG, Nümbrecht, Deutschland	Durchmesser: 9 mm, Länge: 72 mm, PP, Filter: PE, steril, DNA-, DNase-/RNase-, PCR Inhib.-, ATP-frei, pyrogenfrei/endotoxinfrei
Biosphere® Filter Tips 2-100 µL	Sarstedt AG & Co. KG, Nümbrecht, Deutschland	Durchmesser: 8 mm, Länge: 51 mm, PP, Filter: PE, steril, DNA-, DNase-/RNase-, PCR Inhib.-, ATP-frei, pyrogenfrei/endotoxinfrei
EconoSpin® Column	Epoch LifeScience Inc., Sugar Land, TX 77496	EconoSpin Mini Spin Column, mit Deckel
Einmalhandschuhe (Nitril)	VWR International GmbH, Darmstadt, Deutschland	Größe M, puderfrei
Einmalhandschuhe (Vinyl)	VWR International GmbH, Darmstadt, Deutschland	Größe M, puderfrei
Gefrierbeutel	Pely-plastic GmbH & Co. KG, Wahlstedt, Deutschland	Frischhalte-Beutel mit Sicherheitsverschluss, V = 3 L, 27 x 25 cm,
Gefrierbeutel	Quickpack Haushalt + Hygiene GmbH, Renningen, Deutschland	Ja! Gefrierbeutel mit Verschlussclips, V = 3 L, 25 X 32 cm
Klebefolie (qPCR)	Sarstedt AG & Co. KG, Nümbrecht, Deutschland	79x141 mm, optisch klar, DNase-/RNase-frei, lösungsmittelbeständig, Material: Polyolefin
Mikro-Einweg-Laborspatel	Carl Roth GmbH & Co. KG, Karlsruhe, Deutschland	Einweg-Kunststoff-Spatel mikro, antistatisch, unsteril
Parafilm	Bemis Company, Inc., Oshkosh, USA	Laboratory Film, 4inchx125ft.-Rolle, PM-996
Pasteurpipetten	Thermo Fisher Scientific GmbH, Schwerte, Deutschland	Länge: 230 mm/150 mm

Material	Hersteller	Spezifikation
PCR-Einzelgefäß 0,2 mL	Sarstedt AG & Co. KG, Nümbrecht, Deutschland	PCR-Einzelgefäß mit flachem Deckel, Multiply®-Pro, PCR Performance Tested (DNA-, DNase-/RNase-, PCR-Inhib.-frei) 0,2 ml, transparent
PCR-Einzelgefäß 0,5 mL	Fisher Scientific GmbH, Schwerte, Deutschland	Axygen™ PCR-Röhrchen, 0,5 mL, mit flachem Deckel
Pipettenspitzen 10 µL (StackPack)	Sarstedt AG & Co. KG, Nümbrecht, Deutschland	Durchmesser: 6 mm Länge: 31 mm, PP
Pipettenspitzen 1000 µL	Sarstedt AG & Co. KG, Nümbrecht, Deutschland	Durchmesser: 9 mm Länge: 72 mm, PP
Pipettenspitzen 2-200 µL (StackPack)	Sarstedt AG & Co. KG, Nümbrecht, Deutschland	Länge: 51 mm, PP, PCR Performance Tested (DNA-, DNase-/RNase-, PCR-Inhib.-frei)
Präzisionswischtücher	Kimberley-Clark Professional, Koblenz-Rheinhafen, Deutschland	Weiß, 11 x 21 cm
Reaktionsgefäß 1,5 mL	Sarstedt AG & Co. KG, Nümbrecht, Deutschland	Transparent, PP, mit anhängendem Deckel
Reaktionsgefäß 2 mL	Sarstedt AG & Co. KG, Nümbrecht, Deutschland	Transparent, PP, mit anhängendem Deckel
Tubes 15 mL	Sarstedt AG & Co. KG, Nümbrecht, Deutschland	120x17 mm, Spitzboden, PP, steril und Pyrogen/endotoxinfrei
Tubes 50 mL (braun)	Sarstedt AG & Co. KG, Nümbrecht, Deutschland	114x28 mm, Spitzboden, PP, braun, steril und Pyrogen/endotoxinfrei
Tubes 50 mL (klar)	Sarstedt AG & Co. KG, Nümbrecht, Deutschland	114x28 mm, Spitzboden, PP, steril und Pyrogen/endotoxinfrei

9.3 Verwendete Chemikalien, Puffer, Kits und Primer

Die verwendeten Chemikalien und Puffer mit Angaben zu den Herstellern und Spezifikationen sind Tabelle 22 zu entnehmen.

Tabelle 22: Verwendete Chemikalien und Puffer mit Angabe des Herstellers und der Spezifikationen.

Chemikalie	Hersteller	Spezifikation
<i>Accu Melt™ HRM Super Mix</i>	Quantabio, Beverly, USA	2-fach konzentriert Zusammensetzung: firmenspezifisch
Agar-Agar	Carl Roth GmbH + Co. KG, Karlsruhe, Deutschland	Bakteriologisch, CAS-Nr.: 9002-18-0

Chemikalie	Hersteller	Spezifikation
Agarose	SERVA Electrophoresis GmbH, Heidelberg, Deutschland	For DNA electrophoresis, research grade CAS-Nr.: 9012-36-6
Agencourt® AMPure® XP	Beckman Coulter Inc., Brea, USA	---
Ampicillin Natriumsalz	Carl Roth GmbH & Co. KG, Karlsruhe, Deutschland	≥ 99 % CAS-Nr.: 69-52-3
Bindepuffer (DNA-Isolierung mittels CTAB-Methode)	Eigene Herstellung	5,5 M Guanidinhydrochlorid, 20 mM Tris-HCl in Wasser, HCl q.s. pH 6,6
Biozym <i>Taq</i> DNA Polymerase	Biozym Scientific GmbH, Hessisch Oldendorf, Deutschland	5 U/μL, 1000 Units
Biozym <i>Taq</i> DNA Polymerase Puffer	Biozym Scientific GmbH, Hessisch Oldendorf, Deutschland	10-fach konzentriert Enthält 30 mM MgCl ₂
BSA (Bovines Serumalbumin)	New England Biolabs® GmbH, Frankfurt am Main, Deutschland	Purified BSA, 10 mg/mL
Caseinhydrolysat, standard	Carl Roth GmbH & Co. KG, Karlsruhe, Deutschland	Für die Molekularbiologie, säurehydrolysiert CAS-Nr.: 65072-00-6
Cetyltrimethylammoniumchlorid (CTAB)	Carl Roth GmbH & Co. KG, Karlsruhe, Deutschland	≥ 99 %, für die Biochemie CAS-Nr.: 57-09-0
Chloroform	VWR International GmbH, Darmstadt, Deutschland	99 % + 0,6 % Ethanol CAS-Nr.: 67-66-3
<i>CutSmart® Buffer</i>	New England BioLabs® Inc., Ipswich, USA	10-fach konzentriert
D-Glucose wasserfrei	Acrose Organics, New Jersey, US	CAS-Nr.: 50-99-7
dNTPs	Bioline GmbH, Luckenwalde, Deutschland	1:1:1:1 (dATP:dTTP:dGTP:dCTP) 10 mM (4x 2,5 mM)
<i>DreamTaq</i> DNA Polymerase	Thermo Fisher Scientific™, Waltham, USA	5 U/μL, 500 U
<i>DreamTaq</i> DNA Polymerase Puffer	Thermo Fisher Scientific™, Waltham, USA	10-fach konzentriert Enthält 20 mM MgCl ₂
EDTA-disodium	SERVA Electrophoresis GmbH, Heidelberg, Deutschland	Analytical grade CAS-Nr.: 6381-92-6
Essigsäure	Merck KGaA, Darmstadt, Deutschland	96 %, zur Analyse CAS-Nr.: 64-19-7
Ethanol absolut	VWR International GmbH, Darmstadt, Deutschland	100 % CAS-Nr.: 64-17-5
Extraktionspuffer (DNA-Isolierung mittels CTAB-Methode)	Eigene Herstellung	55 mM CTAB, 1,4 M NaCl, 0,1 M Tris- HCl, 20 mM EDTA in Wasser, HCl q.s. pH 8,0
Extraktionspuffer (DNA-Isolierung mittels <i>Rapid Extraction</i>)	Eigene Herstellung	200 mM Tris, 250 mM NaCl, 25 mM EDTA, 0,5 % SDS, HCl q.s. pH 8,5
GelRed™ Nucleic Acid Gel Stain	Biotium, Inc., Fremont, USA	10.000-fach konzentriert, in Wasser

Chemikalie	Hersteller	Spezifikation
Glycerin	Carl Roth GmbH & Co. KG, Karlsruhe, Deutschland	≥ 99,5 % CAS-Nr.: 56-81-5
Guanidinhydrochlorid	Carl Roth GmbH & Co. KG, Karlsruhe, Deutschland	≥ 99,5 % CAS-Nr.: 50-01-1
Hefeextrakt	Carl Roth GmbH & Co. KG, Karlsruhe, Deutschland	Pulver, für die Bakteriologie CAS-Nr.: 8013-01-2
HF-Puffer (für <i>S7 Fusion High-Fidelity DNA</i> Polymerase)	Biozym Scientific GmbH, Hessisch Oldendorf, Deutschland	5-fach konzentriert Enthält 7,5 mM MgCl ₂
Isopropanol	VWR International GmbH, Darmstadt, Deutschland	100 % CAS-Nr.: 67-63-0
Kaliumchlorid	Carl Roth GmbH & Co. KG, Karlsruhe, Deutschland	≥ 99,5 % CAS-Nr.: 7447-40-7
LBamp-Agar	Eigene Herstellung	15 g/L Agar 170 mg/L Ampicillin (nach Autoklavieren) 10 g/L Caseinhydrolysat 5 g/L Hefeextrakt 5 g/L Natriumchlorid Ad 1 L bidest. Wasser
LBamp-Medium	Eigene Herstellung	100 mg/L Ampicillin (nach Autoklavieren) 10 g/L Caseinhydrolysat 5 g/L Hefeextrakt 5 g/L Natriumchlorid Ad 1 L bidest. Wasser
<i>Loading Dye</i>	Eigene Herstellung	49,9975 % Glycerin 49,9975 % TAE-Puffer 0,005% Xylencyanol
Magnesiumchlorid	Sigma-Aldrich Chemie GmbH, München, Deutschland	CAS-Nr.: 7791-18-6
Magnesiumsulfat wasserfrei	Sigma-Aldrich Chemie GmbH, München, Deutschland	CAS-Nr.: 7487-88-9
<i>MfeI</i> (Restriktionsenzym)	New England BioLabs® Inc., Ipswich, USA	10 U/μL, 500 Units
Natriumacetat	Grüssing GmbH, Filsum, Deutschland	99 %, CAS-Nr.: 127-09-3
Natriumchlorid	Carl Roth GmbH + Co. KG, Karlsruhe, Deutschland	≥ 99 % CAS-Nr.: 7647-14-5
Natriumdodecylsulfat (SDS)	Sigma-Aldrich Chemie GmbH, München, Deutschland	99 %, CAS-Nr.: 151-21-3
<i>PeqGold DNA Ladder Mix</i>	VWR International GmbH, Darmstadt, Deutschland	Fragmentlängen: 100-10.000 bp

Chemikalie	Hersteller	Spezifikation
Primer	Invitrogen, Thermo Fisher Scientific™, Waltham, USA	Entsalzt
Primer	Integrated DNA Technologies, Illinois, USA	Entsalzt
<i>S7 Fusion High-Fidelity DNA Polymerase</i>	Biozym Scientific GmbH, Hessisch Oldendorf, Deutschland	2 U/μL, 100 U
Salzsäure	VWR International GmbH, Darmstadt, Deutschland	37 % CAS-Nr.: 7647-01-0
SOC-Medium	Eigene Herstellung	2 % Caseinhydrolysat 20 mM Glucose 0,5 % Hefeextrakt 2,5 mM Kaliumchlorid 10 mM Magnesiumchlorid 10 mM Magnesiumsulfat 10 mM Natriumchlorid Ad 100 mL bidest. Wasser
TAE-Puffer	Eigene Herstellung	50-fach konzentriert 2 M TRIS, 100 mM EDTA, in Wasser, CH ₃ COOH q.s. pH 8,2
<i>Taq Polymerase Lagerpuffer</i>	Eigene Herstellung	2,5 mM Tris-HCl, 150 mM NaCl, 0,05 mM EDTA, 1 mM DTT, in Wasser, 50 % Glycerin, HCl q.s. pH 8,8
TE-Puffer	Eigene Herstellung	10 mM TRIS, 1 mM EDTA, in Wasser
Trizma® Base (TRIS)	Sigma-Aldrich Chemie GmbH, München, Deutschland	CAS-Nr.: 77-86-1
Trizma® hydrochloride (TRIS-HCl)	Sigma-Aldrich Chemie GmbH, München, Deutschland	CAS-Nr.: 1185-53-1
Tween® 20	Carl Roth GmbH & Co. KG, Karlsruhe, Deutschland	CAS-Nr.: 9005-64-5
Waschpuffer (DNA-Isolierung mittels CTAB-Methode)	Eigene Herstellung	20 mM Tris-HCl, 1 mM EDTA, 50 mM NaCl, in Wasser, 50 % Ethanol, HCl q.s. pH 7,4
Xylencyanol	Merck KGaA, Darmstadt, Deutschland	For electrophoresis CAS-Nr.: 2650-17-1

Die genutzten Kits mit Angaben zu den Herstellern sind in Tabelle 23 zusammengefasst.

Tabelle 23: Verwendete Kits mit Angabe des Herstellers.

Kit	Hersteller
<i>AccuMelt™ HRM SuperMix</i>	QuantaBio, Beverly, USA
<i>Agilent DNA 7500 Kit</i>	Agilent Technologies, Inc., Santa Clara, USA
<i>Agilent High Sensitivity DNA Kit</i>	Agilent Technologies, Inc., Santa Clara, USA

Kit	Hersteller
<i>Index Adapter Replacement Caps</i>	Illumina Inc., San Diego, USA
<i>MiSeq[®] v2, Reagent Kit (Box 1 und 2)</i>	Illumina Inc., San Diego, USA
<i>Monarch[®] PCR & DNA Cleanup Kit</i>	New England Biolabs GmbH, Frankfurt am Main, Deutschland
<i>Monarch[®] Plasmid Miniprep Kit</i>	New England Biolabs GmbH, Frankfurt am Main, Deutschland
<i>Nextera XT Library Prep Kit</i>	Illumina Inc., San Diego, USA
<i>Nextera[®] XT Index Kit</i>	Illumina Inc., San Diego, USA
<i>Nextera[®] XT Library Prep Kit 24 samples, Box 1 und 2</i>	Illumina Inc., San Diego, USA
<i>Nextera[™] DNA CD Indexes 24 Indexes, 24 samples</i>	Illumina Inc., San Diego, USA
<i>Nextera[™] DNA Flex Library Prep 24 samples, Box 1-3</i>	Illumina Inc., San Diego, USA
<i>peqGold Plant DNA Mini Kit</i>	VWR International GmbH, Darmstadt, Deutschland
<i>peqGold Gel Extraction Kit</i>	VWR International GmbH, Darmstadt, Deutschland
<i>QuantiFluor[®] dsDNA System</i>	Promega GmbH, Mannheim, Deutschland
<i>Topo[™] TA Cloning[™] Kit for Sequencing, without competent cells</i>	Life Technologies GmbH, Darmstadt, Deutschland

Die für die Analysen eingesetzten Primer inklusive ihrer Spezifikationen, den zu erwartenden Fragmentlängen sowie Annealingtemperaturen sind der folgenden Tabelle zu entnehmen.

Tabelle 24: Verwendete Primer inklusive der Spezifikationen, zu erwartenden Fragmentlängen und gewählten Annealingtemperaturen. Sofern es sich nicht um selbst entwickelte Primer handelt, ist die entsprechende Quelle in der Spalte „Name“ angegeben.

Name	Sequenz (5'-3')	Annealingtemperatur [°C]		Fragmentlänge [bp]
		Taq DNA Polymerase	S7 (HF) DNA Polymerase	
A_FW	ATCCACTTGGCTACATCCGC	54,2	65,0	698
A_RV	TAGCGGCTTGGCCTGTAGTA			
B.1_FW	GAGTACTCTACCGTTGAGTT	60,0	60,0	592/710*
B.1_RV	TGATTGCGTAAACCCAATCA			
B.2_FW	GACTACCACATATTTTGTTCGTATG	45,6	---	84/202*
B.2_RV	CAAACAAATATAATCCGATCGTG			
C_FW	TGGGCTCTGTAAGAATTTGT	63,0	60,0	494/609**
C_RV	GTGATTCCTTTTCTCGCTAC			
D.1_FW	AAAGGATTGAGCCGAATACA	53,5	62,0	670
D.1_RV	TGAACCACCCTTTTCCATA			
D.2_FW	CCGATCATGGATCGAGCC	49,0	---	619
D.2_RV	CTAGAATTGTCACTATCGCTCCAAC			
D.HRMA_FW	TGATACAACAAACACGGCGG	52,2	---	99
D.HRMA_RV	GAGAACCAGGAGTATCACTTAATATGATGG			

Name	Sequenz (5'-3')	Annealingtemperatur [°C]		Fragmentlänge [bp]
		Taq DNA Polymerase	S7 (HF) DNA Polymerase	
ITS-5 (FW) ^[3]	GGAAGGAGAAGTCGTAACAAGG	50,1	---	666
26S-25R (RV) ^[3]	TATGCTTAAACTCAGCGGGT			
WRKY_FP1 ^[69]	TGGMGIAARTAYGGNCARA	50,0	---	---
WRKY_RP2 ^[69]	TGRBYRTGYTTICCYTCRTAIGTDGT			
WRKY-280K1_FW	GGAGGAAGTACGGGCAGAAG	54,7		261
WRKY-280K1_RV	GGTTATGTTTGCCTTCGTAGGTGG			
WRKY-280K2_FW	GTGATTACCGACTTAAGATCATGCG	54,1		233
WRKY-280K2_RV	GGCGGAAGTATGGACAGAAAGTTG			
WRKY-320K3_FW	GTTGTCAAAGGAAATCCAAACCTAGG	65,0		259
WRKY-320K3_RV	GTGAGAGGCCCTTCCACATGC			
WRKY-950K16_FW	GGCGGAAGTATGGACAGAAAGTGG	56,9		948
WRKY-950K16_RV	GGCTGTGCTTGCCCTCGTAC			
SCoT 7 ^[163]	CAACAATGGCTACCACGG	50,0	---	---
SCoT 13 ^[163]	ACGACATGGCGACCATCG	50,0	---	---
SCoT 17 ^[163]	ACCATGGCTACCACCGAG	50,0	---	---
SCoT 19 ^[163]	ACCATGGCTACCACCGGC	50,0	---	---
SCoT 21 ^[163]	ACGACATGGCGACCCACA	46,4	---	---
SCoT 24 ^[163]	CACCATGGCTACCACCAT	50,0	---	---
SCoT 32 ^[163]	CCATGGCTACCACCGCAC	50,0	---	---
SCoT 35 ^[163]	CATGGCTACCACCGCCC	50,0	---	---
SCoT 36 ^[163]	GCAACAATGGCTACCACC	50,0	---	---
RAPD 17 ^[234]	GACCGCTTGT	34,0	---	---
SCAR_TomV_FW	GACCGCTTGTGGCGCAATCC	55,1	---	ca. 350
SCAR_TomV_RV	CCGCTTGTATGATACGATACAGAGTCAC			
SCAR_TomE_FW	TGGCGCAATCCACTGAC	49,6	---	ca. 350
SCAR_TomE_RV	ATGATACGATACAGAGTCACGTTAG			
M13_FW	GTAAAACGACGGCCAG	45,0	---	165 + Länge Insert
M13_RV	CAGGAAACAGCTATGAC			
Tom_K16_FW	GGT GAT GCA GGG GTT GTA ATT GTG G	57,3	---	ca. 950
Tom_K16_RV	GCCCTGGTTCATCCATGATACTACCAC			
Tom_K17_FW	CAGAGTTTGTGCCCTTGTG	51,3	---	ca. 780
Tom_K17_RV	GAGGATCGATACCTCACGTC			
Tom_K18_FW	ATGAAGTAATTGTGTGAACCCATGTTG	51,6	---	ca.593
Tom_K18_RV	ACTTCTTTAAGCACGAATGGAGAAG			
Tom_K20_FW	GATTAGAGTTTATACAGAGTTTGTGCC	49,8	---	ca. 640
Tom_K20_RV	CTATAGCAATACACTTAACCACTTCTAGG			

* z.T. längeres Fragment aufgrund einer Insertion mit einer Länge von 118 bp, ** z.T. längeres Fragment aufgrund einer Insertion mit einer Länge von 115 bp

9.4 Geräteparameter

Die verwendeten Geräte sowie Herstellerangaben und Spezifikationen sind in Tabelle 25 zusammengefasst.

Tabelle 25: Verwendete Geräte mit Angabe des Herstellers und der Spezifikationen.

Gerät	Hersteller	Spezifikation
Analysenwaage	Kern & Sohn GmbH, Balingen-Frommern, Deutschland	Modell: AGB 210-4 Genauigkeit: 0,1 mg
Analysenwaage	Mettler-Toledo GmbH, Gießen, Deutschland	Modellreihe: TLE Genauigkeit: 0,1 mg
DNA-Workstation	Kisker Biotech, Steinfurt, Deutschland	DNA/RNA UV-Cleaner Box UVC/T-M-AR
Eismaschine	ITV Ice Makers, Miami, USA	Ice queen
Elektrophoreseeinheit	Appligene, Illkirch, Frankreich	---
Elektrophorese-Netzgerät	Carl Roth GmbH & Co. KG, Karlsruhe, Deutschland	Electrophoresis Power Supply peqlab EV 231
Elektroporationsgerät	Eppendorf Vertrieb Deutschland GmbH, Wesseling-Berzdorf, Deutschland	Eppendorf Eporator®
Eppendorf Pipetten	Eppendorf AG, Hamburg, Deutschland	100-1000 µL, 10-100 µL, 1-10 µL, 0,1-2,5 µL
Gefrierschrank (klein)	Robert Bosch Hausgeräte GmbH, München, Deutschland	GSP36A31/02, -21 °C
Gefrierschrank (groß)	Liebherr-International Deutschland GmbH, Biberach an der Riß, Deutschland	Comfort, G1213, -21 °C
Geldokumentationseinheit	Biostep® GmbH, Jahnsdorf, Deutschland	Biostep® Dark Hood DH-40/50
Inkubationsschüttler	Edmund Bühler GmbH, Bodelshausen, Deutschland	Inkubationshaube TH 30 und Schüttler SM 30
Inkubationsschüttler	Thermo Scientific, Thermo Electron LED GmbH, Langenselbold, Deutschland	MAXQ 8000
Kapillargelelektrophorese	Agilent Technologies, Inc., Santa Clara, USA	2100 Bioanalyzer
Kühlschrank (klein)	Liebherr-International Deutschland GmbH, Biberach an der Riß, Deutschland	Comfort, TP1410, 5 °C
Kühlschrank (groß)	Liebherr-International Deutschland GmbH, Biberach an der Riß, Deutschland	MediLine, Lkexv3600, 5 °C
Messermühle	Retsch GmbH, Haan, Deutschland	GM300

Gerät	Hersteller	Spezifikation
Mikrowelle	Severin Elektrogeräte GmbH, Sundern, Deutschland	Typ: 7810
Mikro-Zentrifuge/Vortex Mixer	Kisker Biotech GmbH & Co. KG, Steinfurt, Deutschland	Combi-Spin, FVL-2400N
NanoDrop™	DeNovix Inc., Wilmington, USA	DS-11 + Spektrophotometer
pH-Meter	Mettler-Toledo GmbH, Gießen, Deutschland	FiveEasy pH meter F20-Std-Kit
Quantus™ Fluorometer	Promega GmbH, Mannheim, Deutschland	---
<i>Real-time</i> PCR	Bio-Rad Laboratories, Inc., Hercules, USA	C1000 Touch™ Thermal Cycler, CFX96 Deep Well™ Real-Time System
Sequenziergerät	Illumina Inc., San Diego, USA	MiSeq™
Sicherheitswerkbank	Thermo Scientific, Thermo Electron LED GmbH, Langenselbold, Deutschland	Typ: S2020 1.2
Thermocycler	Biometra GmbH, Göttingen, Deutschland	T3000 Thermocycler
Thermocycler	Biometra GmbH, Göttingen, Deutschland	T3 Thermocycler
Thermocycler (Gradienten)	Biometra GmbH, Göttingen, Deutschland	TGradient Thermocycler
Thermoschüttler	bioSan, Riga, Lettland	TS-100
Trockenschrank	Thermo Scientific, Thermo Electron LED GmbH, Langenselbold, Deutschland	Typ B 20
Vortexmischer	peqlab® Biotechnologie GmbH, Erlangen, Deutschland	peQtwist
Vortexmischer	VWR International GmbH, Darmstadt, Deutschland	Modell: 444-1372
Vortexmischer	VWR International GmbH, Darmstadt, Deutschland	Modell: 1719
Vortexmischer	Kisker Biotech GmbH & Co. KG, Steinfurt, Deutschland	Modell: LSM Mixer Uzusio VTX- 3000L
Vortexmischer (für Bioanalyser- Chips)	IKA®-Werke GmbH & Co. KG, Staufen, Deutschland	Modell: MS 3 basic
Waage	Sartorius AG, Göttingen, Deutschland	Modell: 1203 MP
Wasseraufbereitungssystem	Merck Millipore, Merck KGaA, Darmstadt, Deutschland	Rios™ Essential 5 Direct-Q® 3 UV-R, Endfilter: Biopak, Reinstwasser (Typ 1)
Wasserbad	Prüfgeräte-Werk Medingen GmbH, Medingen, Deutschland	Typ: W 6

Gerät	Hersteller	Spezifikation
Wasserbad	Lauda Dr. R. Wobser GmbH & Co. KG, Lauda-Königshofen, Deutschland	Typ: AL2
Zentrifuge	Beckman Coulter® Inc., Brea, USA	Microfuge® 16 Centrifuge, Non-Refrigerated, 50/60 Hz, 220-240 V
Zentrifuge	Labo Gene Aps, Dänemark	Scanspeed 1524 M
Zentrifuge	SIGMA Laborzentrifugen GmbH Osterode am Harz, Deutschland	SIGMA 3-18 KS

9.5 Softwareparameter und Online-Tools

Die verwendeten Programme und Online-Tools mit entsprechenden Hersteller-/Betreiberangaben sowie der Anwendungen sind der Tabelle 26 zu entnehmen.

Tabelle 26: Verwendete Software und Online-Tools mit Angabe des Herstellers und der Anwendung.

Software	Version	Hersteller/Betreiber	Anwendung
2100 Expert	B.02.08.SI648 (SR3)	Agilent Technologies, Inc., Santa Clara, USA	Steuerung der Kapillargelelektrophorese, Datenauswertung
Biostep Argus X1	5.1.11	Biostep® GmbH, Jahnsdorf, Deutschland	Steuerung der Geldokumentationseinheit
BLASTN ^[265]	2.13.0+	National Library of Medicine, 8600 Rockville Pike Bethesda, MD 20894	Datenbankabgleich von Nukleotidsequenzen
CFX Manager™ Software	3.1	Bio-Rad Laboratories, Inc., Hercules, USA	Steuerung der <i>real-time</i> PCR, Datenauswertung
Citavi	6.14.4	Swiss Academic Software GmbH, Wädenswil, Schweiz	Literaturverwaltung
<i>Clustal Omega Alignment Tool</i>	Online-Tool	EMBL-EBI, Hinxton, UK	Erstellung von <i>Alignments</i>
Excel	365 ProPlus	Microsoft Corporation	Datenauswertung, Erstellung von Kalibriergeraden
<i>Genome Compiler</i>	2.2.88	Genome Compiler Corporation, Los Altos, USA	Datenauswertung, Erstellung <i>Alignments</i>
ImageJ	1.52a	Wayne Rasband, National Institutes of Health, USA	Graustufenanalyse
MATLAB	2021b	MathWorks, Inc., Natick, USA	Datenauswertung
MiSeq™ Control Software	2.6.1.1	Illumina Inc., San Diego, USA	Steuerung des MiSeq

Software	Version	Hersteller/Betreiber	Anwendung
<i>Multiple Primer Analyzer</i>	Online-Tool	Thermo Fisher Scientific™, Waltham, USA	Berechnung von Annealingtemperaturen und Prüfung von Primern auf Sekundärstrukturen und Dimer-Bildung
paint.net	4.3.3	dotPDN LLC	Bildbearbeitung
PAST	4.03	Freeware (Øyvind Hammer, University of Oslo)	Datenauswertung
PowerPoint	365 ProPlus	Microsoft Corporation	Erstellen von Grafiken
<i>Precision Melt Analysis</i>	1.2	Bio-Rad Laboratories, Inc., Hercules, USA	Schmelzpunktanalyse
uMELT Quartz	3.6.2 „Quartz“	DNA-UTAH.ORG	Berechnung von Schmelzkurven

9.6 Probenmaterial

In den folgenden drei Tabellen sind alle verwendeten Haselnussproben (Kerne: Tabelle 27 und Tabelle 28, Blätter: Tabelle 29) inklusive der Angaben zu Herkunft, Erntezeitpunkt und Lieferant aufgelistet. Die Erklärung der für die Barcodes verwendeten Abkürzungen ist der Tabelle 30 und Tabelle 31 zu entnehmen.

Tabelle 27: Auflistung aller verwendeten Haselnussproben (Rohware) mit Angabe der Sorte, Herkunft, Lieferant, verwendete Isolierungsmethode und Angabe der Analysemethoden, für welche die Proben eingesetzt wurden.

Nr.	Sorte	Herkunft	Ernte- jahr	Lieferant	Barcode	Analyse ^{a)}
102*	Ata Baba	Aserbaidsschan, Zaqatala	2015	Schlüter & Maack GmbH	15AZZA-NB- 001-ATA-HN	1, 2, 3, 4, 6
123*	Ata Baba	Aserbaidsschan	2014	Biozentrum Klein Flottbek	14AZNB-NB- 002-ATA-HN	1, 2, 3
S2*	Ata Baba	Aserbaidsschan	2014	August Storck KG	14AZNB-NB- 003-ATA-HN	1, 2, 3, 6
64*	Barcelona	Frankreich, Aquitainen	2014	SCA Unicoque	14FRAQ-LG- 004-BAR-HN	3
52*	Barcelona	Frankreich, Pays de la Loire	2014	SCA Unicoque	14FRPL-LQ-001- BAR-HN	3, 6
82*	Barcelona	Frankreich, Midi- Pyrénées	2014	SCA Unicoque	14FRMP-TG- 005-BAR-HN	6
43*	Barcelona	Frankreich, Aquitainen	2014	SCA Unicoque	14FRAQ-GI-004- BAR-HN	1, 6
69*	Butler	Frankreich, Aquitainen	2014	SCA Unicoque	14FRAQ-PA- 001-BUT-HN	4, 7

Nr.	Sorte	Herkunft	Ernte- jahr	Lieferant	Barcode	Analyse ^{a)}
40*	Butler	Frankreich, Aquitanien	2014	SCA Unicoque	14FRAQ-GI-001- BUT-HN	1, 3, 6
109*	Butler	Frankreich, Aquitanien	2015	SCA Unicoque	15FRAQ-PA- 001-BUT-HN	3, 6
57*	Butler	Frankreich, Midi- Pyrénées	2014	SCA Unicoque	14FRMP-LT- 001-BUT-HN	6
104*	Çakildak (Delisava)	Türkei, Ordu	2015	Schlüter & Maack GmbH	15TROR-OR- 005-CAK-HN	1, 6
105*	Çakildak (Delisava)	Türkei, Düzce, Akçakoca	2015	Biozentrum Klein Flottbek	15TRDU-AK- 005-CAK-HN	1, 6
103*	Çakildak (Delisava)	Türkei, Düzce, Akçakoca	2015	Biozentrum Klein Flottbek	15TRDU-AK- 001-CAK-HN	1, 6
132*	Çakildak (Delisava)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 001-CAK-HN	1
212**	Çakildak (Delisava)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 002-CAK-HN	1
213**	Çakildak (Delisava)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 003-CAK-HN	1
214**	Çakildak (Delisava)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 006-CAK-HN	1
215**	Çakildak (Delisava)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 007-CAK-HN	1
216**	Çakildak (Delisava)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 008-CAK-HN	1
F2*	Corabel	Deutschland, Bayern	2014	Bayerische Landesanstalt für Weinbau und Gartenbau	14DEBY-GO- 001-COR-HN	6
85*	Corabel	Frankreich, Poitou-Charentes	2014	SCA Unicoque	14FRPC-VI-001- COR-HN	3
32*	Corabel	Frankreich, Aquitanien	2014	SCA Unicoque	14FRAQ-DO- 001-COR-HN	1, 3, 6
35*	Corabel	Frankreich, Midi- Pyrénées	2014	SCA Unicoque	14FRMP-HG- 001-COR-HN	6
77*	Ennis	Frankreich, Pays de la Loire	2014	SCA Unicoque	14FRPL-SR-001- ENN-HN	4, 6, 7
80*	Ennis	Frankreich, Midi- Pyrénées	2014	SCA Unicoque	14FRMP-TG- 003-ENN-HN	3
46	Ennis	Frankreich, Centre	2014	SCA Unicoque	14FRCE-IN-001- ENN-HN	1, 6
51*	Ennis	Frankreich, Centre	2014	SCA Unicoque	14FRCE-LC-002- ENN-HN	2, 3
59*	Ennis	Frankreich, Midi- Pyrénées	2014	SCA Unicoque	14FRMP-LT- 003-ENN-HN	6

Nr.	Sorte	Herkunft	Ernte- jahr	Lieferant	Barcode	Analyse ^{a)}
142*	Foşa (Yomra)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 002-YOM-HN	1, 2, 3
143*	Foşa (Yomra)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 003-YOM-HN	1, 2, 3
177**	Foşa (Yomra)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 005-YOM-HN	1, 3
178**	Foşa (Yomra)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 001-YOM-HN	1, 3
179**	Foşa (Yomra)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 002-YOM-HN	1
180**	Foşa (Yomra)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 006-YOM-HN	1
181**	Foşa (Yomra)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 003-YOM-HN	1
182**	Foşa (Yomra)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 004-YOM-HN	1
183**	Foşa (Yomra)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 005-YOM-HN	1
184**	Foşa (Yomra)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 006-YOM-HN	1
F10*	Hallsche Riesen	Deutschland, Bayern	2014	Bayerische Landesanstalt für Weinbau und Gartenbau	14DEBY-GO- 001-HRI-HN	1, 2, 3
134*	Karayağlı	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR- 001-KAR-HN	1, 3
133*	Karayağlı	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 001-KAR-HN	1
145*	Karayağlı	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 002-KAR-HN	1
146*	Karayağlı	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 004-KAR-HN	1
147*	Karayağlı	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR- 002-KAR-HN	1
148*	Karayağlı	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR- 003-KAR-HN	1
153*	Karayağlı	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 003-KAR-HN	1
154*	Karayağlı	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 005-KAR-HN	1
185**	Karayağlı	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 006-KAR-HN	1
186**	Karayağlı	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 001-KAR-HN	1

Nr.	Sorte	Herkunft	Ernte- jahr	Lieferant	Barcode	Analyse ^{a)}
187**	Karayağlı	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 002-KAR-HN	1
188**	Karayağlı	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 003-KAR-HN	1
189**	Karayağlı	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 004-KAR-HN	1
190**	Karayağlı	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 005-KAR-HN	1
191**	Karayağlı	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 006-KAR-HN	1
192**	Karayağlı	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 007-KAR-HN	1
193**	Karayağlı	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 007-KAR-HN	1
194**	Karayağlı	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 008-KAR-HN	1
195**	Karayağlı	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 009-KAR-HN	1
196**	Karayağlı	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 010-KAR-HN	1
197**	Karayağlı	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 011-KAR-HN	1
S1*	Mortarella	Italien	2014	August Storck KG	14ITNB-NB-001- MOR-HN	6
106*	Mortarella	Italien, Kampanien	2015	STELLIFERI & ITAVEX s.r.l. Unipersonale	15ITKP-AV-001- MOR-HN	1, 3, 6
107*	Mortarella	Italien, Kampanien	2016	August Storck KG	16ITKP-CT-005- MOR-HN	3, 6
25*	N.A. (AZ)	Aserbaidtschan	2014	Schlüter & Maack GmbH	14AZNB-NB- 001-NB-HN	3
121*	N.A. (AZ)	Aserbaidtschan	2018	August Storck KG	18AZNB-NB- 001-NB-HN	1, 2, 3
122*	N.A. (AZ)	Aserbaidtschan	2018	August Storck KG	18AZNB-NB- 002-NB-HN	1, 2, 3
S4*	N.A. (GE)	Georgien, Gurien	2014	August Storck KG	14GEGU-NB- 002-NB-HN	1, 2, 3, 4
S5*	N.A. (GE)	Georgien, Mingrelien	2014	August Storck KG	14GEME-NB- 003-NB-HN	1, 2, 3, 4
S6*	N.A. (GE)	Georgien, Imeretien	2014	August Storck KG	14GEIM-VA- 004-NB-HN	1, 2, 3, 4
118*	N.A. (GE)	Georgien, Mingrelien	2015	Schlüter & Maack GmbH	15GEME-TS- 001-NB-HN	1, 2, 3

Nr.	Sorte	Herkunft	Ernte- jahr	Lieferant	Barcode	Analyse ^{a)}
162**	N.A. (GE)	Georgien	2015	Schlüter & Maack GmbH	15GENB-NB-001-NB-HN	2
163**	N.A. (GE)	Georgien	2015	Schlüter & Maack GmbH	15GENB-NB-002-NB-HN	2
164**	N.A. (GE)	Georgien	2015	Schlüter & Maack GmbH	15GENB-NB-003-NB-HN	2
165**	N.A. (GE)	Georgien, Mingrelien	2015	Schlüter & Maack GmbH	15GEME-NB-001-NB-HN	2
151*	N.A. (GE)	Georgien, Mingrelien, Tsalenjikha	2018	Schlüter & Maack GmbH	18GEME-TS-001-NB-HN	1, 2, 3, 4
152*	N.A. (GE)	Georgien, Mingrelien, Tsalenjikha	2018	Schlüter & Maack GmbH	18GEME-TS-002-NB-HN	1, 2, 3
13*	N.A. (GE)	Georgien, Gurien	2014	Schlüter & Maack GmbH	14GEGU-NB-001-NB-HN	2, 3
15*	N.A. (GE)	Georgien	2014	Schlüter & Maack GmbH	14GENB-NB-001-NB-HN	2, 3
17*	N.A. (GE)	Georgien, Mingrelien	2014	Schlüter & Maack GmbH	14GEME-NB-002-NB-HN	2
19*	N.A. (GE)	Georgien	2014	Schlüter & Maack GmbH	14GENB-NB-002-NB-HN	2, 3
124*	N.A. (GE)	Georgien	2018	August Storck KG	18GENB-NB-001-NB-HN	1, 2, 3, 4
2*	Negret	Spanien, Katalonien	2014	Coselva SCCL	14ESKA-TA-002-NEG-HN	1, 6
108*	Negret	Spanien, Katalonien	2015	Coselva SCCL	15ESKA-TA-003-NEG-HN	2, 3, 6
116*	Palaz	Türkei	2012	Biozentrum Klein Flottbek	12TRNB-NB-003-PAL-HN	1
117*	Palaz	Türkei, Ordu	2015	Schlüter & Maack GmbH	15TROR-OR-004-PAL-HN	1
128*	Palaz	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR-001-PAL-HN	1, 2, 3
139*	Palaz	Türkei, Ordu	2014	AiF 18040 N (2014-2017)	14TROR-OR-001-PAL-HN	1, 3
140*	Palaz	Türkei, Ordu	2014	AiF 18040 N (2014-2017)	14TROR-OR-002-PAL-HN	1, 2, 3
167**	Palaz	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR-002-PAL-HN	1
168**	Palaz	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR-003-PAL-HN	1, 3

Nr.	Sorte	Herkunft	Ernte- jahr	Lieferant	Barcode	Analyse ^{a)}
169**	Palaz	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR-005-PAL-HN	1
170**	Palaz	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR-006-PAL-HN	1, 3
171**	Palaz	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR-007-PAL-HN	1
172**	Palaz	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR-008-PAL-HN	1
F1*	Pauetet	Deutschland, Bayern	2014	Bayerische Landesanstalt für Weinbau und Gartenbau	14DEBY-GO-001-PAU-HN	6
72*	Pauetet	Frankreich, Aquitanien	2014	SCA Unicoque	14FRAQ-PA-004-PAU-HN	4, 7
31*	Pauetet	Frankreich, Poitou-Charentes	2014	SCA Unicoque	14FRPC-CM-002-PAU-HN	1, 3, 5, 6
33*	Pauetet	Frankreich, Aquitanien	2014	SCA Unicoque	14FRAQ-DO-002-PAU-HN	5, 6
131*	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK-001-MIN-HN	1, 3
149*	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK-006-MIN-HN	1
161*	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK-010-MIN-HN	1
155*	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK-002-MIN-HN	1, 3
157*	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK-004-MIN-HN	1
158*	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK-005-MIN-HN	1
159*	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK-007-MIN-HN	1
160*	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK-008-MIN-HN	1
198**	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK-001-MIN-HN	1
199**	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK-002-MIN-HN	1
200**	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK-003-MIN-HN	1
201**	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK-004-MIN-HN	1
202**	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK-005-MIN-HN	1

Nr.	Sorte	Herkunft	Ernte- jahr	Lieferant	Barcode	Analyse ^{a)}
203**	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 006-MIN-HN	1
204**	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 007-MIN-HN	1
205**	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 008-MIN-HN	1
206**	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 009-MIN-HN	1
207**	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 010-MIN-HN	1
208**	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 011-MIN-HN	1
209**	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 011-MIN-HN	1
210**	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2015	AiF 18040 N (2014-2017)	15TRDU-AK- 012-MIN-HN	1
211**	Sari Yağlı (Mincane)	Türkei, Düzce, Akçakoca	2014	AiF 18040 N (2014-2017)	14TRDU-AK- 012-MIN-HN	1
84*	Segorbe	Frankreich, Midi- Pyrénées	2014	SCA Unicoque	14FRMP-TG- 007-SEG-HN	3
67*	Segorbe	Frankreich, Aquitanien	2014	SCA Unicoque	14FRAQ-LG- 007-SEG-HN	4, 7
34*	Segorbe	Frankreich, Rhône-Alpes	2014	SCA Unicoque	14FRRA-DR- 001-SEG-HN	6
37*	Segorbe	Frankreich, Midi- Pyrénées	2014	SCA Unicoque	14FRMP-HG- 003-SEG-HN	1, 2, 3, 6
45*	Segorbe	Frankreich, Aquitanien	2014	SCA Unicoque	14FRAQ-GI-006- SEG-HN	6
114*	Tombul	Türkei, Ordu	2014	Schlüter & Maack GmbH	14TROR-OR- 002-TOM-HN	1, 6
115*	Tombul	Türkei, Düzce, Akçakoca	2016	Ferrero OHG mbH	16TRDU-AK- 003-TOM-HN	1, 6
119*	Tombul	Türkei, Samsun	2014	Schlüter & Maack GmbH	14TRSA-SA-001- TOM-HN	1, 4, 7
120*	Tombul	Türkei, Samsun	2014	Schlüter & Maack GmbH	14TRSA-CS-001- TOM-HN	1, 6
127*	Tombul	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR- 001-TOM-HN	1
129*	Tombul	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR- 002-TOM-HN	1
141*	Tombul	Türkei, Ordu	2014	AiF 18040 N (2014-2017)	14TROR-OR- 001-TOM-HN	1

Nr.	Sorte	Herkunft	Ernte- jahr	Lieferant	Barcode	Analyse ^{a)}
173**	Tombul	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR-003-TOM-HN	1
174**	Tombul	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR-004-TOM-HN	1
175**	Tombul	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR-005-TOM-HN	1
176**	Tombul	Türkei, Ordu	2015	AiF 18040 N (2014-2017)	15TROR-OR-006-TOM-HN	1
22*	Tonda di Giffoni	Frankreich, Aquitanien	2014	SCA Unicoque	14FRAQ-LG-008-TDG-HN	3, 6
8*	Tonda di Giffoni	Italien, Kampanien	2014	STELLIFERI & ITAVEX s.r.l. Unipersonale	14ITKP-SL-001-TDG-HN	2, 3, 6, 7
F4*	Tonda di Giffoni	Deutschland, Bayern	2014	Bayerische Landesanstalt für Weinbau und Gartenbau	14DEBY-GO-003-TDG-HN	1, 6
F3*	T. G. Romana	Deutschland, Bayern	2014	Bayerische Landesanstalt für Weinbau und Gartenbau	14DEBY-GO-002-TGR-HN	6
1*	T. G. Romana	Italien, Latium	2014	Horst Walberg Trockenfrucht Import GmbH	14ITLA-RO-001-TGR-HN	7
7*	T. G. Romana	Italien, Latium	2014	STELLIFERI & ITAVEX s.r.l. Unipersonale	14ITLA-VT-001-TGR-HN	1, 2, 3, 6
113*	T. G. Romana	Italien, Kampanien	2015	AgroTeamConsulting/ NocciolaRe	15ITKP-VR-001-TGR-HN	6
9*	T. G. Trilobata	Italien, Piemont	2014	STELLIFERI & ITAVEX s.r.l. Unipersonale	14ITPM-AL-001-TGT-HN	1, 3
F6*	T. G. Trilobata	Deutschland, Bayern	2014	Bayerische Landesanstalt für Weinbau und Gartenbau	14DEBY-GO-002-TGT-HN	1
135*	Wildtyp	Türkei, Ordu	2015	AiF 18040 N	15TROR-OR-003-X-HN	1
136*	Wildtyp	Türkei, Ordu	2015	AiF 18040 N	15TROR-OR-001-X-HN	1
137*	Wildtyp	Türkei, Ordu	2015	AiF 18040 N	15TROR-OR-004-X-HN	1
138*	Wildtyp	Türkei, Ordu	2015	AiF 18040 N	15TROR-OR-002-X-HN	1

*: DNA-Isolierung mittels CTAB-Methode, **: DNA-Isolierung mittels *Rapid Extraction*

a) 1: Indels, 2: RFLP, 3: HRMA, 4: ITS, 5: WRKY, 6: SCoT, 7: SCAR

Tabelle 28: Auflistung aller verwendeten verarbeiteten Haselnussproben mit Angabe der Sorte, Herkunft, Lieferant, verwendete Isolierungsmethode und Angabe der Analysenmethoden, für welche die Proben eingesetzt wurden.

Nr.	Beschreibung	Herkunft der Haselnüsse	Jahr	Lieferant	Barcode	Analyse ^{a)}
V1*	Kekse	Frankreich	-	Selbst hergestellt	18FRNB-NB-001-NB-HNP	1
V2*	Geröstete Tombul	Türkei	2018	---	18TRNB-NB-001-TOM-HNG	1
V3*	Geröstete Nocciola Piemonte (g.g.A.)	Italien, Piemont	2018	---	18ITNB-NB-001-TGT-HNG	1, 2, 3
V4*	Geröstete Haselnüsse N.A. (GE)	Georgien	2018	Selbst geröstet	18ITNB-NB-001-TGT-HNG	2, 3
V5*	Geröstete Haselnüsse N.A. (AZ)	Aserbaidshjan	2018	Selbst geröstet	18ITNB-NB-001-TGT-HNG	2, 3
V6*	Schokolade	Türkei	2017	Alfred Ritter GmbH & Co. KG	18FRNB-NB-001-NB-HNP	2, 3

*: DNA-Isolierung mittels CTAB-Methode, **: DNA-Isolierung mittels *Rapid Extraction*

a) 1: Indels, 2: RFLP, 3: HRMA

Tabelle 29: Angaben zu den verwendeten Haselnussblatt-Proben mit Angabe der Sorte, Herkunft, Lieferant und Angabe der Analysenmethoden, für welche die Proben eingesetzt wurden.

Nr.	Sorte bzw. Art	Herkunft	Ernte-jahr	Lieferant	Barcode	Analyse ^{a)}
B1	Ata Baba	Aserbaidshjan, Balakan	2016	Gebr. Jancke GmbH	16AZNB-BK-001-ATA-HNB	5, 8, 9
B2	Ata Baba	Aserbaidshjan, Zaqatala	2016	Gebr. Jancke GmbH	16AZNB-ZA-001-ATA-HNB	5, 9
B3	Barcelona	Frankreich, Aquitanien	2018	SCA Unicoque, Cancon	18FRAQ-CA-001-BAR-HNB	8
B4	Butler	Frankreich, Aquitanien	2018	SCA Unicoque, Cancon	18FRAQ-CA-001-BUT-HNB	8
B5	Corabel	Frankreich, Aquitanien	2018	SCA Unicoque, Cancon	18FRAQ-CA-001-COR-HNB	8
B6	Corabel	Schweiz, Hörhausen	2017	Nuss – Baumschule Gubler GmbH	17SWTU-HH-001-COR-HNB	8
B7	Hallsche Riesen	Deutschland, Bayern	2014	Bayerische Landesanstalt für Weinbau und Gartenbau	14DEBY-GO-001-HRI-HNB	8, 9
B8	Negret	Deutschland, Bayern	2014	Bayerische Landesanstalt für Weinbau und Gartenbau	14DEBY-GO-001-NEG-HNB	8, 9
B9	Palaz	Türkei, Ordu	2016	Gebr. Jancke GmbH	16TROR-NB-001-PAL-HNB	5, 8, 9
B10	Pauetet	Frankreich, Aquitanien	2016	SCA Unicoque, Cancon	16FRAQ-CA-001-PAU-HNB	8

Nr.	Sorte bzw. Art	Herkunft	Ernte-jahr	Lieferant	Barcode	Analyse ^{a)}
B11	Pauetet	Frankreich, Aquitanien	2016	SCA Unicoque, Cancon	16FRAQ-CA-002-PAU-HNB	5, 9
B12	Pauetet	Deutschland, Bayern	2014	Bayerische Landesanstalt für Weinbau und Gartenbau	14DEBY-GO-001-PAU-HNB	5, 9
B13	Segorbe	Deutschland, Bayern	2014	Bayerische Landesanstalt für Weinbau und Gartenbau	14DEBY-GO-001-SEG-HNB	8
B14	Segorbe	Frankreich, Aquitanien	2016	SCA Unicoque, Cancon	16FRAQ-CA-001-SEG-HNB	9
B15	Segorbe	Frankreich, Aquitanien	2016	SCA Unicoque, Cancon	16FRAQ-CA-002-SEG-HNB	9
B16	Tonda di Giffoni	Deutschland, Bayern	2014	Bayerische Landesanstalt für Weinbau und Gartenbau	14DEBY-GO-001-TDG-HNB	8, 9
B17	Tonda di Giffoni	Frankreich, Aquitanien	2016	SCA Unicoque, Cancon (FR)	16FRAQ-CA-001-TDG-HNB	8, 9
B18	Tonda di Giffoni	Frankreich, Aquitanien	2016	SCA Unicoque, Cancon	16FRAQ-CA-002-TDG-HNB	8, 9
B19	Tonda Gentile Romana	Deutschland, Bayern	2014	Bayerische Landesanstalt für Weinbau und Gartenbau	14DEBY-GO-001-TGR-HNB	8, 9
B20	Tonda Gentile Trilobata	Deutschland, Bayern	2014	Bayerische Landesanstalt für Weinbau und Gartenbau	14DEBY-GO-001-TGT-HNB	8
B21	<i>C. avellana</i>	Deutschland, Hamburg	2014	Loki-Schmidt Garten, Botanischer Garten Klein Flottbek der Universität Hamburg	14DEHH-KL-001-NB-HNB	7, 8
B22	<i>C. maxima</i> Purpurea	Deutschland, Hamburg	2014	Loki-Schmidt Garten, Botanischer Garten Klein Flottbek der Universität Hamburg	14DEHH-KL-001-PUR-HNB	7
B23	<i>C. maxima</i> Lambert Filbert	Deutschland, Bayern	2014	Bayerische Landesanstalt für Weinbau und Gartenbau	14DEBY-GO-001-LAF-HNB	7
B24	<i>C. colurna</i>	Deutschland, Hamburg	2014	Loki-Schmidt Garten, Botanischer Garten Klein Flottbek der Universität Hamburg	14DEHH-KL-002-NB-HNB	7

a) 5: WRKY, 7: SCAR, 8: Sequenzierung cpGenom, 9: Sequenzvalidierung

9.7 Kennzeichnung des Probenmaterials

Die Kennzeichnung der Proben erfolgte über Barcodes, die nach dem Schema in Abbildung 54 aufgebaut sind und alle notwendigen Informationen die Probe betreffend enthalten.



Abbildung 54: Schematischer Aufbau der Barcodes für die Probenkennzeichnung.

Die für die Barcodes verwendeten Abkürzungen sind in Tabelle 30 und Tabelle 31 erläutert.

Tabelle 30: Für die Barcodes zur Probenkennzeichnung verwendete Abkürzungen für Staat, Bundesland/Provinz und Region.

Staat		Bundesland/Provinz		Region	
Abk.	Name	Abk.	Name	Abk.	Name
AZ	Aserbaidshon	AQ	Aquitonien	AK	Akçakoca
DE	Deutschland	BY	Bayern	AL	Alba
ES	Spanien	CE	Centre	AV	Avellino
FR	Frankreich	DU	Düzce	BK	Balakan
GE	Georgien	GU	Gurien	CA	Cancon
IT	Italien	HH	Hamburg	CM	Charente-Maritime
SW	Schweiz	IM	Imeretien	CS	Carsamba
TR	Türkei	KA	Katalonien	CT	Caserta
		KP	Kampanien	DO	Dordogne
		LA	Latium	DR	Drôme
		ME	Mingrelien	GI	Gironde
		MP	Midi-Pyrénées	GO	Gonnorsdorf
		NB	unbekannt	HH	Hörhausen
		OR	Ordu	HG	Haute-Garonne
		PC	Poitou-Charentes	KL	Klein Flottbek
		PL	Pays de la Loire	LG	Lot-et-Garonne
		PM	Piemont	LQ	Loire-Atlantique
		RA	Rhône-Alpes	LT	Lot
		SA	Samsun	NB	Unbekannt
		TU	Thurgau	OR	Ordu
		ZA	Zaqatala	PA	Pyrénées-Atlantique
				RO	Rom
				SA	Samsun
				SL	Salerno
				SR	Sarthe
				TA	Tarragonien
				TG	Tarn-et-Garonne
				TS	Tsaljikha
				VA	Vani

Staat		Bundesland/Provinz		Region	
Abk.	Name	Abk.	Name	Abk.	Name
				VI	Vienne
				VR	Vairano
				VT	Viterbo

Tabelle 31: Für die Barcodes zur Probenkennzeichnung verwendete Abkürzungen für Sorte und Form.

Sorte		Form	
Abk.	Name	Abk.	Name
ATA	Ata Baba	HN	Haselnüsse
BAR	Barcelona (Fertile de Coutard)	HNB	Haselnussblatt
BUT	Butler	HNG	Haselnüsse geröstet
CAK	Çakildak (Delisava)	HNP	Haselnüsse verarbeitet
COR	Corabel		
ENN	Ennis		
HRI	Hallsche Riesen		
KAR	Karayağlı		
MIN	Sari Yağlı (Mincane)		
MOR	Mortarella		
NEG	Negret		
PAL	Palaz		
PAU	Pauetet		
SEG	Segorbe		
TDG	Tonda di Giffoni		
TGR	Tonda Gentile Romana		
TGT	Tonda Gentile Trilobata		
TOM	Tombul		
NB	Unbekannt		
X	Wildtyp		
YOM	Yomra (Fosa)		

9.8 Temperaturprogramme

Tabelle 32: Standard-Temperaturprogramm für eine PCR mit einer *Taq* DNA Polymerase (z.B. Biozym *Taq* DNA Polymerase, *DreamTaq* DNA Polymerase).

Prozess	Temperatur / Zyklus	Zeit
Deckelheizung	95 °C	Dauerhaft
Denaturierung	95 °C	5 min
Beginn Zyklus (x35)		
Denaturierung	95 °C	20-40 s*
Annealing	variabel	20-40 s*
Elongation	72 °C	20-40 s*
Ende Zyklus		
Finale Elongation	72 °C	5 min
Kühlung	4 °C	Bis Entnahme

* variiert je nach Amplifikatlänge, bei kurzen Amplifikaten wurden die Schritte verkürzt und bei längeren PCR-Produkten verlängert

Tabelle 33: Standard-Temperaturprogramm für eine PCR mit einer *High-Fidelity* DNA Polymerase (z.B. *S7 Fusion High-Fidelity* DNA Polymerase).

Prozess	Temperatur / Zyklus	Zeit
Deckelheizung	98 °C	Dauerhaft
Denaturierung	98 °C	30 s
Beginn Zyklus (x25)		
Denaturierung	98 °C	7 s
Annealing	variabel	20 s
Elongation	72 °C	30 s
Ende Zyklus		
Finale Elongation	72 °C	5 min
Kühlung	4 °C	Bis Entnahme

Tabelle 34: Temperaturprogramm für die HRMA mit integrierter Schmelzkurve im Anschluss an die PCR.

Prozess	Temperatur / Zyklus	Zeit
Deckelheizung	95 °C	Dauerhaft
Denaturierung	95 °C	5 min
Beginn Zyklus (x40)		
Denaturierung	95 °C	7 s
Annealing	52,2 °C	15 s
Elongation	70 °C	20 s
Ende Zyklus		
Finale Elongation	70 °C	5 min
Schmelzkurve	65-80 °C	Erhöhung um 0,2 °C/0,02 min
Kühlung	4 °C	Bis Entnahme

Tabelle 35: Temperaturprogramm für die Amplifikation der WRKY-Gene mit Hilfe der degenerierten sowie der selbst designten sequenzspezifischen Primer.

Prozess	Temperatur / Zyklus	Zeit
Deckelheizung	99 °C	Dauerhaft
Denaturierung	94 °C	5 min
Beginn Zyklus (x35)		
Denaturierung	94 °C	30 s
Annealing	50 °C	45 s
Elongation	72 °C	2,5 min
Ende Zyklus		
Finale Elongation	72 °C	20 min
Kühlung	4 °C	Bis Entnahme

Tabelle 36: Temperaturprogramm der PCR für die Analyse der SCoT Polymorphismen nach Collard und Mackill (2009).^[163]

Prozess	Temperatur / Zyklus	Zeit
Deckelheizung	94 °C	Dauerhaft
Denaturierung	94 °C	3 min
Beginn Zyklus (x35)		
Denaturierung	94 °C	1 min
Annealing	variabel	1 min
Elongation	72 °C	2 min
Ende Zyklus		
Finale Elongation	72 °C	5 min
Kühlung	4 °C	Bis Entnahme

9.9 Parameter für die statistische Bewertung von Kalibriergeraden

Im Folgenden sind die Parameter und die zugrunde liegenden Berechnungen aufgeführt, die für die statistische Bewertung der Regressionsgeraden eingesetzt wurden.

Lineare Regression:

Ein lineares Regressionsmodell ermöglicht die Darstellung eines linearen Zusammenhangs zweier Variablen auf Basis der kleinsten Fehlerquadrate. Die lineare Regression kann mit Hilfe folgender Formel dargestellt werden:

$$y = m \times x + b \quad \text{Gleichung 3}$$

- y Zielgröße (z.B. relative Area)
- m Steigung der Regressionsgerade (Empfindlichkeit)
- x Einflussgröße (z.B. Anteil georgischer Haselnüsse)
- b Achsenabschnitt der Regressionsgerade

Korrelationskoeffizient:

Der Korrelationskoeffizient r stellt ein Maß für den linearen Zusammenhang zweier Messwerte dar. Er kann mit der Funktion KORREL von Microsoft Excel oder nach Gleichung 4 berechnet werden.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x}) \times (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \times \sum_{i=1}^n (y_i - \bar{y})^2}} \quad \text{Gleichung 4}$$

r	Korrelationskoeffizient
x_i	Einflussgröße an i -ter Stelle
\bar{x}	Mittelwert der Einflussgrößen
y_i	Zielgröße an i -ter Stelle
\bar{y}	Mittelwert der Zielgröße

Standardabweichung:

Die Standardabweichung stellt ein Maß für die Streuung von Messwerten hinsichtlich des zugehörigen Mittelwerts dar. Sie kann mit Hilfe der Formel STABW.N von Microsoft Excel oder nach Gleichung 5 berechnet werden.

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} \quad \text{Gleichung 5}$$

σ	Standardabweichung
x	Einflussgröße
\bar{x}	Mittelwert der Einflussgrößen
n	Stichprobenumfang

Variationskoeffizient:

Der Variationskoeffizient stellt ein Maß für die relative Streuung von Messwerten bezüglich des zugehörigen Mittelwerts dar. Er kann nach folgender Gleichung berechnet werden.

$$\text{VarK} = \frac{\sigma}{x} \times 100 \quad \text{Gleichung 6}$$

VarK	Variationskoeffizient [%]
σ	Standardabweichung
x	Stichprobenmittelwert

Reststandardabweichung:

Die Reststandardabweichung s_y ist ein Maß für die Streuung von Messwerten um eine Regressionsgerade. Sie kann mit Hilfe von Microsoft Excel und der Funktion STFEHLERYX berechnet werden. Die zugrunde liegende Gleichung lautet wie folgt:

$$s_y = \sqrt{\frac{1}{(n-2)} \times \left[\sum_{i=1}^n (y_i - \bar{y})^2 - \frac{[\sum_{i=1}^n (x_i - \bar{x}) \times (y_i - \bar{y})]^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right]}$$
Gleichung 7

s_y	Reststandardabweichung
n	Anzahl der Messwerte
x_i	Einflussgröße an i-ter Stelle
\bar{x}	Mittelwert der Einflussgrößen
y_i	Zielgröße an i-ter Stelle
\bar{y}	Mittelwert der Zielgrößen

Verfahrensstandardabweichung:

Die Verfahrensstandardabweichung ist ein Maß für die Streuung einer Methode und wird nach Gleichung 8 berechnet.

$$s_{x0} = \frac{s_y}{b}$$
Gleichung 8

s_{x0}	Verfahrensstandardabweichung
s_y	Reststandardabweichung
b	Empfindlichkeit (Steigung der Regressionsgerade)

Verfahrensvariationskoeffizient:

Der Verfahrensvariationskoeffizient oder auch relative Verfahrensstandardabweichung dient als Präzisionsmaß und wird wie folgt berechnet:

$$v_{x0} = \frac{s_{x0}}{\bar{x}} \times 100\%$$
Gleichung 9

v_{x0}	Verfahrensvariationskoeffizient [%]
s_{x0}	Verfahrensstandardabweichung
\bar{x}	Mittelwert der Einflussgrößen

9.10 Ergänzungen zum Abschnitt 6 – Ergebnisse und Diskussion

9.10.1 Sequenzierung des cpGenoms mittels NGS

In der folgenden Tabelle sind alle Haselnussproben aufgelistet, die mittels NGS sequenziert wurden. Die Rohdaten wurden in der Datenbank des *National Center for Biotechnology Information* (NCBI) veröffentlicht. Die zugehörigen *Accession numbers* sind der Tabelle 37 zu entnehmen.

Tabelle 37: *Accession numbers* der veröffentlichten Rohdaten der Sequenzierung des cpGenoms von 15 Haselnussproben.

Haselnussorte	Probennr.	<i>Accession number</i>
Ata Baba	B1	SAMN1738335
Barcelona	B3	SAMN1738336
Butler	B4	SAMN1738337
Corabel	B5	SAMN1738338
Hallsche Riesen	B7	SAMN1738339
Negret	B8	SAMN1738340
Palaz	B9	SAMN1738341
Pauetet	B10	SAMN1738342
Segorbe	B13	SAMN1738343
Tonda di Giffoni	B16	SAMN1738344
Tonda di Giffoni	B17	SAMN1738345
Tonda di Giffoni	B18	SAMN1738346
Tonda Gentile Romana	B19	SAMN1738347
Tonda Gentile Trilobata	B20	SAMN1738348
N.A.	B21	SAMN1738349

9.10.2 Identifizierung und Validierung von Sequenzunterschieden

Im Folgenden sind vier Abschnitte des cpGenoms dargestellt, die mittels Sanger-Sequenzierung resequenziert wurden, um potentielle Sequenzunterschiede zu überprüfen. Die aufgeführten Sequenzabschnitte stammen von der veröffentlichten Sequenz einer *C. avellana* unbekannter Sorte (*Accession No.* KX822768.2). Darauf basierende Primerbindungsstellen sind entsprechend gekennzeichnet.

Bereich A:

```

5' ATCCACTTGGCTACATCCGC CCCTATACTCCATTTAGTACAAATTACATGGGGATTCATCTCTACCATTTCATCATG
   A_FW
TTTACCTTTCTTCTAAGATACAAGAGACAGAAAAGAAATTGAAACCTTTATCTTTTCATTTATCTTTTCATTTATCT
TTTCATTTTGAAATAAAAAAGATAATAATAATTAATAATACTAAACCAAACCTTCACAAAAAACGGATAAGAACATA
TAAAAAAATGACTGCATAACGAAAATATGATACTCAAATACTTTTTTCGTATGTTAAAGGAAAGAAAGAAAAGACT
TATGTAAGTAAAAGACAACATAATAGTAATAAAAGGCTAATAAAGGAGCAATAGCAATCCTCTTGAAAGAACAAGA
AGTTGATTATTGCTCCTTTATTTTCAAAAACCTCATATATACTAATACTAAAATACTAATACTAAAACCGAAGTCTTA
TCCATTTGTAGATGGAGCTTCGACAGCAGCTAAGTCTAGAGGGAAATTATGAGCATTACGTTTCATGCATAACTTCCA
TACCAAGGTTAGCGCGGTTAATAATATCAGCCCAGGTATTAATTACACGCCCTTGACTATCAACTACGGATTGGTTG
AAATTGAAACCATTTAAGTTGAAAGCCATAGTCTGATACCTAAAGCGGTGAACCAGATACCTACTACAGGCCAAGC
   A_RV
CGCTA 3'

```

Abbildung 55: Region A (Pos. 76-773 bp) des cpGenoms einer *C. avellana* (*Accession No.* KX822768.2). Die Bindungsstellen des Primerpaares A_FW/RV sind grau hinterlegt.

Bereich B:

```

5' GAGTACTCTACCGTTGAGTT AGCAACCCGAATCAAAATGTATAGATCTAATCGTAATCAAAAAAGAGATTGAATTC
   B.1_FW
ACAACGCAATCAAAATATTAATAAAAAAAATATTTCTAATCGATTCTGAACATAACATAAAAATGAACATATCAG
ATGCAAATACAATGACGTCTAAGATAAGAATATAGATTAGGATAAAAATATAGATAAAAATAAAATTTATA GACTAC
CACATATTTTGTTCGTATG TTATCCATAGAACAGAGATAAATATATACATTTATT CACGATCGGATTATATTTGTTT
   B.2_FW                                     B.2_RV
GATATACTGTTGTCAATATAAAAGTTAATGTTGAGAAAAGAACATAATGTAGAAAACCATAAATTCAAATAAAAAAT
TATTCAATATTTTCTTATTAATTTATTAATTTCAATATAATATTTAATTACTATAACTATAAATAAATAGAAATGA
AAAAATAAAAAAACTAATGACTTGTATTAAATTAGCACTACATATTTAATAAAGATAAATAAAAAAGGTATAGGCGT
AAAAAAAATTGGAGAGAAGTATAAAAAAATCTTGATTGGGTTTACGCAATCA 3'
   B.1_RV

```

Abbildung 56: Region B (Pos. 4454-5045 bp) des cpGenoms einer *C. avellana* (*Accession No.* KX822768.2). Die Bindungsstellen des Primerpaares B.1_FW/RV sind grau und die des Primerpaares B.2_FW/RV rot hinterlegt. Die Position, an der bei manchen Sorten eine Insertion eingefügt ist, ist mit einem Pfeil gekennzeichnet.

Bereich C:

```

5' TGGGCTCTGTAAGAATTTGTCGATATTACTGATTAAAGATTCTTATTTATAGATGGTTTATAGGAATAAATTGATAT
   C_FW
TTCTTCTTCTACAGAGAAAGGCTTATTACAAGTCACAACATACGAGCCATTAAAATATCTTTTTTTTTTATTTTA
ATTCCTCGAACACAAGACATGATCAATTTAGATATAAATATTTTTAATTATATCTAAATAAGATTATATAATTTAT
CAGATCATCGATTTCATGTACCAAATATTTCCATATCGATTTTTTATTATTAATTAATACAATATTAATAAAAAAAT
                                     ↓
GAAAAATGAAAGTATAACGTAATAAATATATAATACTATAGTAGTTAATACACATATAAATTTGAAGGAATACAT
AAAAATATTTTTTTTTTATCAATCTCACAACGAGATTCAAGAATAAGATTAGTTGATGGAACGAGAGGAATATGTCT
GGGGATCAACTGGTAGCGAGAAAAGGAATCAC 3'
               C_RV

```

Abbildung 57: Region C (Pos. 35742-36235 bp) des cpGenoms einer *C. avellana* (Accession No. KX822768.2). Die Bindungsstellen des Primerpaares C_FW/RV sind grau hinterlegt. Die Position, an der bei manchen Sorten eine Insertion eingefügt ist, ist mit einem Pfeil gekennzeichnet.

Bereich D:

```

5' AAAGGATTGAGCCGAATACAAAACAAAGATACTATTGTACATATTTGTATATATTTTATAGATACATACTTATCTAG
   D.1_FW
ATATACAGGATCTTAAATAAAAAAAGCTATATCTAAAACAAAACAACTTAAACGTTTTTATTGTTGTGGTCCA
CAATGAATTCTATGGGTCCTTAGGGTTGGTGTTCCTTTCTATCCCGTAGTTTCCGATCATGGATCGAGCCACATA
                                     D.2_FW
TCACAACCTCTTCTACCCATCCCGTATATTGTCTTTTCGTTCTGTGTTGGGCTAGAACTTATTAGTAGACGAGAT
TTTACGAAAAAAGTTCTTTATATTCTTTTATATTCATAGGATATAACAATTATTTTCATTTTTCGATGTTAATTTGAT
ACAACAAAACACGGCGGGATCACTATTTTAAAATTGAAAAATAATGAAATCATTAATAAAAAAGGATTCCATCATATTAA
   D.HRMA_FW                                     D.HRMA_RV
GTGATACTCCTGGTTCTCACAAAAGAGAATCATTTTTTTTCAATTTCAACACCTACTTTTTTTGTTAATATAATCCT
AGTAATTGGATTTATATGCATATTCTGATAGTAAATGAAATATTCAAATGATTTTTTATCGAATGACTATTCATCTA
TTGTATTTTCATGTAAATAGGGGCAAAAAGGCTTTTATGGAAAAAGGGTGGTTCAAATTCGATGTTGTCCAAAAGAAAAG
                                     D.1_RV
TTAGAATACAGTTGTAGGCTAAGTAAATCAATGGATAGTCTTGGTCCTATTGAAAATACCGGTGTAAGTGAAGACCC
GATTACAAATGATATAGATAAAAAACACTCTTAGTTGGAGCGATAGTGACAATTCTAG 3'
                                     D.2_RV

```

Abbildung 58: Region D (Pos. 60677-61503 bp) des cpGenoms einer *C. avellana* (Accession No. KX822768.2). Die Bindungsstellen des Primerpaares D.1_FW/RV sind grau, die des Primerpaares D.2_FW/RV rot und die des Primerpaares D.HRMA_FW/RV blau hinterlegt. Die Erkennungssequenz des Restriktionsenzym *MfeI* (5'-CAATTG-3') ist grün hervorgehoben. Aufgrund eines SNPs ist die Restriktionsschnittstelle im vorliegenden Beispiel nicht intakt.

Die folgende Tabelle enthält eine Übersicht aller innerhalb der Bereiche A-D identifizierten Sequenzunterschiede.

Tabelle 38: Ergebnis der Resequenzierung vier ausgewählter Bereiche (A-D) des cpGenoms.

	Bereich A		Bereich B			Bereich C				Bereich D				
	SNP 1	Indel 1	SNP 2*	SNP 3	Indel 2	Indel 3	SNP 4*	Indel 4	SNP 5	SNP 6	SNP 7	Indel 5		
Ata Baba	T	Insert (118 bp)	T	G	Deletion	Insert (115 bp)	G	Deletion	T	A	T	Deletion		
Palaz	T	Insert (118 bp)	T	G	Deletion	Insert (115 bp)	G	Deletion	T	A	T	Deletion		
Segorbe	G	Deletion	-	G	Insert (T)	Deletion	-	Insert (A)	A	A	G	Deletion		
Pauetet	G	Deletion	-	G	Insert (T)	Deletion	-	Insert (A)	A	A	G	Deletion		
Hallische Riesen	G	Deletion	-	G	Insert (T)	Deletion	-	Insert (A)	A	A	G	Deletion		
Negret	G	Deletion	-	G	Insert (T)	Deletion	-	Insert (A)	A	A	G	Deletion		
Tonda di Giffoni	G	Deletion	-	G	Insert (T)	Deletion	-	Insert (A)	A	A	G	Deletion		
T. G. Romana	G	Deletion	-	G	Insert (T)	Deletion	-	Insert (A)	A	A	G	Deletion		
T. G. Trilobata	G	Deletion	-	G	Insert (T)	Deletion	-	Insert (A)	A	A	G	Deletion		
N.A. (GE)	T	Insert (118 bp)	G	T	Insert (T)	Insert (115 bp)	T	Insert (A)	T	C	T	Insert (T)		

* SNP befindet sich innerhalb des großen Indels des jeweiligen Bereichs

Im Folgenden sind Alignments der mittels Sanger-Sequenzierung erhaltenen Sequenzen für die Bereiche A-D aufgeführt, denen die Unterschiede entnommen werden können.

Alignment Bereich A:

Hallsche R.	(EU)	CTACCATTCATCATGTTTACCTTTCTTCTAAGATACAAGAGACAGAAAAGAAATTGAAAC
Tonda di Gif.	(EU)	CTACCATTCATCATGTTTACCTTTCTTCTAAGATACAAGAGACAGAAAAGAAATTGAAAC
Ata Baba	(AZ)	CTACCATTCATCATGTTTACCTTTCTTCTAAGATACAAGAGACAGAAAAGAAATTGAAAC
Pauetet	(EU)	CTACCATTCATCATGTTTACCTTTCTTCTAAGATACAAGAGACAGAAAAGAAATTGAAAC
Segorbe	(EU)	CTACCATTCATCATGTTTACCTTTCTTCTAAGATACAAGAGACAGAAAAGAAATTGAAAC
T.G.Trilobata	(EU)	CTACCATTCATCATGTTTACCTTTCTTCTAAGATACAAGAGACAGAAAAGAAATTGAAAC
Negret	(EU)	CTACCATTCATCATGTTTACCTTTCTTCTAAGATACAAGAGACAGAAAAGAAATTGAAAC
Palaz	(TR)	CTACCATTCATCATGTTTACCTTTCTTCTAAGATACAAGAGACAGAAAAGAAATTGAAAC
T.G.Romana	(EU)	CTACCATTCATCATGTTTACCTTTCTTCTAAGATACAAGAGACAGAAAAGAAATTGAAAC
N.A.	(GE)	CTACCATTCATCATGTTTACCTTTCTTCTAAGATACAAGAGACAGAAAAGAAATTGAAAC *****
← SNP 1		
Hallsche R.	(EU)	CTTTATCTTTTCATTTATCTTTTCATTTATCTTTTCATTTTGAATAAAAAAGATAATAA
Tonda di Gif.	(EU)	CTTTATCTTTTCATTTATCTTTTCATTTATCTTTTCATTTTGAATAAAAAAGATAATAA
Ata Baba	(AZ)	CTTTATCTTTTCATTTATCTTTTCATTTATCTTTTCATTTTGAATAAAAAAGATAATAA
Pauetet	(EU)	CTTTATCTTTTCATTTATCTTTTCATTTATCTTTTCATTTTGAATAAAAAAGATAATAA
Segorbe	(EU)	CTTTATCTTTTCATTTATCTTTTCATTTATCTTTTCATTTTGAATAAAAAAGATAATAA
T.G.Trilobata	(EU)	CTTTATCTTTTCATTTATCTTTTCATTTATCTTTTCATTTTGAATAAAAAAGATAATAA
Negret	(EU)	CTTTATCTTTTCATTTATCTTTTCATTTATCTTTTCATTTTGAATAAAAAAGATAATAA
Palaz	(TR)	CTTTATCTTTTCATTTATCTTTTCATTTATCTTTTCATTTTGAATAAAAAAGATAATAA
T.G.Romana	(EU)	CTTTATCTTTTCATTTATCTTTTCATTTATCTTTTCATTTTGAATAAAAAAGATAATAA
N.A.	(GE)	CTTTATCTTTTCATTTATCTTTTCATTTATCTTTTCATTTTGAATAAAAAAGATAATAA *****
Hallsche R.	(EU)	TAATTAATAATACTAAACCAAACCTTCACAAAAACGGATAAGAACATATAAAAAAATG
Tonda di Gif.	(EU)	TAATTAATAATACTAAACCAAACCTTCACAAAAACGGATAAGAACATATAAAAAAATG
Ata Baba	(AZ)	TAATTAATAATACTAAACCAAACCTTCACAAAAACGGATAAGAACATATAAAAAAATG
Pauetet	(EU)	TAATTAATAATACTAAACCAAACCTTCACAAAAACGGATAAGAACATATAAAAAAATG
Segorbe	(EU)	TAATTAATAATACTAAACCAAACCTTCACAAAAACGGATAAGAACATATAAAAAAATG
T.G.Trilobata	(EU)	TAATTAATAATACTAAACCAAACCTTCACAAAAACGGATAAGAACATATAAAAAAATG
Negret	(EU)	TAATTAATAATACTAAACCAAACCTTCACAAAAACGGATAAGAACATATAAAAAAATG
Palaz	(TR)	TAATTAATAATACTAAACCAAACCTTCACAAAAACGGATAAGAACATATAAAAAAATG
T.G.Romana	(EU)	TAATTAATAATACTAAACCAAACCTTCACAAAAACGGATAAGAACATATAAAAAAATG
N.A.	(GE)	TAATTAATAATACTAAACCAAACCTTCACAAAAACGGATAAGAACATATAAAAAAATG *****
Hallsche R.	(EU)	ACTGCATAACGAAAATATGATACTCAAATACTTTTTCGTATGTTAAAGGAAAGAAAGAAA
Tonda di Gif.	(EU)	ACTGCATAACGAAAATATGATACTCAAATACTTTTTCGTATGTTAAAGGAAAGAAAGAAA
Ata Baba	(AZ)	ACTGCATAACGAAAATATGATACTCAAATACTTTTTCGTATGTTAAAGGAAAGAAAGAAA
Pauetet	(EU)	ACTGCATAACGAAAATATGATACTCAAATACTTTTTCGTATGTTAAAGGAAAGAAAGAAA
Segorbe	(EU)	ACTGCATAACGAAAATATGATACTCAAATACTTTTTCGTATGTTAAAGGAAAGAAAGAAA
T.G.Trilobata	(EU)	ACTGCATAACGAAAATATGATACTCAAATACTTTTTCGTATGTTAAAGGAAAGAAAGAAA
Negret	(EU)	ACTGCATAACGAAAATATGATACTCAAATACTTTTTCGTATGTTAAAGGAAAGAAAGAAA
Palaz	(TR)	ACTGCATAACGAAAATATGATACTCAAATACTTTTTCGTATGTTAAAGGAAAGAAAGAAA
T.G.Romana	(EU)	ACTGCATAACGAAAATATGATACTCAAATACTTTTTCGTATGTTAAAGGAAAGAAAGAAA
N.A.	(GE)	ACTGCATAACGAAAATATGATACTCAAATACTTTTTCGTATGTTAAAGGAAAGAAAGAAA *****
Hallsche R.	(EU)	AAGACTTATGTAAGTAAAAGACAACCTAAATAGTAATTAAGGCTAATAAAGGAGCAATA
Tonda di Gif.	(EU)	AAGACTTATGTAAGTAAAAGACAACCTAAATAGTAATTAAGGCTAATAAAGGAGCAATA
Ata Baba	(AZ)	AAGACTTATGTAAGTAAAAGACAACCTAAATAGTAATTAAGGCTAATAAAGGAGCAATA
Pauetet	(EU)	AAGACTTATGTAAGTAAAAGACAACCTAAATAGTAATTAAGGCTAATAAAGGAGCAATA
Segorbe	(EU)	AAGACTTATGTAAGTAAAAGACAACCTAAATAGTAATTAAGGCTAATAAAGGAGCAATA
T.G.Trilobata	(EU)	AAGACTTATGTAAGTAAAAGACAACCTAAATAGTAATTAAGGCTAATAAAGGAGCAATA
Negret	(EU)	AAGACTTATGTAAGTAAAAGACAACCTAAATAGTAATTAAGGCTAATAAAGGAGCAATA
Palaz	(TR)	AAGACTTATGTAAGTAAAAGACAACCTAAATAGTAATTAAGGCTAATAAAGGAGCAATA
T.G.Romana	(EU)	AAGACTTATGTAAGTAAAAGACAACCTAAATAGTAATTAAGGCTAATAAAGGAGCAATA
N.A.	(GE)	AAGACTTATGTAAGTAAAAGACAACCTAAATAGTAATTAAGGCTAATAAAGGAGCAATA *****
Hallsche R.	(EU)	GCAATCCTCTTGAAAGAACAAGAAGTTGATTATTGCTCCTTTATTTTCAAAAACTCATAT
Tonda di Gif.	(EU)	GCAATCCTCTTGAAAGAACAAGAAGTTGATTATTGCTCCTTTATTTTCAAAAACTCATAT
Ata Baba	(AZ)	GCAATCCTCTTGAAAGAACAAGAAGTTGATTATTGCTCCTTTATTTTCAAAAACTCATAT
Pauetet	(EU)	GCAATCCTCTTGAAAGAACAAGAAGTTGATTATTGCTCCTTTATTTTCAAAAACTCATAT
Segorbe	(EU)	GCAATCCTCTTGAAAGAACAAGAAGTTGATTATTGCTCCTTTATTTTCAAAAACTCATAT
T.G.Trilobata	(EU)	GCAATCCTCTTGAAAGAACAAGAAGTTGATTATTGCTCCTTTATTTTCAAAAACTCATAT
Negret	(EU)	GCAATCCTCTTGAAAGAACAAGAAGTTGATTATTGCTCCTTTATTTTCAAAAACTCATAT
Palaz	(TR)	GCAATCCTCTTGAAAGAACAAGAAGTTGATTATTGCTCCTTTATTTTCAAAAACTCATAT
T.G.Romana	(EU)	GCAATCCTCTTGAAAGAACAAGAAGTTGATTATTGCTCCTTTATTTTCAAAAACTCATAT
N.A.	(GE)	GCAATCCTCTTGAAAGAACAAGAAGTTGATTATTGCTCCTTTATTTTCAAAAACTCATAT *****

Hallsche R.	(EU)	ATACTAATACTAAAATACTAATACTAAAACCGAAGTNTATCCNNNNNGNNGNNNNN-----
Tonda di Gif.	(EU)	ATACTAATACTAAAATACTAATACTAAAACCGAAGTNTATCCNNNNNGNNGNNNNN-----
Ata Baba	(AZ)	ATACTAATACTAAAATACTAATACTAAAACCGAAGTNTATCCNNNNNNNNNNNNNN-----
Pauetet	(EU)	ATACTAATACTAAAATACTAATACTAAAACCGAAGTNTNTCCNNNNNNNNNNNNNNN---
Segorbe	(EU)	ATACTAATACTAAAATACTAATACTAAAACCGAAGTCTTATCCATTTGTAGATGGAGCTT
T.G.Trilobata	(EU)	ATACTAATACTAAAATACTAATACTAAAACCGAAGTCTTATCCATTTGTAGATGGAGCTT
Negret	(EU)	ATACTAATACTAAAATACTAATACTAAAACCGAAGTCTTATCCATTTGTAGATGGAGCTT
Palaz	(TR)	ATACTAATACTAAAATACTAATACTAAAACCGAAGTCTTATCCATTTGTAGATGGAGCTT
T.G.Romana	(EU)	ATACTAATACTAAAATACTAATACTAAAACCGAAGTCTTATCCATTTGTAGATGGAGCTT
N.A.	(GE)	ATACTAATACTAAAATACTAATACTAAAACCGAAGTCTTATCCATTTGTAGATGGAGCTT

Alignment Bereich B:

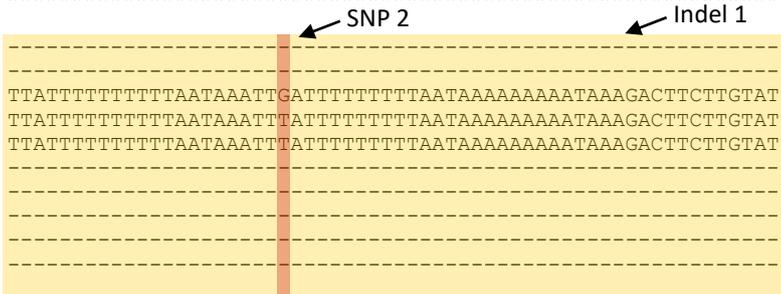
Segorbe	(EU)	TTAGCAACCCGAATCAAAATGTATAGATCTAATCGTAATCAAAAAAGAGATTGAATTTCA
Tonda di Gif.	(EU)	TTAGCAACCCGAATCAAAATGTATAGATCTAATCGTAATCAAAAAAGAGATTGAATTTCA
N.A.	(GE)	TTAGCAACCCGAATCAAAATGTATAGATCTAATCGTAATCAAAAAAGAGATTGAATTTCA
Ata Baba	(AZ)	TTAGCAACCCGAATCAAAATGTATAGATCTAATCGTAATCAAAAAAGAGATTGAATTTCA
Palaz	(TR)	TTAGCAACCCGAATCAAAATGTATAGATCTAATCGTAATCAAAAAAGAGATTGAATTTCA
Negret	(EU)	-----NNNNNNNNNNNNNNNATCNNNNCGTAATCAAAAAAGAGATNGAATTTCA
Pauetet	(EU)	-----NNNNNNNNNGNNNNNATCNNNNCGTAATCAAAAAAGAGATTGAATTTCA
T.G.Romana	(EU)	-----NNNNNNNNNGNNNNNATCTAATCGTAATCAAAAAAGAGATTGAATTTCA
T.G.Trilobata	(EU)	-----NNNNNNNNNNNNNNNNNNATCTANNCNNAATCAAAAAAGAGATTGAATTTCA
Hallsche R.	(EU)	-----NNNNNNNNNGNATNGATCTANTCGTAATCAAAAAAGAGATTGAATTTCA

Segorbe	(EU)	CAACGCAATCAAAATATTAATAAAAAAAAAATTTCTAATCGATTCTGAACATAACATA
Tonda di Gif.	(EU)	CAACGCAATCAAAATATTAATAAAAAAAAAATTTCTAATCGATTCTGAACATAACATA
N.A.	(GE)	CAACGCAATCAAAATATTAATAAAAAAAAAATTTCTAATCGATTCTGAACATAACATA
Ata Baba	(AZ)	CAACGCAATCAAAATATTAATAAAAAAAAAATTTCTAATCGATTCTGAACATAACATA
Palaz	(TR)	CAACGCAATCAAAATATTAATAAAAAAAAAATTTCTAATCGATTCTGAACATAACATA
Negret	(EU)	CAACGCAATCAAAATATTAATAAAAAAAAAATTTCTAATCGATTCTGAACATAACATA
Pauetet	(EU)	CAACGCAATCAAAATATTAATAAAAAAAAAATTTCTAATCGATTCTGAACATAACATA
T.G.Romana	(EU)	CAACGCAATCAAAATATTAATAAAAAAAAAATTTCTAATCGATTCTGAACATAACATA
T.G.Trilobata	(EU)	CAACGCAATCAAAATATTAATAAAAAAAAAATTTCTAATCGATTCTGAACATAACATA
Hallsche R.	(EU)	CAACGCAATCAAAATATTAATAAAAAAAAAATTTCTAATCGATTCTGAACATAACATA

Segorbe	(EU)	AAAATGAACATATCAGATGCAAATACAATGACGTCTAAGATAAGAATATAGATTAGGATA
Tonda di Gif.	(EU)	AAAATGAACATATCAGATGCAAATACAATGACGTCTAAGATAAGAATATAGATTAGGATA
N.A.	(GE)	AAAATGAACATATCAGATGCAAATACAATGACGTCTAAGATAAGAATATAGATTAGGATA
Ata Baba	(AZ)	AAAATGAACATATCAGATGCAAATACAATGACGTCTAAGATAAGAATATAGATTAGGATA
Palaz	(TR)	AAAATGAACATATCAGATGCAAATACAATGACGTCTAAGATAAGAATATAGATTAGGATA
Negret	(EU)	AAAATGAACATATCAGATGCAAATACAATGACGTCTAAGATAAGAATATAGATTAGGATA
Pauetet	(EU)	AAAATGAACATATCAGATGCAAATACAATGACGTCTAAGATAAGAATATAGATTAGGATA
T.G.Romana	(EU)	AAAATGAACATATCAGATGCAAATACAATGACGTCTAAGATAAGAATATAGATTAGGATA
T.G.Trilobata	(EU)	AAAATGAACATATCAGATGCAAATACAATGACGTCTAAGATAAGAATATAGATTAGGATA
Hallsche R.	(EU)	AAAATGAACATATCAGATGCAAATACAATGACGTCTAAGATAAGAATATAGATTAGGATA

Segorbe	(EU)	AAAATATAGATAAAAATAAAAATTTATAGACTACCACATATTTTGTTCGTATGTTATCCAT
Tonda di Gif.	(EU)	AAAATATAGATAAAAATAAAAATTTATAGACTACCACATATTTTGTTCGTATGTTATCCAT
N.A.	(GE)	AAAATATAGATAAAAATAAAAATTTATAGACTACCACATATTTTGTTCGTATGTTATCCAT
Ata Baba	(AZ)	AAAATATAGATAAAAATAAAAATTTATAGACTACCACATATTTTGTTCGTATGTTATCCAT
Palaz	(TR)	AAAATATAGATAAAAATAAAAATTTATAGACTACCACATATTTTGTTCGTATGTTATCCAT
Negret	(EU)	AAAATATAGATAAAAATAAAAATTTATAGACTACCACATATTTTGTTCGTATGTTATCCAT
Pauetet	(EU)	AAAATATAGATAAAAATAAAAATTTATAGACTACCACATATTTTGTTCGTATGTTATCCAT
T.G.Romana	(EU)	AAAATATAGATAAAAATAAAAATTTATAGACTACCACATATTTTGTTCGTATGTTATCCAT
T.G.Trilobata	(EU)	AAAATATAGATAAAAATAAAAATTTATAGACTACCACATATTTTGTTCGTATGTTATCCAT
Hallsche R.	(EU)	AAAATATAGATAAAAATAAAAATTTATAGACTACCACATATTTTGTTCGTATGTTATCCAT

Segorbe	(EU)	-----
Tonda di Gif.	(EU)	-----
N.A.	(GE)	TTATTTTTTTTTTAATAAATTGATTTTTTTTTTAATAAAAAAAAAATAAAGACTTCTTGAT
Ata Baba	(AZ)	TTATTTTTTTTTTAATAAATTGATTTTTTTTTTAATAAAAAAAAAATAAAGACTTCTTGAT
Palaz	(TR)	TTATTTTTTTTTTAATAAATTGATTTTTTTTTTAATAAAAAAAAAATAAAGACTTCTTGAT
Negret	(EU)	-----
Pauetet	(EU)	-----
T.G.Romana	(EU)	-----
T.G.Trilobata	(EU)	-----
Hallsche R.	(EU)	-----



Segorbe	(EU)	-----AG
Tonda di Gif.	(EU)	-----AG
N.A.	(GE)	CCCAGCAAATCCAATGAACCCACCATTTGAATAAAGTAAAATAAAAAACCAAGTCTATAG
Ata Baba	(AZ)	CCCAGCAAATCCAATGAACCCACCATTTGAATAAAGTAAAATAAAAAACCAAGTCTATAG
Palaz	(TR)	CCCAGCAAATCCAATGAACCCACCATTTGAATAAAGTAAAATAAAAAACCAAGTCTATAG
Negret	(EU)	-----AG
Pauetet	(EU)	-----AG
T.G.Romana	(EU)	-----AG
T.G.Trilobata	(EU)	-----AG
Hallsche R.	(EU)	-----AG
		**
Segorbe	(EU)	AACAGAGATAAAATATATACATTTATTCACGATCGGATTATATTTGTTTGTATATACTGTTG
Tonda di Gif.	(EU)	AACAGAGATAAAATATATACATTTATTCACGATCGGATTATATTTGTTTGTATATACTGTTG
N.A.	(GE)	AACAGAGATAAAATATATACATTTATTCACGATCGGATTATATTTGTTTGTATATACTGTTG
Ata Baba	(AZ)	AACAGAGATAAAATATATACATTTATTCACGATCGGATTATATTTGTTTGTATATACTGTTG
Palaz	(TR)	AACAGAGATAAAATATATACATTTATTCACGATCGGATTATATTTGTTTGTATATACTGTTG
Negret	(EU)	AACAGAGATAAAATATATACATTTATTCACGATCGGATTATATTTGTTTGTATATACTGTTG
Pauetet	(EU)	AACAGAGATAAAATATATACATTTATTCACGATCGGATTATATTTGTTTGTATATACTGTTG
T.G.Romana	(EU)	AACAGAGATAAAATATATACATTTATTCACGATCGGATTATATTTGTTTGTATATACTGTTG
T.G.Trilobata	(EU)	AACAGAGATAAAATATATACATTTATTCACGATCGGATTATATTTGTTTGTATATACTGTTG
Hallsche R.	(EU)	AACAGAGATAAAATATATACATTTATTCACGATCGGATTATATTTGTTTGTATATACTGTTG

Segorbe	(EU)	TCAATATAAAAGTAAATGTTGAGAAAAGAACAATAATGTAGAAAACCATAAATCAAATAA
Tonda di Gif.	(EU)	TCAATATAAAAGTAAATGTTGAGAAAAGAACAATAATGTAGAAAACCATAAATCAAATAA
N.A.	(GE)	TCAATATAAAAGTAAATGTTGAGAAAAGAACAATAATGTAGAAAACCATAAATCAAATAA
Ata Baba	(AZ)	TCAATATAAAAGTAAATGTTGAGAAAAGAACAATAATGTAGAAAACCATAAATCAAATAA
Palaz	(TR)	TCAATATAAAAGTAAATGTTGAGAAAAGAACAATAATGTAGAAAACCATAAATCAAATAA
Negret	(EU)	TCAATATAAAAGTAAATGTTGAGAAAAGAACAATAATGTAGAAAACCATAAATCAAATAA
Pauetet	(EU)	TCAATATAAAAGTAAATGTTGAGAAAAGAACAATAATGTAGAAAACCATAAATCAAATAA
T.G.Romana	(EU)	TCAATATAAAAGTAAATGTTGAGAAAAGAACAATAATGTAGAAAACCATAAATCAAATAA
T.G.Trilobata	(EU)	TCAATATAAAAGTAAATGTTGAGAAAAGAACAATAATGTAGAAAACCATAAATCAAATAA
Hallsche R.	(EU)	TCAATATAAAAGTAAATGTTGAGAAAAGAACAATAATGTAGAAAACCATAAATCAAATAA

Segorbe	(EU)	AAAATTATTCAATATTTTCTTATTAATTTATTAATTCATATAATATTTAATTAATTAAT
Tonda di Gif.	(EU)	AAAATTATTCAATATTTTCTTATTAATTTATTAATTCATATAATATTTAATTAATTAAT
N.A.	(GE)	AAAATTATTCAATATTTTCTTATTAATTTATTAATTCATATAATATTTAATTAATTAAT
Ata Baba	(AZ)	AAAATTATTCAATATTTTCTTATTAATTTATTAATTCATATAATATTTAATTAATTAAT
Palaz	(TR)	AAAATTATTCAATATTTTCTTATTAATTTATTAATTCATATAATATTTAATTAATTAAT
Negret	(EU)	AAAATTATTCAATATTTTCTTATTAATTTATTAATTCATATAATATTTAATTAATTAAT
Pauetet	(EU)	AAAATTATTCAATATTTTCTTATTAATTTATTAATTCATATAATATTTAATTAATTAAT
T.G.Romana	(EU)	AAAATTATTCAATATTTTCTTATTAATTTATTAATTCATATAATATTTAATTAATTAAT
T.G.Trilobata	(EU)	AAAATTATTCAATATTTTCTTATTAATTTATTAATTCATATAATATTTAATTAATTAAT
Hallsche R.	(EU)	AAAATTATTCAATATTTTCTTATTAATTTATTAATTCATATAATATTTAATTAATTAAT

Segorbe	(EU)	AACTATAAATAAATAGAAATGAAAAATAAAAAACTAATGACTTGTATTAAATTAGCACT
Tonda di Gif.	(EU)	AACTATAAATAAATAGAAATGAAAAATAAAAAACTAATGACTTGTATTAAATTAGCACT
N.A.	(GE)	AACTATAAATAAATAGAAATGAAAAATAAAAAACTAATGACTTGTATTAAATTAGCACT
Ata Baba	(AZ)	AACTATAAATAAATAGAAATGAAAAATAAAAAACTAATGACTTGTATTAAATTAGCACT
Palaz	(TR)	AACTATAAATAAATAGAAATGAAAAATAAAAAACTAATGACTTGTATTAAATTAGCACT
Negret	(EU)	AACTATAAATAAATAGAAATGAAAAATAAAAAACTAATGACTTGTATTAAATTAGCACT
Pauetet	(EU)	AACTATAAATAAATAGAAATGAAAAATAAAAAACTAATGACTTGTATTAAATTAGCACT
T.G.Romana	(EU)	AACTATAAATAAATAGAAATGAAAAATAAAAAACTAATGACTTGTATTAAATTAGCACT
T.G.Trilobata	(EU)	AACTATAAATAAATAGAAATGAAAAATAAAAAACTAATGACTTGTATTAAATTAGCACT
Hallsche R.	(EU)	AACTATAAATAAATAGAAATGAAAAATAAAAAACTAATGACTTGTATTAAATTAGCACT

Segorbe	(EU)	ACATATTTAATAAAGATAAATACAAAAAGGTATAGGCCGTAAAAAAAATTCGGAGAGAAGT
Tonda di Gif.	(EU)	ACATATTTAATAAAGATAAATACAAAAAGGTATAGGCCGTAAAAAAAATTCGGAGAGAAGT
N.A.	(GE)	ACATATTTAATAAAGATAAATACAAAAAGGTATAGGCCGTAAAAAAAATTCGGAGAGAAGT
Ata Baba	(AZ)	ACATATTTAATAAAGATAAATACAAAAAGGTATAGGCCGTAAAAAAAATTCGGAGAGAAGT
Palaz	(TR)	ACATATTTAATAAAGATAAATACAAAAAGGTATAGGCCGTAAAAAAAATTCGGAGAGAAGT
Negret	(EU)	ACATATTTAATAAAGATAAATACAAAAAGGTATAGGCCGTAAAAAAAATTCGGAGAGAAGT
Pauetet	(EU)	ACATATTTAATAAAGATAAATACAAAAAGGTATAGGCCGTAAAAAAAATTCGGAGAGAAGT
T.G.Romana	(EU)	ACATATTTAATAAAGATAAATACAAAAAGGTATAGGCCGTAAAAAAAATTCGGAGAGAAGT
T.G.Trilobata	(EU)	ACATATTTAATAAAGATAAATACAAAAAGGTATAGGCCGTAAAAAAAATTCGGAGAGAAGT
Hallsche R.	(EU)	ACATATTTAATAAAGATAAATACAAAAAGGTATAGGCCGTAAAAAAAATTCGGAGAGAAGT

SNP 3



Alignment Bereich C:

Ata Baba	(AZ)	NNNNTTCTTATTTATAGATGGTTTATAGGAATAAATTGATATTTCCCTTCTCTACAGAGA
N.A.	(GE)	TAGATTCCTTATTTATAGATGGTTTATAGGAATAAATTGATATTTCCCTTCTCTACAGAGA
Palaz	(TR)	AAGATTCCTTATTTATAGATGGTTTATAGGAATAAATTGATATTTCCCTTCTCTACAGAGA
Segorbe	(EU)	NNNATTCNATTTATAGATGGTTTATAGGAATAAATTGATATTTCCCTTCTCTACAGAGA
T.G.Romana	(EU)	NNNNTTNNTATTTNATAGATGGTTTATAGGAATAAATTGATATTTCCCTTCTCTACAGAGA
Hallsche R.	(EU)	NNNNTTCTTATTTATAGATGGTTTATAGGAATAAATTGATATTTCCCTTCTCTACAGAGA
Pauetet	(EU)	NNNANTNTTATTTATAGATGGTTTATAGGAATAAATTGATATTTCCCTTCTCTACAGAGA
T.G.Trilobata	(EU)	NNNNNNNNNATTTNNTAGATGGTTNATAGGNNTAANTTGATATTTCCCTTCTCTACAGAGA
Negret	(EU)	NNAGNNCTTATTTATAGATGGTTTATAGGAATAAATTGATATTTCCCTTCTCTACAGAGA
Tonda di Gif.	(EU)	NANNNNTTATTTATAGATGGTTTATAGGAATAAATTGATATTTCCCTTCTCTACAGAGA

*** ***** ***** ** Indel 2

Ata Baba	(AZ)	AAGGCTTATTACAAGTCACAACATACGAGCCATTTAAAAATATCTTTTTT	TTTATTTTAA
N.A.	(GE)	AAGGCTTATTACAAGTCACAACATACGAGCCATTTAAAAATATCTTTTTT	TTTATTTTAA
Palaz	(TR)	AAGGCTTATTACAAGTCACAACATACGAGCCATTTAAAAATATCTTTTTT	TTTATTTTAA
Segorbe	(EU)	AAGGCTTATTACAAGTCACAACATACGAGCCATTTAAAAATATCTTTTTT	TTTATTTTAA
T.G.Romana	(EU)	AAGGCTTATTACAAGTCACAACATACGAGCCATTTAAAAATATCTTTTTT	TTTATTTTAA
Hallsche R.	(EU)	AAGGCTTATTACAAGTCACAACATACGAGCCATTTAAAAATATCTTTTTT	TTTATTTTAA
Pauetet	(EU)	AAGGCTTATTACAAGTCACAACATACGAGCCATTTAAAAATATCTTTTTT	TTTATTTTAA
T.G.Trilobata	(EU)	AAGGCTTATTACAAGTCACAACATACGAGCCATTTAAAAATATCTTTTTT	TTTATTTTAA
Negret	(EU)	AAGGCTTATTACAAGTCACAACATACGAGCCATTTAAAAATATCTTTTTT	TTTATTTTAA
Tonda di Gif.	(EU)	AAGGCTTATTACAAGTCACAACATACGAGCCATTTAAAAATATCTTTTTT	TTTATTTTAA

Ata Baba	(AZ)	TTCCCGAACACAAAGACATGATCAATTTAGATATAAATATTTTTAATTTATATCTAAATAA
N.A.	(GE)	TTCCCGAACACAAAGACATGATCAATTTAGATATAAATATTTTTAATTTATATCTAAATAA
Palaz	(TR)	TTCCCGAACACAAAGACATGATCAATTTAGATATAAATATTTTTAATTTATATCTAAATAA
Segorbe	(EU)	TTCCCGAACACAAAGACATGATCAATTTAGATATAAATATTTTTAATTTATATCTAAATAA
T.G.Romana	(EU)	TTCCCGAACACAAAGACATGATCAATTTAGATATAAATATTTTTAATTTATATCTAAATAA
Hallsche R.	(EU)	TTCCCGAACACAAAGACATGATCAATTTAGATATAAATATTTTTAATTTATATCTAAATAA
Pauetet	(EU)	TTCCCGAACACAAAGACATGATCAATTTAGATATAAATATTTTTAATTTATATCTAAATAA
T.G.Trilobata	(EU)	TTCCCGAACACAAAGACATGATCAATTTAGATATAAATATTTTTAATTTATATCTAAATAA
Negret	(EU)	TTCCCGAACACAAAGACATGATCAATTTAGATATAAATATTTTTAATTTATATCTAAATAA
Tonda di Gif.	(EU)	TTCCCGAACACAAAGACATGATCAATTTAGATATAAATATTTTTAATTTATATCTAAATAA

Ata Baba	(AZ)	GATTTATATAATTTATCAGATCATCGATTCATGTACCAAATATTTCCATATCGAT	ACATG
N.A.	(GE)	GATTTATATAATTTATCAGATCATCGATTCATGTACCAAATATTTCCATATCGAT	ACATG
Palaz	(TR)	GATTTATATAATTTATCAGATCATCGATTCATGTACCAAATATTTCCATATCGAT	ACATG
Segorbe	(EU)	GATTTATATAATTTATCAGATCATCGATTCATGTACCAAATATTTCCATATCGAT	-----
T.G.Romana	(EU)	GATTTATATAATTTATCAGATCATCGATTCATGTACCAAATATTTCCATATCGAT	-----
Hallsche R.	(EU)	GATTTATATAATTTATCAGATCATCGATTCATGTACCAAATATTTCCATATCGAT	-----
Pauetet	(EU)	GATTTATATAATTTATCAGATCATCGATTCATGTACCAAATATTTCCATATCGAT	-----
T.G.Trilobata	(EU)	GATTTATATAATTTATCAGATCATCGATTCATGTACCAAATATTTCCATATCGAT	-----
Negret	(EU)	GATTTATATAATTTATCAGATCATCGATTCATGTACCAAATATTTCCATATCGAT	-----
Tonda di Gif.	(EU)	GATTTATATAATTTATCAGATCATCGATTCATGTACCAAATATTTCCATATCGAT	-----

***** Indel 3

Ata Baba	(AZ)	CGATATTTTGTTCGACAACGGGATGGAGAATGGATGCGAGAAAGATACTTTCATTTAC
N.A.	(GE)	CGATATTTTGTTCGACAACGGGATGGAGAATGGATGCGAGAAAGATACTTTCATTTAC
Palaz	(TR)	CGATATTTTGTTCGACAACGGGATGGAGAATGGATGCGAGAAAGATACTTTCATTTAC
Segorbe	(EU)	-----
T.G.Romana	(EU)	-----
Hallsche R.	(EU)	-----
Pauetet	(EU)	-----
T.G.Trilobata	(EU)	-----
Negret	(EU)	-----
Tonda di Gif.	(EU)	-----

Ata Baba	(AZ)	AGTTTTTTTATTAATATTAATAA	CAAAAAATATTAATAAAAAGACAAT	TTTTATTATT
N.A.	(GE)	AGTTTTTTTATTAATATTAATAA	CAAAAAATATTAATAAAAAGACAAT	TTTTATTATT
Palaz	(TR)	AGTTTTTTTATTAATATTAATAA	CAAAAAATATTAATAAAAAGACAAT	TTTTATTATT
Segorbe	(EU)	-----	-----	TTTTATTATT
T.G.Romana	(EU)	-----	-----	TTTTATTATT
Hallsche R.	(EU)	-----	-----	TTTTATTATT
Pauetet	(EU)	-----	-----	TTTTATTATT
T.G.Trilobata	(EU)	-----	-----	TTTTATTATT
Negret	(EU)	-----	-----	TTTTATTATT
Tonda di Gif.	(EU)	-----	-----	TTTTATTATT

SNP 4

Indel 4


Ata Baba	(AZ)	AAATTAATACAATATTTAAAAA-AAAAATGAAAAATGAAAGTATAACGTAATAAAATATAT
N.A.	(GE)	AAATTAATACAATATTTAAAAA-AAAAATGAAAAATGAAAGTATAACGTAATAAAATATAT
Palaz	(TR)	AAATTAATACAATATTTAAAAA-AAAAATGAAAAATGAAAGTATAACGTAATAAAATATAT
Segorbe	(EU)	AAATTAATACAATATTTAAAAA-AAAAATGAAAAATGAAAGTATAACGTAATAAAATATAT
T.G.Romana	(EU)	AAATTAATACAATATTTAAAAA-AAAAATGAAAAATGAAAGTATAACGTAATAAAATATAT
Hallsche R.	(EU)	AAATTAATACAATATTTAAAAA-AAAAATGAAAAATGAAAGTATAACGTAATAAAATATAT
Pauetet	(EU)	AAATTAATACAATATTTAAAAA-AAAAATGAAAAATGAAAGTATAACGTAATAAAATATAT
T.G.Trilobata	(EU)	AAATTAATACAATATTTAAAAA-AAAAATGAAAAATGAAAGTATAACGTAATAAAATATAT
Negret	(EU)	AAATTAATACAATATTTAAAAA-AAAAATGAAAAATGAAAGTATAACGTAATAAAATATAT
Tonda di Gif.	(EU)	AAATTAATACAATATTTAAAAA-AAAAATGAAAAATGAAAGTATAACGTAATAAAATATAT

Ata Baba	(AZ)	AATACTATAGTAGTTTAAATACACATATAAAATTTGAAGGAATACATAAAAAATATTTTTTTT
N.A.	(GE)	AATACTATAGTAGTTTAAATACACATATAAAATTTGAAGGAATACATAAAAAATATTTTTTTT
Palaz	(TR)	AATACTATAGTAGTTTAAATACACATATAAAATTTGAAGGAATACATAAAAAATATTTTTTTT
Segorbe	(EU)	AATACTATAGTAGTTTAAATACACATATAAAATTTGAAGGAATACATAAAAAATATTTTTTTT
T.G.Romana	(EU)	AATACTATAGTAGTTTAAATACACATATAAAATTTGAAGGAATACATAAAAAATATTTTTTTT
Hallsche R.	(EU)	AATACTATAGTAGTTTAAATACACATATAAAATTTGAAGGAATACATAAAAAATATTTTTTTT
Pauetet	(EU)	AATACTATAGTAGTTTAAATACACATATAAAATTTGAAGGAATACATAAAAAATATTTTTTTT
T.G.Trilobata	(EU)	AATACTATAGTAGTTTAAATACACATATAAAATTTGAAGGAATACATAAAAAATATTTTTTTT
Negret	(EU)	AATACTATAGTAGTTTAAATACACATATAAAATTTGAAGGAATACATAAAAAATATTTTTTTT
Tonda di Gif.	(EU)	AATACTATAGTAGTTTAAATACACATATAAAATTTGAAGGAATACATAAAAAATATTTTTTTT

Ata Baba	(AZ)	TATCAATCTCACAACGAGATTCAAGAATAAGATTAGTTGATGGAACGAGAGGAATATGT
N.A.	(GE)	TATCAATCTCACAACGAGATTCAAGAATAAGATTAGTTGATGGAACGAGAGGAATATGT
Palaz	(TR)	TATCAATCTCACAACGAGATTCAAGAATAAGATTAGTTGATGGAACGAGAGGAATATGT
Segorbe	(EU)	TATCAATCTCACAACGAGATTCAAGAATAAGATTAGTTGATGGAACGAGAGGAATATGT
T.G.Romana	(EU)	TATCAATCTCACAACGAGATTCAAGAATAAGATTAGTTGATGGAACGAGAGGAATATGT
Hallsche R.	(EU)	TATCAATCTCACAACGAGATTCAAGAATAAGATTAGTTGATGGAACGAGAGGAATATGT
Pauetet	(EU)	TATCAATCTCACAACGAGATTCAAGAATAAGATTAGTTGATGGAACGAGAGGAATATGT
T.G.Trilobata	(EU)	TATCAATCTCACAACGAGATTCAAGAATAAGATTAGTTGATGGAACGAGAGGAATATGT
Negret	(EU)	TATCAATCTCACAACGAGATTCAAGAATAAGATTAGTTGATGGAACGAGAGGAATATGT
Tonda di Gif.	(EU)	TATCAATCTCACAACGAGATTCAAGAATAAGATTAGTTGATGGAACGAGAGGAATATGT

Alignment Bereich D:

N.A.	(GE)	-----AAAAATTGGTACTATTTGTATATATTTTATAGATACATACTTATCTAGATATA
Negret	(EU)	-NNNNNNNNNNNGNNNNNNNNNTGNNNANATTTNATAGATACATACTTATCTAGATATA
Ata Baba	(AZ)	NNNNNNNNNNNNNNNNNNNTTGTATATATTTTATAGATACATACTTATCTAGATATA
Palaz	(TR)	----NNNNNNNNNTTNGNNNNNTTGNATATATTTTATAGATACATACTTATCTAGATATA
Hallsche R.	(EU)	-----NNNNNNNTTNNNNNNNATTTGTATATATTTTATAGATACATACTTATCTAGATATA
T.G.Trilobata	(EU)	-----NNNNNNNNNNNNNNNTTGNATNTATTTTATAGATACATACTTATCTAGATATA
Tonda di Gif.	(EU)	---NNNNNNNNNNNGNNNNNNNTTGTATATATTTTATAGATACATACTTATCTAGATATA
Segorbe	(EU)	-----NNNNNNNNNNNNNATTTGTATATATTTTATAGATACATACTTATCTAGATATA
T.G.Romana	(EU)	----NNNNNNNNNNNNNNNNNTTGTATATATTTTATAGATACATACTTATCTAGATATA
Pauetet	(EU)	-----NNNNNNNNNNNNNNNTTGTATATATTTTATAGATACATACTTATCTAGATATA
		**** *****

N.A.	(GE)	CAGGATCTTAAATAAAAAAGCTATATCTAAAACAAAACAACCTAAACGTTTTTATTGTT
Negret	(EU)	CAGGATCTTAAATAAAAAAGCTATATCTAAAACAAAACAACCTAAACGTTTTTATTGTT
Ata Baba	(AZ)	CAGGATCTTAAATAAAAAAGCTATATCTAAAACAAAACAACCTAAACGTTTTTATTGTT
Palaz	(TR)	CAGGATCTTAAATAAAAAAGCTATATCTAAAACAAAACAACCTAAACGTTTTTATTGTT
Hallsche R.	(EU)	CAGGATCTTAAATAAAAAAGCTATATCTAAAACAAAACAACCTAAACGTTTTTATTGTT
T.G.Trilobata	(EU)	CAGGATCTTAAATAAAAAAGCTATATCTAAAACAAAACAACCTAAACGTTTTTATTGTT
Tonda di Gif.	(EU)	CAGGATCTTAAATAAAAAAGCTATATCTAAAACAAAACAACCTAAACGTTTTTATTGTT
Segorbe	(EU)	CAGGATCTTAAATAAAAAAGCTATATCTAAAACAAAACAACCTAAACGTTTTTATTGTT
T.G.Romana	(EU)	CAGGATCTTAAATAAAAAAGCTATATCTAAAACAAAACAACCTAAACGTTTTTATTGTT
Pauetet	(EU)	CAGGATCTTAAATAAAAAAGCTATATCTAAAACAAAACAACCTAAACGTTTTTATTGTT

N.A.	(GE)	GTGTTGGGTCCACAATGAATTCATGGGCCTTAGGGTTGGTGTTCCTTTTCTATCCCG
Negret	(EU)	GTGTTGGGTCCACAATGAATTCATGGGCCTTAGGGTTGGTGTTCCTTTTCTATCCCG
Ata Baba	(AZ)	GTGTTGGGTCCACAATGAATTCATGGGCCTTAGGGTTGGTGTTCCTTTTCTATCCCG
Palaz	(TR)	GTGTTGGGTCCACAATGAATTCATGGGCCTTAGGGTTGGTGTTCCTTTTCTATCCCG
Hallsche R.	(EU)	GTGTTGGGTCCACAATGAATTCATGGGCCTTAGGGTTGGTGTTCCTTTTCTATCCCG
T.G.Trilobata	(EU)	GTGTTGGGTCCACAATGAATTCATGGGCCTTAGGGTTGGTGTTCCTTTTCTATCCCG
Tonda di Gif.	(EU)	GTGTTGGGTCCACAATGAATTCATGGGCCTTAGGGTTGGTGTTCCTTTTCTATCCCG
Segorbe	(EU)	GTGTTGGGTCCACAATGAATTCATGGGCCTTAGGGTTGGTGTTCCTTTTCTATCCCG
T.G.Romana	(EU)	GTGTTGGGTCCACAATGAATTCATGGGCCTTAGGGTTGGTGTTCCTTTTCTATCCCG
Pauetet	(EU)	GTGTTGGGTCCACAATGAATTCATGGGCCTTAGGGTTGGTGTTCCTTTTCTATCCCG

N.A. (GE) TAGTTTCCGATCATGGATCGAGCCACATATCACAACCTTCTTCTACCCATCCCGTATATTG
 Negret (EU) TAGTTTCCGATCATGGATCGAGCCACATATCACAACCTTCTTCTACCCATCCCGTATATTG
 Ata Baba (AZ) TAGTTTCCGATCATGGATCGAGCCACATATCACAACCTTCTTCTACCCATCCCGTATATTG
 Palaz (TR) TAGTTTCCGATCATGGATCGAGCCACATATCACAACCTTCTTCTACCCATCCCGTATATTG
 Hallsche R. (EU) TAGTTTCCGATCATGGATCGAGCCACATATCACAACCTTCTTCTACCCATCCCGTATATTG
 T.G.Trilobata (EU) TAGTTTCCGATCATGGATCGAGCCACATATCACAACCTTCTTCTACCCATCCCGTATATTG
 Tonda di Gif. (EU) TAGTTTCCGATCATGGATCGAGCCACATATCACAACCTTCTTCTACCCATCCCGTATATTG
 Segorbe (EU) TAGTTTCCGATCATGGATCGAGCCACATATCACAACCTTCTTCTACCCATCCCGTATATTG
 T.G.Romana (EU) TAGTTTCCGATCATGGATCGAGCCACATATCACAACCTTCTTCTACCCATCCCGTATATTG
 Pauetet (EU) TAGTTTCCGATCATGGATCGAGCCACATATCACAACCTTCTTCTACCCATCCCGTATATTG

N.A. (GE) TCCTTTTCGTTCTGTGTGGGCTAGAAACTTATTAGTAGACGAGATTTTACGAAAAAAGT
 Negret (EU) TCCTTTTCGTTCTGTGTGGGCTAGAAACTTATTAGTAGACGAGATTTTACGAAAAAAGT
 Ata Baba (AZ) TCCTTTTCGTTCTGTGTGGGCTAGAAACTTATTAGTAGACGAGATTTTACGAAAAAAGT
 Palaz (TR) TCCTTTTCGTTCTGTGTGGGCTAGAAACTTATTAGTAGACGAGATTTTACGAAAAAAGT
 Hallsche R. (EU) TCCTTTTCGTTCTGTGTGGGCTAGAAACTTATTAGTAGACGAGATTTTACGAAAAAAGT
 T.G.Trilobata (EU) TCCTTTTCGTTCTGTGTGGGCTAGAAACTTATTAGTAGACGAGATTTTACGAAAAAAGT
 Tonda di Gif. (EU) TCCTTTTCGTTCTGTGTGGGCTAGAAACTTATTAGTAGACGAGATTTTACGAAAAAAGT
 Segorbe (EU) TCCTTTTCGTTCTGTGTGGGCTAGAAACTTATTAGTAGACGAGATTTTACGAAAAAAGT
 T.G.Romana (EU) TCCTTTTCGTTCTGTGTGGGCTAGAAACTTATTAGTAGACGAGATTTTACGAAAAAAGT
 Pauetet (EU) TCCTTTTCGTTCTGTGTGGGCTAGAAACTTATTAGTAGACGAGATTTTACGAAAAAAGT

N.A. (GE) TCTTTATATTCCTTTATATTCATAGGATATAACAATTATTTTCATTTTTCGATGTTAATTT
 Negret (EU) TCTTTATATTCCTTTATATTCATAGGATATAACAATTATTTTCATTTTTCGATGTTAATTT
 Ata Baba (AZ) TCTTTATATTCCTTTATATTCATAGGATATAACAATTATTTTCATTTTTCGATGTTAATTT
 Palaz (TR) TCTTTATATTCCTTTATATTCATAGGATATAACAATTATTTTCATTTTTCGATGTTAATTT
 Hallsche R. (EU) TCTTTATATTCCTTTATATTCATAGGATATAACAATTATTTTCATTTTTCGATGTTAATTT
 T.G.Trilobata (EU) TCTTTATATTCCTTTATATTCATAGGATATAACAATTATTTTCATTTTTCGATGTTAATTT
 Tonda di Gif. (EU) TCTTTATATTCCTTTATATTCATAGGATATAACAATTATTTTCATTTTTCGATGTTAATTT
 Segorbe (EU) TCTTTATATTCCTTTATATTCATAGGATATAACAATTATTTTCATTTTTCGATGTTAATTT
 T.G.Romana (EU) TCTTTATATTCCTTTATATTCATAGGATATAACAATTATTTTCATTTTTCGATGTTAATTT
 Pauetet (EU) TCTTTATATTCCTTTATATTCATAGGATATAACAATTATTTTCATTTTTCGATGTTAATTT

N.A. (GE) GATACAACAAACACGGCGGGATCACTATTTTCAAATTGAAAAATAATTGAAATCATTAAAA
 Negret (EU) GATACAACAAACACGGCGGGATCACTATTTTAAAATTGAAAAATAATGGAAATCATTAAAA
 Ata Baba (AZ) GATACAACAAACACGGCGGGATCACTATTTTAAAATTGAAAAATAATGGAAATCATTAAAA
 Palaz (TR) GATACAACAAACACGGCGGGATCACTATTTTAAAATTGAAAAATAATGGAAATCATTAAAA
 Hallsche R. (EU) GATACAACAAACACGGCGGGATCACTATTTTAAAATTGAAAAATAATGGAAATCATTAAAA
 T.G.Trilobata (EU) GATACAACAAACACGGCGGGATCACTATTTTAAAATTGAAAAATAATGGAAATCATTAAAA
 Tonda di Gif. (EU) GATACAACAAACACGGCGGGATCACTATTTTAAAATTGAAAAATAATGGAAATCATTAAAA
 Segorbe (EU) GATACAACAAACACGGCGGGATCACTATTTTAAAATTGAAAAATAATGGAAATCATTAAAA
 T.G.Romana (EU) GATACAACAAACACGGCGGGATCACTATTTTAAAATTGAAAAATAATGGAAATCATTAAAA
 Pauetet (EU) GATACAACAAACACGGCGGGATCACTATTTTAAAATTGAAAAATAATGGAAATCATTAAAA

SNP 5 SNP 6 SNP 7

N.A. (GE) AAAGGATTCATCATATTAAGTGATACTCCTGGTTCTCACAAAAGAGAATCATTTTTTT
 Negret (EU) AAAGGATTCATCATATTAAGTGATACTCCTGGTTCTCACAAAAGAGAATCATTTTTT-TT
 Ata Baba (AZ) AAAGGATTCATCATATTAAGTGATACTCCTGGTTCTCACAAAAGAGAATCATTTTTT-TT
 Palaz (TR) AAAGGATTCATCATATTAAGTGATACTCCTGGTTCTCACAAAAGAGAATCATTTTTT-TT
 Hallsche R. (EU) AAAGGATTCATCATATTAAGTGATACTCCTGGTTCTCACAAAAGAGAATCATTTTTT-TT
 T.G.Trilobata (EU) AAAGGATTCATCATATTAAGTGATACTCCTGGTTCTCACAAAAGAGAATCATTTTTT-TT
 Tonda di Gif. (EU) AAAGGATTCATCATATTAAGTGATACTCCTGGTTCTCACAAAAGAGAATCATTTTTT-TT
 Segorbe (EU) AAAGGATTCATCATATTAAGTGATACTCCTGGTTCTCACAAAAGAGAATCATTTTTT-TT
 T.G.Romana (EU) AAAGGATTCATCATATTAAGTGATACTCCTGGTTCTCACAAAAGAGAATCATTTTTT-TT
 Pauetet (EU) AAAGGATTCATCATATTAAGTGATACTCCTGGTTCTCACAAAAGAGAATCATTTTTT-TT

Indel 5

N.A. (GE) TCAATTTCAACACCTACTTTTTTTGTTAATATAATCCTAGTAATTGGATTTATATGCATA
 Negret (EU) TCAATTTCAACACCTACTTTTTTTGTTAATATAATCCTAGTAATTGGATTTATATGCATA
 Ata Baba (AZ) TCAATTTCAACACCTACTTTTTTTGTTAATATAATCCTAGTAATTGGATTTATATGCATA
 Palaz (TR) TCAATTTCAACACCTACTTTTTTTGTTAATATAATCCTAGTAATTGGATTTATATGCATA
 Hallsche R. (EU) TCAATTTCAACACCTACTTTTTTTGTTAATATAATCCTAGTAATTGGATTTATATGCATA
 T.G.Trilobata (EU) TCAATTTCAACACCTACTTTTTTTGTTAATATAATCCTAGTAATTGGATTTATATGCATA
 Tonda di Gif. (EU) TCAATTTCAACACCTACTTTTTTTGTTAATATAATCCTAGTAATTGGATTTATATGCATA
 Segorbe (EU) TCAATTTCAACACCTACTTTTTTTGTTAATATAATCCTAGTAATTGGATTTATATGCATA
 T.G.Romana (EU) TCAATTTCAACACCTACTTTTTTTGTTAATATAATCCTAGTAATTGGATTTATATGCATA
 Pauetet (EU) TCAATTTCAACACCTACTTTTTTTGTTAATATAATCCTAGTAATTGGATTTATATGCATA

```

N.A. (GE) TTCTGATAGTAAATGAAATATTCAAATGATTTTTTATCGAATGACTATTCATCTATTGTA
Negret (EU) TTCTGATAGTAAATGAAATATTCAAATGATTTTTTATCGAATGACTATTCATCTATTGTA
Ata Baba (AZ) TTCTGATAGTAAATGAAATATTCAAATGATTTTTTATCGAATGACTATTCATCTATTGTA
Palaz (TR) TTCTGATAGTAAATGAAATATTCAAATGATTTTTTATCGAATGACTATTCATCTATTGTA
Hallsche R. (EU) TTCTGATAGTAAATGAAATATTCAAATGATTTTTTATCGAATGACTATTCATCTATTGTA
T.G.Trilobata (EU) TTCTGATAGTAAATGAAATATTCAAATGATTTTTTATCGAATGACTATTCATCTATTGTA
Tonda di Gif. (EU) TTCTGATAGTAAATGAAATATTCAAATGATTTTTTATCGAATGACTATTCATCTATTGTA
Segorbe (EU) TTCTGATAGTAAATGAAATATTCAAATGATTTTTTATCGAATGACTATTCATCTATTGTA
T.G.Romana (EU) TTCTGATAGTAAATGAAATATTCAAATGATTTTTTATCGAATGACTATTCATCTATTGTA
Pauetet (EU) TTCTGATAGTAAATGAAATATTCAAATGATTTTTTATCGAATGACTATTCATCTATTGTA
*****

```

9.10.3 Entwicklung und Anwendung von *targeted*-Methoden auf Basis polymorpher Bereiche des cpGenoms

9.10.4 Nachweis der Indels 1 und 3 in den Bereichen B und C des cpGenoms

9.10.4.1 Qualitativer Nachweis

Der folgenden Tabelle sind die Ergebnisse der Analyse der Bereiche B und C auf Vorhandensein der Insertionen zu entnehmen. Es wurden alle analysierten Proben aufgenommen mit Ausnahme der doppelten Proben, welche bereits im Abschnitt 6.3.1.1 unter Tabelle 4 zusammengefasst sind.

Tabelle 39: Ergebnisse der Analyse der Bereiche B und C des cpGenoms bzgl. der Indels. 1: Probe wurde nur hinsichtlich eines Bereichs analysiert, D: nicht auswertbar, da Doppelbande auf dem Gel, D+: kennzeichnet die stärkere Bande, bei Detektion einer Doppelbande.

Sorte	Herkunft	Probennr.	Insertionen vorhanden	Insertionen nicht vorhanden
Ata Baba	Aserbaidshan	123	X	
Ata Baba	Aserbaidshan	52	X	
Ata Baba	Aserbaidshan, Zaqatala	102	X	
Barcelona	Frankreich, Aquitanien	43		X
Butler	Frankreich, Aquitanien	40		X
Çakildak	Türkei, Düzce, Akçakoca	132		X
Çakildak	Türkei, Düzce, Akçakoca	103		X
Çakildak	Türkei, Ordu	104		X
Çakildak	Türkei, Düzce, Akçakoca	105		X
Çakildak	Türkei, Düzce, Akçakoca	212		X
Çakildak	Türkei, Düzce, Akçakoca	213		X
Çakildak	Türkei, Düzce, Akçakoca	214		X
Çakildak	Türkei, Düzce, Akçakoca	216		X
Corabel	Frankreich, Aquitanien	32		X
Ennis	Frankreich, Centre	46		X
Hallsche Riesen	Deutschland, Bayern	F10		X
Karayağlı	Türkei, Düzce, Akçakoca	133		X
Karayağlı	Türkei, Ordu	134	X	
Karayağlı	Türkei, Düzce, Akçakoca	145	X	
Karayağlı	Türkei, Düzce, Akçakoca	146		X
Karayağlı ¹	Türkei, Ordu	147	X	
Karayağlı ¹	Türkei, Ordu	148	X	
Karayağlı	Türkei, Düzce, Akçakoca	153		X

Sorte	Herkunft	Probennr.	Insertionen vorhanden	Insertionen nicht vorhanden
Karayağlı	Türkei, Düzce, Akçakoca	154		X
Karayağlı	Türkei, Düzce, Akçakoca	185		X
Karayağlı	Türkei, Düzce, Akçakoca	186		X
Karayağlı	Türkei, Düzce, Akçakoca	191	X	
Karayağlı	Türkei, Düzce, Akçakoca	192		X
Karayağlı	Türkei, Düzce, Akçakoca	193		X
Karayağlı	Türkei, Düzce, Akçakoca	194		X
Karayağlı	Türkei, Düzce, Akçakoca	195		X
Karayağlı	Türkei, Düzce, Akçakoca	196		X
Mortarella	Italien, Kampanien	106		X
NB (AZ)	Aserbaidshan	121	X	
NB (AZ)	Aserbaidshan	122	X	
NB (GE)	Georgien	124	X	
NB (GE)	Georgien, Tsalenjikha	151	X	
NB (GE) ¹	Georgien, Tsalenjikha	152	X	
NB (GE)	Georgien, Imeretien	56	X	
NB (GE)	Georgien, Mingrelien	55	X	
NB (GE)	Georgien, Gurien	54	X	
NB (GE)	Georgien, Mingrelien	118	X	
NB (Wildtyp)	Türkei, Ordu	135		X
NB (Wildtyp)	Türkei, Ordu	136	X	
NB (Wildtyp)	Türkei, Ordu	137		X
NB (Wildtyp)	Türkei, Ordu	138	X	
Negret	Spanien, Katalonien	2		X
Palaz	Türkei, Ordu	128	X	
Palaz	Türkei, Ordu	139	X	
Palaz	Türkei, Ordu	140	X	
Palaz	Türkei, Ordu	116	D+	D
Palaz	Türkei, Ordu	117	X	
Palaz	Türkei, Ordu	168	X	
Palaz	Türkei, Ordu	169		X
Palaz	Türkei, Ordu	170	X	
Palaz	Türkei, Ordu	172	X	
Pauetet	Frankreich, Poitou-Charentes	31		X
Sariyağlı	Türkei, Düzce, Akçakoca	131	X	
Sariyağlı	Türkei, Düzce, Akçakoca	149	X	
Sariyağlı	Türkei, Düzce, Akçakoca	155	X	
Sariyağlı	Türkei, Düzce, Akçakoca	157	X	
Sariyağlı	Türkei, Düzce, Akçakoca	158	X	
Sariyağlı	Türkei, Düzce, Akçakoca	159	X	
Sariyağlı	Türkei, Düzce, Akçakoca	160	X	
Sariyağlı	Türkei, Düzce, Akçakoca	161	X	
Sariyağlı	Türkei, Düzce, Akçakoca	199	X	
Sariyağlı	Türkei, Düzce, Akçakoca	204	X	
Sariyağlı	Türkei, Düzce, Akçakoca	207	X	
Sariyağlı	Türkei, Düzce, Akçakoca	208	X	
Sariyağlı	Türkei, Düzce, Akçakoca	209	X	

Sorte	Herkunft	Probennr.	Insertionen vorhanden	Insertionen nicht vorhanden
Segorbe	Frankreich, Midi-Pyrénées	37		X
Tombul	Türkei, Ordu	127		X
Tombul	Türkei, Ordu	129		X
Tombul	Türkei, Ordu	141	X	
Tombul	Türkei, Ordu	114	D	D+
Tombul	Türkei, Düzce, Akçakoca	115	D	D+
Tombul	Türkei, Samsun	119	D	D
Tombul	Türkei, Samsun	120	D	D
Tombul	Türkei, Ordu	173		X
Tombul	Türkei, Ordu	174		X
Tombul	Türkei, Ordu	175		X
Tonda di Giffoni	Deutschland, Bayern	F4		X
T. G. Romana	Italien, Latium	7		X
T. G. Trilobata	Deutschland, Bayern	F6		X
T. G. Trilobata	Italien, Piemont	9		X
Yomra	Türkei, Düzce, Akçakoca	142	X	
Yomra	Türkei, Düzce, Akçakoca	143	X	
Yomra	Türkei, Düzce, Akçakoca	177	X	
Yomra	Türkei, Düzce, Akçakoca	178	X	
Yomra	Türkei, Düzce, Akçakoca	179	X	
Yomra	Türkei, Düzce, Akçakoca	180	X	
Yomra	Türkei, Düzce, Akçakoca	183	X	

9.10.4.2 Quantifizierung von Haselnussmischungen

Detektion mittels CGE:

Die folgende Tabelle enthält die prozentualen relativen Areas, die für die Erstellung der Kalibriergeraden für die semi-quantitative Abschätzung der Beimischung georgischer Haselnüsse auf Basis der identifizierten Indels genutzt wurden. Die zugehörige grafische Darstellung der Kalibriergerade ist Abbildung 19 in Abschnitt 6.3.1.3 zu entnehmen.

Tabelle 40: Daten für die Erstellung der Kalibriergerade zur semi-quantitativen Abschätzung von Beimischungen georgischer Haselnüsse auf Basis des identifizierten Indels (Indel 1) im Bereich B des cpGenoms nach Detektion mittels CGE.

Anteil georgischer Haselnüsse [%]	Relative Area [%]				Standardabweichung σ	Variationskoeffizient [%]
	Kalibrierung 1	Kalibrierung 2	Kalibrierung 3	Mittelwert		
80	89,96	87,97	90,61	89,51	1,12	1,25
60	77,17	74,16	71,5	74,28	2,31	3,12
40	59,98	54,39	55,91	56,76	2,36	4,16
20	32,18	29,62	33,94	31,92	1,78	5,55
10	17,6	17,52	19,07	18,06	0,71	3,96
5	8,78	4,77	9,86	7,8	2,19	28,09
2,5	3,78	2,7	4,36	3,61	0,69	19,11

Detektion mittels AGE:

Die folgende Abbildung zeigt die Agarosegele der Kalibriersets 2 und 3 nach erfolgter PCR, durchgeführt wie unter 7.11.1.2 beschrieben. Das Agarosegel des Kalibriersets 1 ist Abbildung 20 unter 6.3.1.3 zu entnehmen.

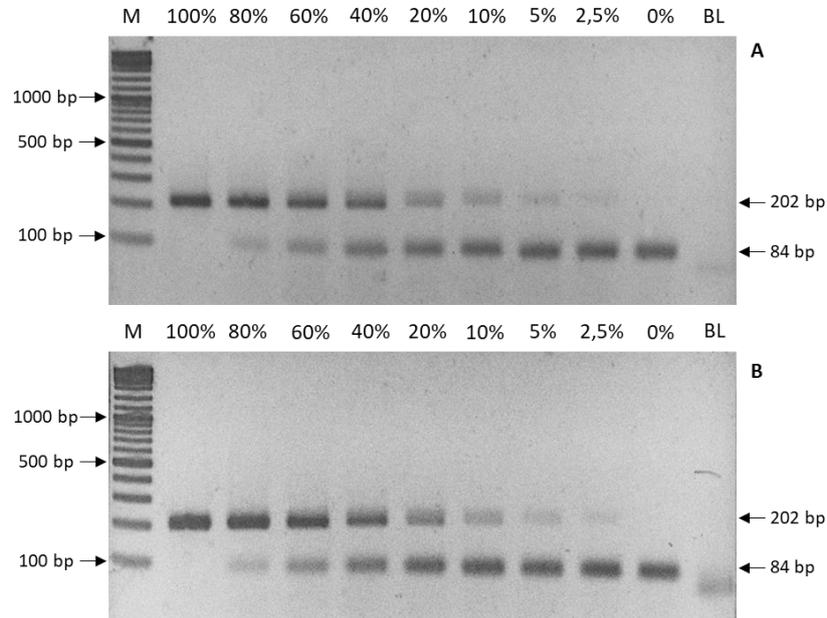


Abbildung 59: Agarosegele der Kalibriersets 2 (A) und 3 (B). Die angegebenen prozentualen Anteile beziehen sich auf den Anteil georgischer Haselnüsse in den Mischungen. Das kurze Fragment (84 bp) ist den italienischen Haselnüssen der Sorte T. G. Trilobata zuzuordnen und das lange Fragment (202 bp) den georgischen Haselnüssen. Bis zu einer Beimischung von 5% ist eine deutliche Bande auf dem Gel zu erkennen. M = DNA-Leiter, BL = Blindwert

In der nachfolgenden Tabelle sind die Ergebnisse der Graustufenanalyse der Agarosegele zusammengefasst. Darauf basierend wurde die unter 6.3.1.3 in Abbildung 21 dargestellte Kalibriergerade erstellt.

Tabelle 41: Daten für die Erstellung der Kalibriergerade zur semi-quantitativen Abschätzung von Beimischungen georgischer Haselnüsse auf Basis des identifizierten Indels (Indel 1) im Bereich B des cpGenoms nach Detektion der PCR-Produkte mittels AGE und anschließender Graustufenanalyse der Gelbilder.

Anteil georgischer Haselnüsse [%]	Relative Area [%]				Standard- abweichung σ	Variations- koeffizient [%]
	Kalibrierung 1	Kalibrierung 2	Kalibrierung 3	Mittelwert		
80	78,28	83,01	77,01	79,43	2,58	3,25
60	66,51	65,83	65,04	65,79	0,6	0,91
40	52,98	53,18	57,15	54,44	1,92	3,53
20	28,65	35,88	46,32	36,95	7,25	19,62
10	18,77	24,29	32,56	25,21	5,67	22,48
5	16,56	15,71	25,84	19,37	4,59	23,69
2,5	8,85	11,45	15,81	12,04	2,87	23,85

9.10.5 PCR-RFLP für den Nachweis georgischer Haselnüsse

Detektion mittels CGE:

Die folgende Tabelle beinhaltet alle Daten, die für die Erstellung der in Abbildung 26 dargestellten Kalibriergerade für den semi-quantitativen Nachweis der Beimischung georgischer Haselnüsse herangezogen wurden. Sie beruhen auf der Auswertung der Elektropherogramme, die durch Messung der drei Kalibriersets nach erfolgter PCR-RFLP, durchgeführt wie unter 7.11.2.2 erläutert, mittels CGE erhalten wurden.

Tabelle 42: Daten für die Erstellung der Kalibriergerade für die semi-quantitative Abschätzung von Beimischungen georgischer Haselnüsse mittels PCR-RFLP auf Basis eines identifizierten SNPs (SNP 6, Bereich D des cpGenoms) nach Detektion mittels CGE. Die relative Area bezieht sich dabei auf das Fragment 2 (413 bp).

Anteil georgischer Haselnüsse [%]	Relative Area [%]				Standard- abweichung σ	Variations- koeffizient [%]
	Kalibrierung 1	Kalibrierung 2	Kalibrierung 3	Mittelwert		
80	32,45	28,86	28,17	29,83	1,88	6,3
60	19,03	18,86	18,24	18,71	0,34	1,82
40	11,7	10,82	10,99	11,17	0,38	3,4
20	4,37	3,67	5,34	4,46	0,69	15,4
10	1,32	1,99	2,23	1,84	0,39	20,9
5	0,68	0,82	1,04	0,85	0,15	17,19
2,5	0,27	0,52	0,57	0,45	0,13	29,11

Detektion mittels AGE:

Die folgende Abbildung zeigt die Agarosegele der drei Kalibriersets nach PCR und enzymatischem Verdau, durchgeführt wie unter 7.11.2.2 beschrieben.

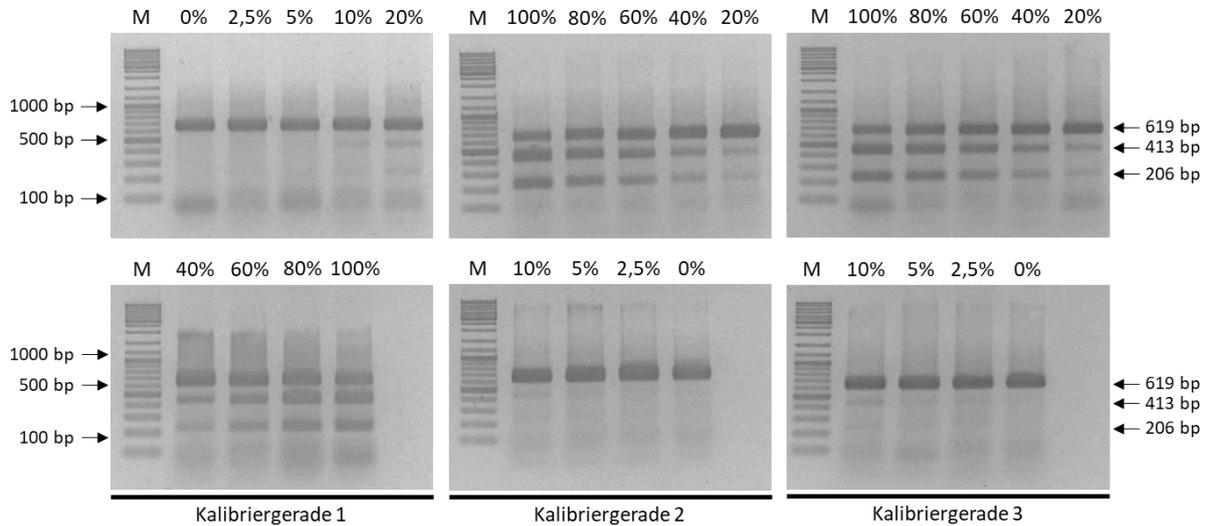


Abbildung 60: Agarosegele der drei Kalibriersets. Die prozentualen Anteile geben den Anteil georgischer Haselnüsse in den Mischungen an. Mit zunehmendem Anteil georgischer Haselnüsse ist eine Zunahme der Intensität der Teil-Fragmente (Fragment 1: 206 bp, Fragment 2: 413 bp), erzeugt durch den enzymatischen Verdau des Ausgangstemplates (619 bp), zu erkennen.

In der nachfolgenden Tabelle sind die Ergebnisse der Graustufenanalyse der Agarosegele zusammengefasst. Mit Hilfe dieser Daten wurde die in Abbildung 27 dargestellte Kalibriergerade für die semi-quantitative Abschätzung der Beimischung georgischer Haselnüsse erstellt.

Tabelle 43: Daten für die Erstellung der Kalibriergerade für die semi-quantitative Abschätzung von Beimischungen georgischer Haselnüsse mittels PCR-RFLP auf Basis eines identifizierten SNPs (SNP 6, Bereich D des cpGenoms) nach Detektion mittels AGE. Die relative Area bezieht sich dabei auf das Fragment 2 (413 bp).

Anteil georgischer Haselnüsse [%]	Relative Area [%]				Standard- abweichung σ	Variations- koeffizient [%]
	Kalibrierung 1	Kalibrierung 2	Kalibrierung 3	Mittelwert		
80	30,52	32,62	32,3	31,81	0,92	2,9
60	26,75	28,64	30,4	28,6	1,49	5,21
40	22,47	22,09	24,54	23,03	1,07	4,67
20	12,82	14,53	15,46	14,27	1,1	7,68
10	5,95	6,56	9,17	7,23	1,4	19,35
5	2,89	3,17	5,28	3,78	1,07	28,21
2,5	1,73	1,8	3,41	2,31	0,78	33,52

9.10.6 HRMA für die Differenzierung von Haselnüssen unterschiedlichen geographischen Ursprungs

9.10.6.1 Berechnung von Schmelzkurven

In Tabelle 44 sind die Parameter, die für die Berechnung von Schmelzkurven im Zuge der Evaluierung des Einflusses dieser ausgewählt wurden, zusammengefasst.

Tabelle 44: Parameter, die für die Berechnung von Schmelzkurven ausgewählt wurden. Als Grundlage dient die unter 6.3.3.2 angegebene 70 bp lange Sequenz. Für alle Berechnungen war der Parameter „SL & Hicks (2004)“ für die *Salt Correction* vorbelegt und „Medium – 0,50 °C“ für *Resolution* ausgewählt.

Berechnung Nr.	Thermodynamics	Free [Mg ⁺⁺] [mM]	[Mono ⁺] [mM]	DMSO [%]
1	Unified SL (1998)	3	20	0
2	Unified SL (1998)	6	20	0
3	Unified SL (1998)	3	40	0
4	Unified SL (1998)	3	20	2
5	Blake & Delcourt (1998)	3	20	0
6	Sugimoto et al. (1995)	3	20	0
7	Huguet et al. (2010)	3	20	0
8	Weber et al. (2015)	3	20	0

In Abbildung 61 sind die erhaltenen Schmelzkurven gezeigt, die auf denselben thermodynamischen Annahmen beruhen (Unified SL (1998)) jedoch auf Basis unterschiedlicher Reaktionsbedingungen berechnet wurden.

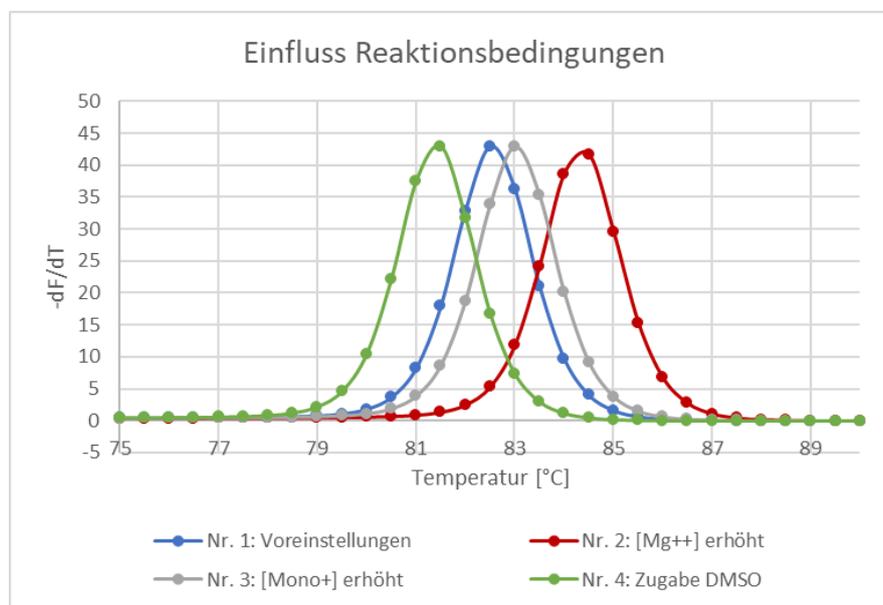


Abbildung 61: Berechnete Schmelzkurven für die unter 6.3.3.2 angegebene 70 bp lange DNA-Sequenz und den in Tabelle 44 angegebenen Parametern.

9.10.6.2 Methodenentwicklung und -validierung

Im Folgenden sind einige Abbildungen, Bezug nehmend auf Abschnitt 6.3.3.3, aufgeführt.

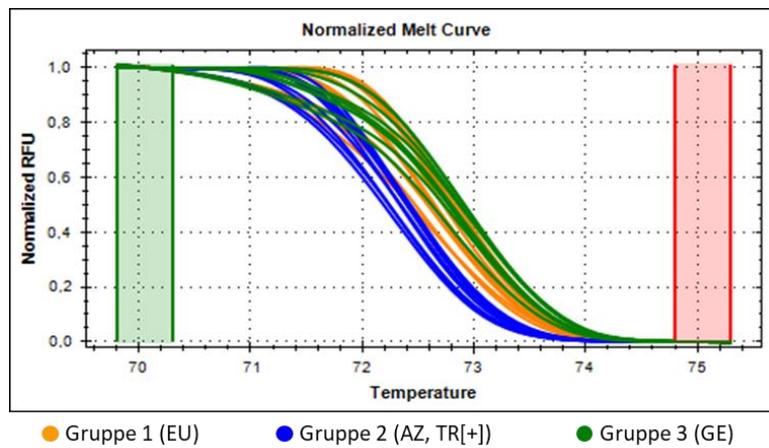


Abbildung 62: Normalisierte Schmelzkurven der in Tabelle 9 gekennzeichneten Proben (*), die für die Berechnung der in Abbildung 31 dargestellten Differenzkurven herangezogen wurden.

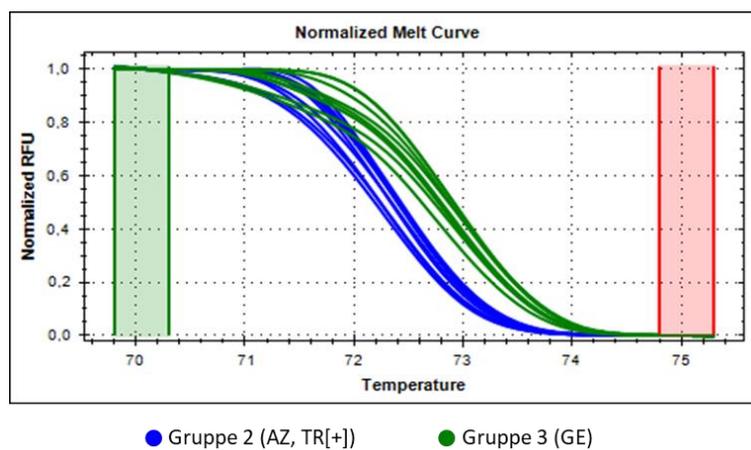


Abbildung 63: Normalisierte Schmelzkurven der in Tabelle 9 gekennzeichneten Proben (*) der Gruppen 2 und 3, die für die Berechnung der in Abbildung 64 dargestellten Differenzkurven herangezogen wurden.

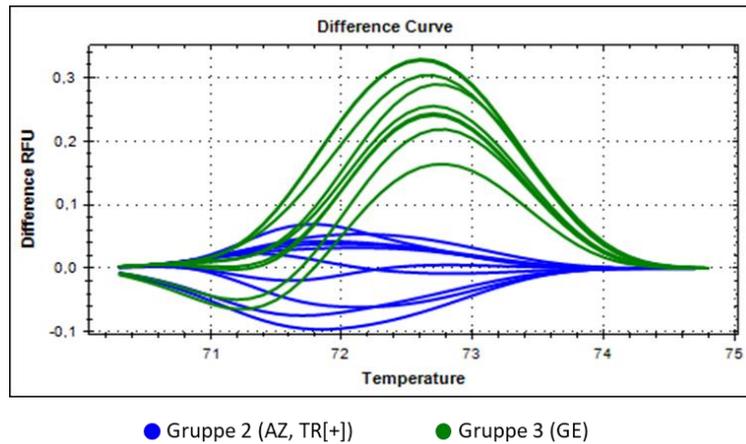


Abbildung 64: Differenzkurven berechnet über die in Abbildung 63 dargestellten normalisierten Schmelzkurven basierend auf den Kurven der Gruppe 2 (blau) als Referenz.

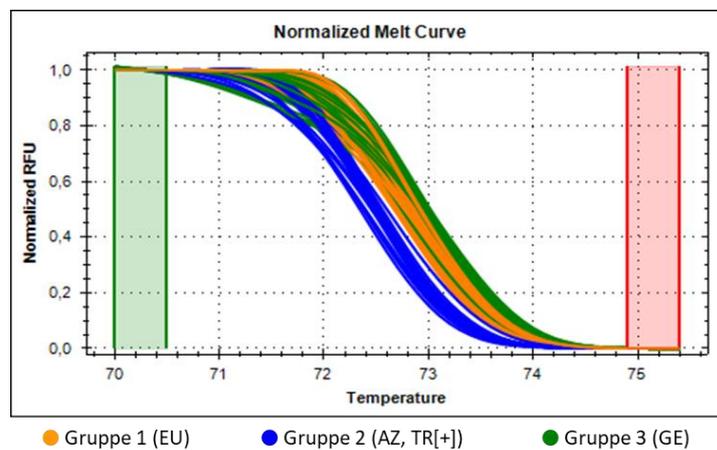


Abbildung 65: Normalisierte Schmelzkurven, die für die Berechnung der in Abbildung 32 dargestellten Differenzkurven herangezogen wurden. Die Abbildung enthält alle in Tabelle 10 aufgeführten Messungen.

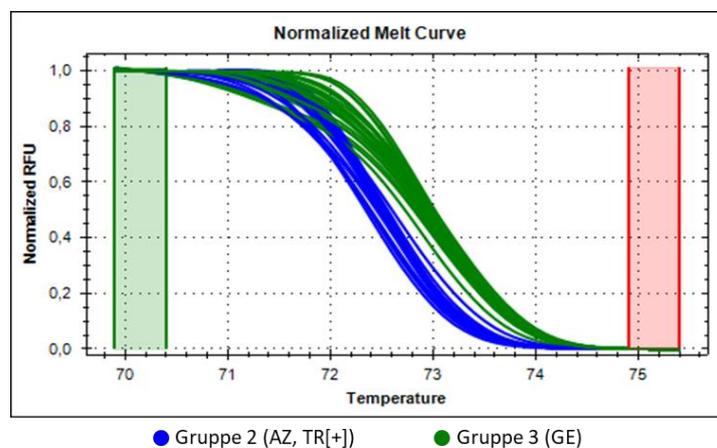


Abbildung 66: Normalisierte Schmelzkurven, die für die Berechnung der in Abbildung 67 dargestellten Differenzkurven herangezogen wurden. Die Abbildung enthält alle in Tabelle 10 aufgeführten Messungen der Proben der Gruppe 2 und 3.

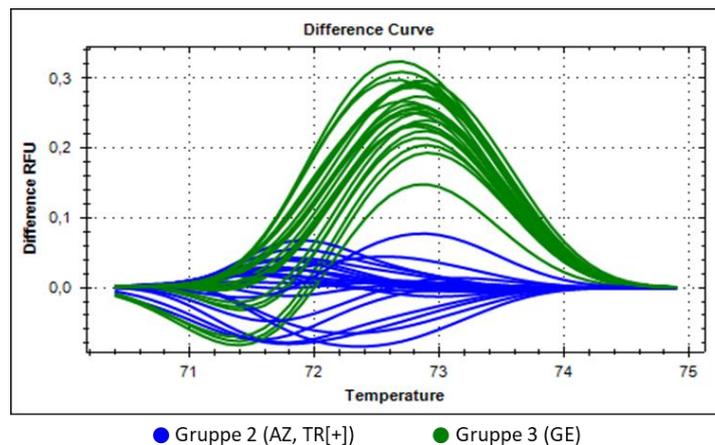


Abbildung 67: Differenzkurven berechnet über die in Abbildung 66 dargestellten normalisierten Schmelzkurven basierend auf den Kurven der Gruppe 2 (blau) als Referenz.

9.10.6.3 Quantifizierung mittels HRMA

Die folgenden Tabellen fassen die Werte zusammen, die für die Erstellung der Kalibriergeraden in Abbildung 36 und Abbildung 38 herangezogen wurden.

Tabelle 45: Daten für die Erstellung der in Abbildung 36 dargestellten Kalibriergeraden.

Anteil georgischer Haselnüsse [%]	ΔF
100	0,266
60	0,249
40	0,173
20	0,080
10	0,02
5	0,017
2,5	0

Tabelle 46: Daten für die Erstellung der in Abbildung 38 dargestellten Kalibriergeraden.

Anteil georgischer Haselnüsse [%]	ΔF	$\Delta F / \Delta F \text{ max. } [\%]$
100	0,286	---
70	0,157	55,033
50	0,095	33,160
30	0,063	22,070
10	0,031	11,010
0	0	---

Die folgenden beiden Abbildungen stellen den *Difference plot* und die daraus resultierende Kalibriergerade dar anhand derer die Möglichkeit der Quantifizierung unterhalb von 10% Beimischung untersucht wurde.

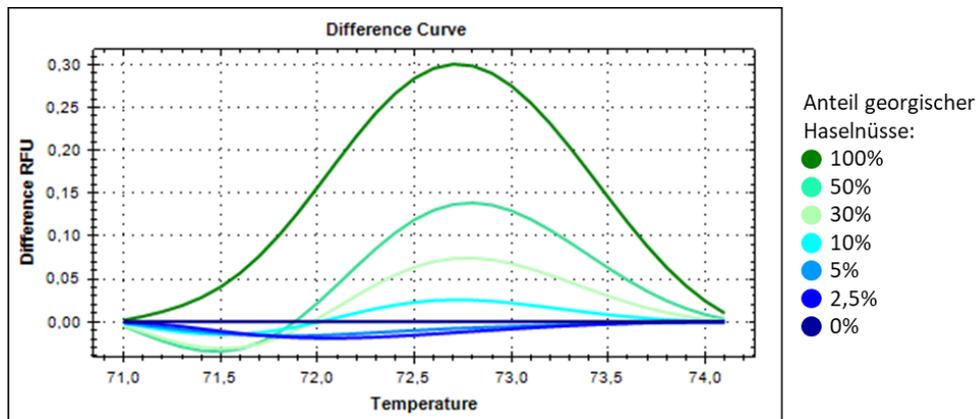


Abbildung 68: *Differenz plot* der Kalibrierreihe (Matrixmischungen) im Bereich von 2,5-100% bezogen auf den Anteil georgischer Haselnüsse.

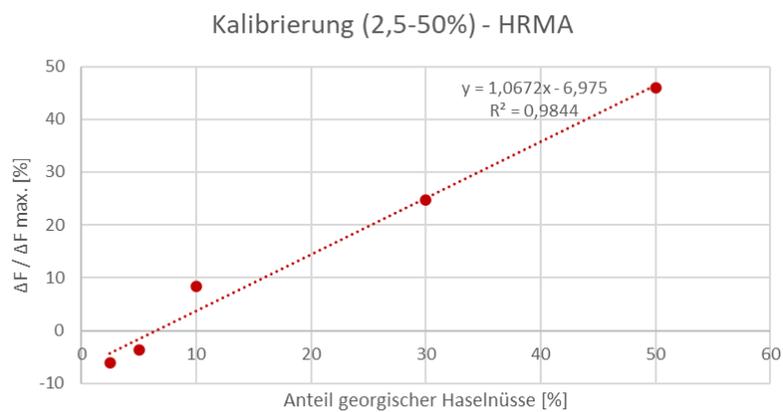


Abbildung 69: Kalibriergerade für den Bereich von 2,5-50% bezogen auf den Anteil georgischer Haselnüsse. Sie wurde auf Basis des in Abbildung 68 dargestellten *Difference plots* erstellt.

9.10.7 Analyse der ITS-Bereiche

Im Folgenden ist das Alignment veröffentlichter und selbst mittels Sanger-Sequenzierung generierter Sequenzen der ITS-Bereiche dargestellt. Die zugehörige Auswertung erfolgte im Abschnitt 6.4.1.

Casina	(EU)	TTTCCGTAGGTGAACCTGCGGAAGGATCATTGTGCGAAACCTGCCAGCAGAACGACCCGC
Tombul G.	(EU)	TTTCCGTAGGTGAACCTGCGGAAGGATCATTGTGCGAAACCTGCCAGCAGAACGACCCGC
Barcelona	(EU)	TTTCCGTAGGTGAACCTGCGGAAGGATCATTGTGCGAAACCTGCCAGCAGAACGACCCGC
Pauetet	(EU)	-----GNNNNNGNNACTGCGGAGNTCATTGTGCGAAACCTGCCAGCAGAACGACCCGC
Butler	(EU)	-----GNNNNNNNNACTGCGGAGGACATTGTGCGAAACCTGCCAGCAGAACGACCCGC
Ennis	(EU)	-----GNNNNNNNNNNNTGCGGAGANATTGTGCGAAACCTGCCAGCAGAACGACCCGC
Tombul	(TR)	-----CNNNNNNNNNTGCGGAGACATTGTGCGAAACCTGCCAGCAGAACGACCCGC
Segorbe	(EU)	-----GNNNNNNNNACTGCGGAGNTCATTGTGCGAAACCTGCCAGCAGAACGACCCGC
N.A.	(GE)	-----TTGTGCGAAACCTGCCAGCAGAACGACCCGC
N.A.	(GE)	-----TTGTGCGAAACCTGCCAGCAGAACGACCCGC
Ata Baba	(AZ)	----CGCTTAAGGGACTGCGGAAGGACATTGTGCGAAACCTGCCAGCAGAACGACCCGC

* *****

Casina	(EU)	GAACTTGATAAACAACCGGGGCGGGGGCGTCTCGCCCCGTGCCCCGAACGGCGGG
Tombul G.	(EU)	GAACTTGATAAACAACCGGGGCGGGGGCGTCTCGCCCCGTGCCCCGAACGGCGGG
Barcelona	(EU)	GAACTTGATAAACAACCGGGGCGGGGGCGTCTCGCCCCGTGCCCCGAACGGCGGG
Pauetet	(EU)	GAACTTGATAAACAACCGGGGCGGGGGCGTCTCGCCCCGTGCCCCGAACGGCGGG
Butler	(EU)	GAACTTGATAAACAACCGGGGCGGGGGCGTCTCGCCCCGTGCCCCGAACGGCGGG
Ennis	(EU)	GAACTTGATAAACAACCGGGGCGGGGGCGTCTCGCCCCGTGCCCCGAACGGCGGG
Tombul	(TR)	GAACTTGATAAACAACCGGGGCGGGGGCGTCTCGCCCCGTGCCCCGAACGGCGGG
Segorbe	(EU)	GAACTTGATAAACAACCGGGGCGGGGGCGTCTCGCCCCGTGCCCCGAACGGCGGG
N.A.	(GE)	GAACTTGATAAACAACCGGGGCGGGGGCGTCTCGCCCCGTGCCCCGAACGGCGGG
N.A.	(GE)	GAACTTGATAAACAACCGGGGCGGGGGCGTCTCGCCCCGTGCCCCGAACGGCGGG
Ata Baba	(AZ)	GAACTTGATAAACAACCGGGGCGGGGGCGTCTCGCCCCGTGCCCCGAACGGCGGG

Casina	(EU)	GAGACACTCGTGCCCTTCTTGCCGAACAACGAACCCCGGCGGGTCTGCGCCAAGGAAC
Tombul G.	(EU)	GAGACACTCGTGCCCTTCTTGCCGAACAACGAACCCCGGCGGGTCTGCGCCAAGGAAC
Barcelona	(EU)	GAGACACTCGTGCCCTTCTTGCCGAACAACGAACCCCGGCGGGTCTGCGCCAAGGAAC
Pauetet	(EU)	GAGACACTCGTGCCCTTCTTGCCGAACAACGAACCCCGGCGGGTCTGCGCCAAGGAAC
Butler	(EU)	GAGACACTCGTGCCCTTCTTGCCGAACAACGAACCCCGGCGGGTCTGCGCCAAGGAAC
Ennis	(EU)	GAGACACTCGTGCCCTTCTTGCCGAACAACGAACCCCGGCGGGTCTGCGCCAAGGAAC
Tombul	(TR)	GAGACACTCGTGCCCTTCTTGCCGAACAACGAACCCCGGCGGGTCTGCGCCAAGGAAC
Segorbe	(EU)	GAGACACTCGTGCCCTTCTTGCCGAACAACGAACCCCGGCGGGTCTGCGCCAAGGAAC
N.A.	(GE)	GAGACACTCGTGCCCTTCTTGCCGAACAACGAACCCCGGCGGGTCTGCGCCAAGGAAC
N.A.	(GE)	GAGACACTCGTGCCCTTCTTGCCGAACAACGAACCCCGGCGGGTCTGCGCCAAGGAAC
Ata Baba	(AZ)	GAGACACTCGTGCCCTTCTTGCCGAACAACGAACCCCGGCGGGTCTGCGCCAAGGAAC

Casina	(EU)	CAACTAAAGAGTGCCTCCGGTCGCTCGGAAACGGCGTGCCTGCGGGGGCGAATCTTGT
Tombul G.	(EU)	CAACTAAAGAGTGCCTCCGGTCGCTCGGAAACGGCGTGCCTGCGGGGGCGAATCTTGT
Barcelona	(EU)	CAACTAAAGAGTGCCTCCGGTCGCTCGGAAACGGCGTGCCTGCGGGGGCGAATCTTGT
Pauetet	(EU)	CAACTAAAGAGTGCCTCCGGTCGCTCGGAAACGGCGTGCCTGCGGGGGCGAATCTTGT
Butler	(EU)	CAACTAAAGAGTGCCTCCGGTCGCTCGGAAACGGCGTGCCTGCGGGGGCGAATCTTGT
Ennis	(EU)	CAACTAAAGAGTGCCTCCGGTCGCTCGGAAACGGCGTGCCTGCGGGGGCGAATCTTGT
Tombul	(TR)	CAACTAAAGAGTGCCTCCGGTCGCTCGGAAACGGCGTGCCTGCGGGGGCGAATCTTGT
Segorbe	(EU)	CAACTAAAGAGTGCCTCCGGTCGCTCGGAAACGGCGTGCCTGCGGGGGCGAATCTTGT
N.A.	(GE)	CAACTAAAGAGTGCCTCCGGTCGCTCGGAAACGGCGTGCCTGCGGGGGCGAATCTTGT
N.A.	(GE)	CAACTAAAGAGTGCCTCCGGTCGCTCGGAAACGGCGTGCCTGCGGGGGCGAATCTTGT
Ata Baba	(AZ)	CAACTAAAGAGTGCCTCCGGTCGCTCGGAAACGGCGTGCCTGCGGGGGCGAATCTTGT

Casina	(EU)	GCAAAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTA
Tombul G.	(EU)	GCAAAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTA
Barcelona	(EU)	GCAAAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTA
Pauetet	(EU)	GCAAAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTA
Butler	(EU)	GCAAAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTA
Ennis	(EU)	GCAAAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTA
Tombul	(TR)	GCAAAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTA
Segorbe	(EU)	GCAAAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTA
N.A.	(GE)	GCAAAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTA
N.A.	(GE)	GCAAAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTA
Ata Baba	(AZ)	GCAAAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTA

Casina (EU) GCGAAATGCGATACTTGGTGTGAATTGCAGAATCCC CGGAATCATCGAGTCTTTGAACGC
 Tombul G. (EU) GCGAAATGCGATACTTGGTGTGAATTGCAGAATCCC CGGAATCATCGAGTCTTTGAACGC
 Barcelona (EU) GCGAAATGCGATACTTGGTGTGAATTGCAGAATCCC CGGAATCATCGAGTCTTTGAACGC
 Pauetet (EU) GCGAAATGCGATACTTGGTGTGAATTGCAGAATCCC CGGAATCATCGAGTCTTTGAACGC
 Butler (EU) GCGAAATGCGATACTTGGTGTGAATTGCAGAATCCC CGGAATCATCGAGTCTTTGAACGC
 Ennis (EU) GCGAAATGCGATACTTGGTGTGAATTGCAGAATCCC CGGAATCATCGAGTCTTTGAACGC
 Tombul (TR) GCGAAATGCGATACTTGGTGTGAATTGCAGAATCCC CGGAATCATCGAGTCTTTGAACGC
 Segorbe (EU) GCGAAATGCGATACTTGGTGTGAATTGCAGAATCCC CGGAATCATCGAGTCTTTGAACGC
 N.A. (GE) GCGAAATGCGATACTTGGTGTGAATTGCAGAATCCC CGGAATCATCGAGTCTTTGAACGC
 N.A. (GE) GCGAAATGCGATACTTGGTGTGAATTGCAGAATCCC CGGAATCATCGAGTCTTTGAACGC
 Ata Baba (AZ) GCGAAATGCGATACTTGGTGTGAATTGCAGAATCCC CGGAATCATCGAGTCTTTGAACGC

Casina (EU) AAGTTGCGCCCCGAAGCCATCTGGTCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGC
 Tombul G. (EU) AAGTTGCGCCCCGAAGCCATCTGGTCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGC
 Barcelona (EU) AAGTTGCGCCCCGAAGCCATCTGGTCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGC
 Pauetet (EU) AAGTTGCGCCCCGAAGCCATCTGGTCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGC
 Butler (EU) AAGTTGCGCCCCGAAGCCATCTGGTCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGC
 Ennis (EU) AAGTTGCGCCCCGAAGCCATCTGGTCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGC
 Tombul (TR) AAGTTGCGCCCCGAAGCCATCTGGTCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGC
 Segorbe (EU) AAGTTGCGCCCCGAAGCCATCTGGTCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGC
 N.A. (GE) AAGTTGCGCCCCGAAGCCATCTGGTCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGC
 N.A. (GE) AAGTTGCGCCCCGAAGCCATCTGGTCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGC
 Ata Baba (AZ) AAGTTGCGCCCCGAAGCCATCTGGTCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGC

Casina (EU) CCCCACCCCATCGCCTCTCCAAGAGACGAGGGCGGTCTGCGGGGCGGACATTGGCCTCC
 Tombul G. (EU) CCCCACCCCATCGCCTCTCCAAGAGACGAGGGCGGTCTGCGGGGCGGACATTGGCCTCC
 Barcelona (EU) CCCCACCCCATCGCCTCTCCAAGAGACGAGGGCGGTCTGCGGGGCGGACATTGGCCTCC
 Pauetet (EU) CCCCACCCCATCGCCTCTCCAAGAGACGAGGGCGGTCTGCGGGGCGGACATTGGCCTCC
 Butler (EU) CCCCACCCCATCGCCTCTCCAAGAGACGAGGGCGGTCTGCGGGGCGGACATTGGCCTCC
 Ennis (EU) CCCCACCCCATCGCCTCTCCAAGAGACGAGGGCGGTCTGCGGGGCGGACATTGGCCTCC
 Tombul (TR) CCCCACCCCATCGCCTCTCCAAGAGACGAGGGCGGTCTGCGGGGCGGACATTGGCCTCC
 Segorbe (EU) CCCCACCCCATCGCCTCTCCAAGAGACGAGGGCGGTCTGCGGGGCGGACATTGGCCTCC
 N.A. (GE) CCCCACCCCATCGCCTCTCCAAGAGACGAGGGCGGTCTGCGGGGCGGACATTGGCCTCC
 N.A. (GE) CCCCACCCCATCGCCTCTCCAAGAGACGAGGGCGGTCTGCGGGGCGGACATTGGCCTCC
 Ata Baba (AZ) CCCCACCCCATCGCCTCTCCAAGAGACGAGGGCGGTCTGCGGGGCGGACATTGGCCTCC

Casina (EU) CGTGAGCTTCGCTCGCGGCTGGCCTAAAAGCGAGTCCCTCGGCACGAGCGCCACGACAA
 Tombul G. (EU) CGTGAGCTTCGCTCGCGGCTGGCCTAAAAGCGAGTCCCTCGGCACGAGCGCCACGACAA
 Barcelona (EU) CGTGAGCTTCGCTCGCGGCTGGCCTAAAAGCGAGTCCCTCGGCACGAGCGCCACGACAA
 Pauetet (EU) CGTGAGCTTCGCTCGCGGCTGGCCTAAAAGCGAGTCCCTCGGCACGAGCGCCACGACAA
 Butler (EU) CGTGAGCTTCGCTCGCGGCTGGCCTAAAAGCGAGTCCCTCGGCACGAGCGCCACGACAA
 Ennis (EU) CGTGAGCTTCGCTCGCGGCTGGCCTAAAAGCGAGTCCCTCGGCACGAGCGCCACGACAA
 Tombul (TR) CGTGAGCTTCGCTCGCGGCTGGCCTAAAAGCGAGTCCCTCGGCACGAGCGCCACGACAA
 Segorbe (EU) CGTGAGCTTCGCTCGCGGCTGGCCTAAAAGCGAGTCCCTCGGCACGAGCGCCACGACAA
 N.A. (GE) CGTGAGCTTCGCTCGCGGCTGGCCTAAAAGCGAGTCCCTCGGCACGAGCGCCACGACAA
 N.A. (GE) CGTGAGCTTCGCTCGCGGCTGGCCTAAAAGCGAGTCCCTCGGCACGAGCGCCACGACAA
 Ata Baba (AZ) CGTGAGCTTCGCTCGCGGCTGGCCTAAAAGCGAGTCCCTCGGCACGAGCGCCACGACAA

Casina (EU) TCGGTGGTTGACAAACCCTCGTGTCCCCTCGTGC GCGGCTCGTGCCTCGTCTTTGTGCTCT
 Tombul G. (EU) TCGGTGGTTGACAAACCCTCGTGTCCCCTCGTGC GCGGCTCGTGCCTCGTCTTTGTGCTCT
 Barcelona (EU) TCGGTGGTTGACAAACCCTCGTGTCCCCTCGTGC GCGGCTCGTGCCTCGTCTTTGTGCTCT
 Pauetet (EU) TCGGTGGTTGACAAACCCTCGTGTCCCCTCGTGC GCGGCTCGTGCCTCGTCTTTGTGCTCT
 Butler (EU) TCGGTGGTTGACAAACCCTCGTGTCCCCTCGTGC GCGGCTCGTGCCTCGTCTTTGTGCTCT
 Ennis (EU) TCGGTGGTTGACAAACCCTCGTGTCCCCTCGTGC GCGGCTCGTGCCTCGTCTTTGTGCTCT
 Tombul (TR) TCGGTGGTTGACAAACCCTCGTGTCCCCTCGTGC GCGGCTCGTGCCTCGTCTTTGTGCTCT
 Segorbe (EU) TCGGTGGTTGACAAACCCTCGTGTCCCCTCGTGC GCGGCTCGTGCCTCGTCTTTGTGCTCT
 N.A. (GE) TCGGTGGTTGACAAACCCTCGTGTCCCCTCGTGC GCGGCTCGTGCCTCGTCTTTGTGCTCT
 N.A. (GE) TCGGTGGTTGACAAACCCTCGTGTCCCCTCGTGC GCGGCTCGTGCCTCGTCTTTGTGCTCT
 Ata Baba (AZ) TCGGTGGTTGACAAACCCTCGTGTCCCCTCGTGC GCGGCTCGTGCCTCGTCTTTGTGCTCT

Casina (EU) GTGACCCGTAGCGTCGCGCTCGCGACTTCCAACGCGACCCAGGTCAGGCGGGACT-
 Tombul G. (EU) GTGACCCGTAGCGTCGCGCTCGCGACTTCCAACGCGACCCAGGTCAGGCGGGACT-
 Barcelona (EU) GTGACCCGTAGCGTCGCGCTCGCGACTTCCAACGCGACCCAGGTCAGGCGGGACT-
 Pauetet (EU) GTGACCCGTAGCGTCGCGCTCGCGACTTCCAACGCGACCCAGGTCAGGCGGGACTA
 Butler (EU) GTGACCCGTAGCGTCGCGCTCGCGACTTCCAACGCGACCCAGGTCAGGCGGGACTA
 Ennis (EU) GTGACCCGTAGCGTCGCGCTCGCGACTTCCAACGCGACCCAGGTCAGGCGGGACTA
 Tombul (TR) GTGACCCGTAGCGTCGCGCTCGCGACTTCCAACGCGACCCAGGTCAGGCGGGACTA
 Segorbe (EU) GTGACCCGTAGCGTCGCGCTCGCGACTTCCAACGCGACCCAGGTCAGGCGGGACTA
 N.A. (GE) GTGACCCGTAGCGTCGCGCTCGCGACTTCCAACGCGACCCAGGTCAGGCGGGACTA
 N.A. (GE) GTGACCCGTAGCGTCGCGCTCGCGACTTCCAACGCGACCCAGGTCAGGCGGGACTA
 Ata Baba (AZ) GTGACCCGTAGCGTCGCGCTCGCGACTTCCAACGCGACCCAGGTCAGGCGGGACTA

9.10.8 Analyse der WRKY-Gene

WRKY-280K1_FW

5' CCCTCACTAAAGGGACTAGTCTGCAGGTTTAAACGAATTTCGCCCTTGGAGGAAGTACGGGCAGAAG
 GTTGTCAAGGGAAATCCATATCCCAGGTTGAGATTTGTGTACATTTTATGTCGATACGATCTCTTAAA
 AGTCTGAAGTGTGGTCTTTCCCTTCCAGAGAATATTTCTTAAGCATCACTTGATCTGTCTTGCAGAA
 GTTACTACCGATGTACCAGTCTCAAATGCAACGTGCGCAAACATGTGGAAAGAGCATCAGATGATCCA
 AGAGCTTTTGTCAACACCTACGAAGGCAAACATAACCAAAGGGCGAATTTCGCGGCCGCT 3'

WRKY-280K1_RV

Abbildung 70: Sequenz der Kolonie K1 der 280 bp langen Bande nach erfolgter Klonierung inklusive des darauf basierenden entwickelten Primerpaares WRKY-280K1_FW/RV. Die Primerbindungsstellen sind rot hervorgehoben. In Grau die Sequenz des für die Klonierung eingesetzten Vektors.

WRKY-280K2_FW

5' CGAATTCGCCCTTTGGTTCGTGTTTGCCTTCATAGGTTGTGATTACCGACTTAGGATCATGCCATGCCCT
 CTCCACATGCTTCCTTACTGTGCAACCTGCACTGGTGCCTTGTAGTAACTCCTGCATAAACAAGTTGT
 CCCCATTTTTCAGTTTGGCAACAAGTTGTCCCCATTTATGAGGACAGACGAGAAGATAAATGAGTTCAG
 GTGATTATTATAGAAACCTTGGATTTGGATTTCTTTTCAACTTTCTGTCCATACTTCCGCCAAAGGG
 CGAATTCGCGGCCGCTAAATTCAATT 3'

WRKY-280K2_RV

Abbildung 71: Sequenz der Kolonie K2 der 280 bp langen Bande nach erfolgter Klonierung inklusive des darauf basierenden entwickelten Primerpaares WRKY-280K2_FW/RV. Die Primerbindungsstellen sind rot hervorgehoben. In Grau die Sequenz des für die Klonierung eingesetzten Vektors.

WRKY-320K3_FW

5' TTTAAACGAATTTCGCCCTTTGGCGGAAATACGGGCAAAAATTGTGTCAAAGGAAATCCAAACCCTAGGTA
 TCAAAATATAAAAATGCTCCATGTGTGTGTTCTCCATTGAAATAAAAATACCTCACCGGAGTGATTGGG
 AGTTGTGTGGTGTGTATCCTCAAATTACATACTTTTATTCATGTTTTACTTGTGTTGATAGTTGGTTTT
 GAGTATTAAGCTTATGTCTCTCCTACAGGAGCTACTACAAATGCACGAGTGCTGGATGTTCTGTGCGGA
 AGCATGTGGAAAGGGCCTCTCACAAATATAAATGTGTTCTTACAACCTACGAGGGCAAACACAACCAA
 GGGCGAATTTCGCGGCCGCTAA 3'

WRKY-320K3_RV

Abbildung 72: Sequenz der Kolonie K3 der 320 bp langen Bande nach erfolgter Klonierung inklusive des darauf basierenden entwickelten Primerpaares WRKY-320K3_FW/RV. Die Primerbindungsstellen sind rot hervorgehoben. In Grau die Sequenz des für die Klonierung eingesetzten Vektors.

WRKY-950K16 FW

5' TCCTAAAGGGACTAGTCTGTCAGGTTTAAACGAATTCGCCCTTGGCGGAAGTATGGACAGAAAAGTGG
 TTAAGGGCAATCCTTATCCAAGGTAGTACTACTTTTCTTATCTTTTGGTCCCCCTTCACCAATGTGGT
 TAAGTTTTCTTTATAATCTATAACAATTTACATTTAGCTCATCGAAGTATATTCATAGAGCTATCTCTT
 AACAAATCTAGTGTTCATATTCTGAGATCTTTTTGCACACCATCGTACTGAAGAGGTTGATAATAG
 ATGTATTCTTATAAAAAATTGCAAAATGAAAAGGGGGATTACATCTTATTCAAAATCTTTAACTCCGTG
 AGACTTAAATCCTCTCAGCTCTTATAACCATGTTTAAAACCTGCCTCTCTTAAATCAAATATTGTCAT
 TGCTATCTCTAACTCCCCCATCATCTCTGCCATATTTTGGCCATACTGAAACTGTCATGACTGCCTTC
 ATGCTTGTTTAGTATCTATCAACCGGCAGTCTCTGATTTTTTTTCTTACACTAGAGTATTTATTCCTAG
 CCTATTTGAGTTGAGCTAATCCCCATGGATGCAGTAATCACCTGCAAACCTACATGCATCCCCAAGCC
 TGGGGAGTCGATCAATGTCTCAACTAAATCATTACCTAAATGGGTGCGAAAACAAAAATTTGGAATT
 TTTAAGATCATGTTCGTGCATGAGAACTCTTTTAGTTGCTTTTTTGAATATGTTGATGTCAAATGTGT
 TTTTCATCATGCCTAATTAATCCTCTCAATGGATGCTTAAAAAAGTAATTAATCTATTCTCAGTGATTG
 GTTTAATTTCTTGTTCATGGGTTGCTTATCTGGAAATTCCTTATGCATCTTTTTTTCTGATCTGCA
 GGAGCTATTATAAATGCACAAGCCCGGGATGTAATGTCCGTAAGCACGTCGAGAGGGCTGCGTCAGAT
 CCTAAAGCTGTCATTACAACGTACGAGGGCAAGCACAGCCAAAGGGCGAATTCGCGGCCGCTAAA 3'

WRKY-950K16_RV

Abbildung 73: Sequenz der Kolonie K16 der 950 bp langen Bande nach erfolgter Klonierung inklusive des darauf basierenden entwickelten Primerpaares WRKY-950K16_FW/RV. Die Primerbindungsstellen sind rot hervorgehoben. In Grau die Sequenz des für die Klonierung eingesetzten Vektors.

Im Folgenden sind die Alignments der Amplifikate für jeweils drei Proben nach der PCR mit den Primerpaaren WRKY-280K2_FW/RV und WRKY-320K3_FW/RV dargestellt.

280 bp - K2:

WRKY_280K2_Pauetet (FR)	-----GGTGCACCTGTAGTAACCTCCTGCATAAACAAGTTGTCCCCATTTTCAGTTTGCCA
WRKY_280K2_Palaz (TR)	GCACTGGTGCACCTGTAGTAACCTCCTGCATAAACAAGTTGTCCCCATTTTCAGTTTGCCA
WRKY_280K2_Ata Baba (AZ)	GCACTGGTGCACCTGTAGTAACCTCCTGCATAAACAAGTTGTCCCCATTTTCAGTTTGCCA

WRKY_280K2_Pauetet (FR)	ACAAGTTGTCCCCATTTTAAAGAGGACAAACGAGAAGATAAATGAGTTCACGGTGATTATTA
WRKY_280K2_Palaz (TR)	ACAAGTTGTCCCCATTTTAAAGAGGACAAACGAGAAGATAAATGAGTTCACGGTGATTATTA
WRKY_280K2_Ata Baba (AZ)	ACAAGTTGTCCCCATTTTAAAGAGGACAAACGAGAAGATAAATGAGTTCACGGTGATTATTA

WRKY_280K2_Pauetet (FR)	TAGAAACCTTGGATTTGGATTTTCCTTTTACAACCTTCTGTCCATACTTCCGC
WRKY_280K2_Palaz (TR)	TAGAAACCTTGGATTTGGATTTTCCTTTTACAACCTTCTGTCCATACTTCCGC
WRKY_280K2_Ata Baba (AZ)	TAGAAACCTTGGATTTGGATTTTCCTTTTACAACCTTCTGTCCATACTTCCGC

320 bp - K3:

```

WRKY_320K3_Ata Baba (AZ)      ----GTGTGTTCTCCCATTTGAAATAAAAATACCTCACCAGAGTGATTGGGAGTTGTGTGGT
WRKY_320K3_Pauetet (FR)      -TATGTGTGTTCTCCCATTTGAAATAAAAATACCTCACCAGAGTGATTGGGAGTTGTGTGGT
WRKY_320K3_Palaz (TR)        ATATGTGTGTTCTCCCATTTGAAATAAAAATACCTCACCAGAGTGATTGGGAGTTGTGTGGT
                               *****

WRKY_320K3_Ata Baba (AZ)      GTTTATCCTCAAATTACATACTTTTATTCCATGTTTTACTTGTTTGATAGTTGGTTTTGA
WRKY_320K3_Pauetet (FR)      GTTTATCCTCAAATTACATACTTTTATTCCATGTTTTACTTGTTTGATAGTTGGTTTTGA
WRKY_320K3_Palaz (TR)        GTTTATCCTCAAATTACATACTTTTATTCCATGTTTTACTTGTTTGATAGTTGGTTTTGA
                               *****

WRKY_320K3_Ata Baba (AZ)      GTATTAAGCTTATGTCTCTCCTACAGGAGCTACTACAAATGCACGAGTGCTGGATGTTCT
WRKY_320K3_Pauetet (FR)      GTATTAAGCTTATGTCTCTCCTACAGGAGCTACTACAAATGCACGAGTGCTGGATGTTCT
WRKY_320K3_Palaz (TR)        GTATTAAGCTTATGTCTCTCCTACAGGAGCTACTACAAATGCACGAGTGCTGGATGTTCT
                               *****

WRKY_320K3_Ata Baba (AZ)      GTGCGGAAGCATGTGGAAGGGCCTCTC
WRKY_320K3_Pauetet (FR)      GTGCGGAAGCATGTGGAAGGGCCTCTC
WRKY_320K3_Palaz (TR)        GTGCGGAAGCATGTGGAAGGGCCTCTC
                               *****
    
```

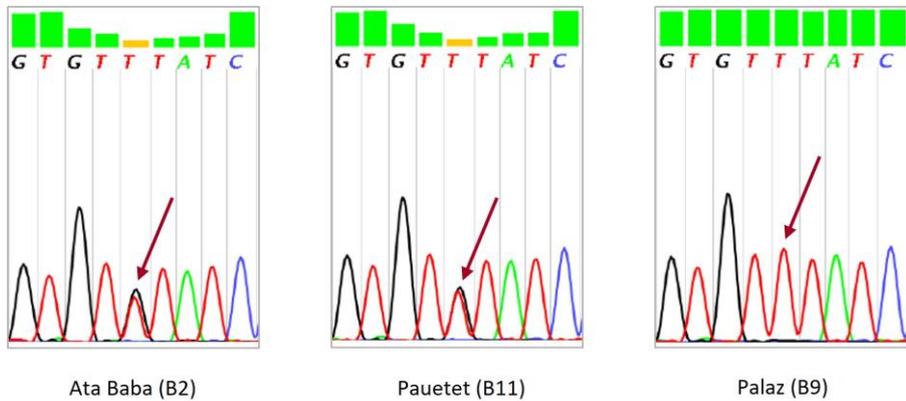


Abbildung 74: Ausschnitt der Sequenzierungsergebnisse der PCR-Produkte amplifiziert mit dem Primerpaar WRKY-320K3_FW/RV. Bei den Proben Ata Baba (B2) und Pauetet (B11) ist eine deutliche Signalüberlagerung der Basen T und G zu erkennen. Die entsprechende Base ist mit einem Pfeil gekennzeichnet. Rechts im Vergleich das Ergebnis der Sequenzierung der Sorte Palaz (B9) ohne Signalüberlagerung.

9.10.9 Analyse von SCoT Polymorphismen

Die folgenden Tabellen beinhalten die 1/0-Matrix für die Auswertung der SCoT-Analyse mit Hilfe multivariater Methoden.

Tabelle 47: 1/0-Matrix für die multivariate Auswertung der SCoT-Analyse. Der Tabellenüberschrift ist jeweils in der ersten Zeile in eckiger Klammer der Primer und in der zweiten Zeile die Bandenlänge in [bp] zu entnehmen.

Nr.	Sorte	[19]	[19]	[19]	[19]	[19]	[19]	[19]	[19]	[19]	[19]	[19]	[19]
		4000	2600	2500	2400	2200	2000	1900	1800	1850	1750	1600	1500
S2	Ata Baba	0	0	1	0	1	0	1	0	0	0	0	0
102	Ata Baba	0	0	1	0	1	0	1	0	0	0	0	0
104	Çakildak	1	0	1	0	1	0	1	0	0	0	0	0
103	Çakildak	0	0	1	0	1	0	1	0	0	0	0	0
105	Çakildak	0	0	1	0	1	0	1	0	0	0	0	0
107	Mortarella	0	0	0	0	1	0	1	0	0	0	0	1
106	Mortarella	0	0	0	0	1	0	1	0	0	0	0	1
S1	Mortarella	0	0	0	0	1	0	1	0	0	0	0	1
2	Negret	0	1	0	0	1	0	0	0	1	0	1	0
108	Negret	0	0	0	0	1	0	0	0	1	0	1	0
40	Butler	0	0	0	0	1	0	0	0	1	0	0	1
57	Butler	0	1	0	0	1	0	0	0	1	0	0	1
109	Butler	0	0	0	0	1	0	0	0	0	0	0	0
F2	Corabel	0	0	0	0	1	0	1	0	0	0	1	0
32	Corabel	0	0	0	0	1	0	0	1	0	0	1	0
35	Corabel	0	0	0	0	1	0	0	1	0	0	1	0
46	Ennis	0	0	0	0	1	0	0	1	0	0	1	0
77	Ennis	0	0	0	1	0	0	1	0	0	0	0	0
59	Ennis	0	0	0	0	1	0	0	1	0	0	0	0
43	Barcelona	0	0	0	0	0	0	1	0	0	0	0	0
52	Barcelona	0	0	0	0	0	0	1	0	0	0	0	0
82	Barcelona	0	0	0	0	0	0	1	0	0	0	0	0
F1	Pauetet	0	0	0	0	1	0	1	0	0	0	1	0
31	Pauetet	0	0	0	0	1	0	1	0	0	0	1	0
33	Pauetet	0	0	0	0	1	0	1	0	0	0	0	0
34	Segorbe	0	0	0	0	1	0	1	0	0	0	0	0
37	Segorbe	0	0	0	0	1	0	1	0	0	0	0	0
45	Segorbe	0	0	0	0	1	0	1	0	0	0	0	0
112	Tombul	0	0	0	0	1	0	1	0	0	0	0	1
111	Tombul	0	0	0	0	1	0	1	0	0	0	0	1
120	Tombul	0	0	0	0	1	0	1	0	0	0	0	1
F3	T. G. Romana	0	0	0	0	1	0	0	0	0	0	0	0
7	T. G. Romana	0	0	0	0	0	1	0	0	0	0	0	0
113	T. G. Romana	0	0	0	0	0	1	0	0	0	0	0	0
F4	Tonda di Giffoni	0	0	0	0	1	1	0	0	0	0	1	0
8	Tonda di Giffoni	0	0	0	0	1	1	0	0	0	1	0	1
22	Tonda di Giffoni	0	0	0	0	1	1	0	0	0	0	0	0

Tabelle 48: Fortsetzung von Tabelle 47.

Nr.	Sorte	[19]	[19]	[19]	[19]	[19]	[19]	[19]	[19]	[19]	[19]	[19]	[13]
		1400	1350	1300	1200	1150	1100	1000	950	900	800	650	3000
S2	Ata Baba	0	1	0	1	0	0	0	1	0	1	1	0
102	Ata Baba	0	0	0	1	0	0	0	1	0	1	1	0
104	Çakildak	0	1	0	1	0	0	0	1	0	1	1	0
103	Çakildak	0	0	0	1	0	0	0	1	0	1	1	0
105	Çakildak	0	0	0	1	0	0	0	1	0	1	1	0
107	Mortarella	0	0	0	1	0	1	0	1	0	1	1	0
106	Mortarella	0	0	0	1	0	1	0	1	0	0	1	0
S1	Mortarella	0	0	0	1	0	1	0	0	0	0	1	0
2	Negret	0	0	0	1	0	0	0	1	0	1	1	0
108	Negret	0	0	0	1	0	0	0	1	0	1	1	0
40	Butler	0	0	0	1	0	0	0	0	1	1	1	0
57	Butler	0	0	1	1	0	0	0	0	1	1	1	0
109	Butler	1	0	0	1	0	0	0	0	1	1	1	0
F2	Corabel	0	0	1	0	0	1	0	0	1	0	1	0
32	Corabel	0	0	0	0	0	0	0	0	1	0	1	0
35	Corabel	0	0	0	0	0	0	0	0	1	1	1	0
46	Ennis	0	0	1	1	0	0	0	0	1	0	1	0
77	Ennis	0	0	0	0	0	0	0	0	1	1	1	1
59	Ennis	0	0	0	1	0	0	0	0	1	0	1	0
43	Barcelona	0	0	0	1	0	0	0	0	1	0	1	0
52	Barcelona	0	0	0	1	0	0	0	0	1	0	1	0
82	Barcelona	0	0	0	1	0	0	0	0	1	0	1	0
F1	Pauetet	0	0	0	1	0	0	0	0	1	0	1	0
31	Pauetet	0	0	0	1	0	0	0	0	1	0	1	0
33	Pauetet	0	0	0	1	0	0	0	0	1	0	1	0
34	Segorbe	0	0	0	0	0	1	0	0	0	0	1	0
37	Segorbe	0	0	0	0	0	1	0	0	0	0	1	0
45	Segorbe	0	0	0	0	0	1	0	0	0	0	1	0
112	Tombul	0	0	1	1	0	0	0	0	0	0	1	0
111	Tombul	0	0	1	1	0	0	1	0	0	0	1	0
120	Tombul	0	0	0	1	0	0	0	0	0	0	1	0
F3	T. G. Romana	0	0	0	1	0	0	0	1	0	0	1	0
7	T. G. Romana	0	0	0	1	0	0	1	0	0	0	1	0
113	T. G. Romana	0	0	0	1	0	0	0	0	0	0	0	0
F4	Tonda di Giffoni	0	0	0	0	1	0	1	0	0	0	1	0
8	Tonda di Giffoni	0	0	1	0	0	1	1	0	0	1	1	0
22	Tonda di Giffoni	0	0	0	1	0	0	1	0	0	1	1	0

Tabelle 49: Fortsetzung von Tabelle 48.

Nr.	Sorte	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[13]
		2700	2600	2500	2450	2400	2350	2250	2200	2100	2000	1900	1800
S2	Ata Baba	0	0	0	0	0	0	0	1	0	0	1	0
102	Ata Baba	0	0	0	0	0	0	0	1	0	0	1	0
104	Çakildak	0	0	0	0	0	0	0	1	0	0	0	0
103	Çakildak	0	0	0	0	0	0	0	1	0	0	0	0
105	Çakildak	0	0	0	0	0	0	0	1	0	0	0	0
107	Mortarella	0	0	1	0	0	0	0	1	0	0	0	0
106	Mortarella	0	0	0	0	0	0	0	1	0	0	0	0
S1	Mortarella	0	0	0	0	0	0	0	1	0	0	0	0
2	Negret	0	0	0	1	0	0	0	0	0	1	0	0
108	Negret	0	1	0	1	0	0	0	0	1	0	0	0
40	Butler	0	0	0	0	0	0	0	0	0	0	0	0
57	Butler	0	0	0	0	0	0	0	1	0	0	0	0
109	Butler	0	0	0	0	0	0	0	0	0	0	0	0
F2	Corabel	0	0	0	1	0	0	0	1	0	0	0	0
32	Corabel	0	0	0	0	0	0	0	0	0	0	0	0
35	Corabel	0	0	0	0	0	0	0	1	0	0	0	0
46	Ennis	1	0	0	0	0	0	0	1	0	0	0	0
77	Ennis	0	0	0	0	0	0	0	1	0	0	0	0
59	Ennis	0	0	0	0	0	0	1	0	0	0	0	0
43	Barcelona	0	0	1	0	0	0	0	0	0	0	0	0
52	Barcelona	0	0	0	0	0	1	0	0	0	0	0	0
82	Barcelona	0	0	0	0	0	0	0	0	0	0	0	0
F1	Pauetet	0	0	1	0	0	0	0	0	0	0	0	0
31	Pauetet	0	0	0	0	0	1	0	0	0	0	0	0
33	Pauetet	0	0	1	0	0	0	0	0	0	0	0	0
34	Segorbe	0	0	0	0	1	0	0	0	0	0	0	0
37	Segorbe	0	0	0	0	0	0	0	0	0	0	0	0
45	Segorbe	0	0	1	0	0	0	0	0	0	0	0	0
112	Tombul	0	0	1	0	0	0	0	0	0	0	0	1
111	Tombul	0	0	1	0	0	0	0	0	0	1	0	1
120	Tombul	0	0	0	0	0	0	0	0	0	1	0	1
F3	T. G. Romana	0	0	0	0	0	0	0	1	0	0	0	0
7	T. G. Romana	0	0	0	0	0	0	0	1	0	0	0	0
113	T. G. Romana	0	0	0	0	0	0	0	0	0	0	1	0
F4	Tonda di Giffoni	0	0	0	0	0	0	0	1	0	0	0	0
8	Tonda di Giffoni	0	0	0	0	0	0	0	1	0	0	1	0
22	Tonda di Giffoni	0	0	0	0	1	0	0	0	0	0	1	0

Tabelle 50: Fortsetzung von Tabelle 49.

Nr.	Sorte	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[13]
		1750	1700	1600	1550	1500	1400	1300	1350	1200	1100	1050	1000
S2	Ata Baba	0	0	1	0	1	0	0	1	0	0	0	1
102	Ata Baba	0	0	1	0	1	0	0	1	0	0	0	1
104	Çakildak	0	0	1	0	1	0	0	1	0	0	0	1
103	Çakildak	0	0	0	0	1	0	0	0	0	0	0	1
105	Çakildak	0	0	1	0	1	0	0	1	0	0	0	0
107	Mortarella	0	1	0	0	1	0	0	0	0	0	0	1
106	Mortarella	0	0	0	0	1	0	0	1	0	0	0	1
S1	Mortarella	0	1	0	0	1	0	0	1	0	0	0	1
2	Negret	0	0	0	0	1	0	0	0	1	0	0	1
108	Negret	1	0	0	0	1	0	0	0	1	0	0	1
40	Butler	0	0	1	0	1	0	0	0	1	0	0	1
57	Butler	1	0	0	0	1	0	0	1	1	0	0	1
109	Butler	0	0	0	0	1	0	0	1	0	0	0	1
F2	Corabel	0	0	0	0	1	0	0	1	0	0	0	1
32	Corabel	0	0	0	0	1	0	0	1	0	0	0	1
35	Corabel	0	0	1	0	1	0	0	0	0	0	0	1
46	Ennis	0	0	0	0	1	0	0	0	1	0	0	1
77	Ennis	0	0	0	0	1	0	0	1	0	0	0	1
59	Ennis	0	0	0	0	1	0	0	1	0	0	0	1
43	Barcelona	0	0	0	1	0	1	0	0	1	0	0	1
52	Barcelona	0	0	0	0	1	0	0	0	1	0	0	1
82	Barcelona	0	0	0	0	1	0	0	0	1	0	0	1
F1	Pauetet	1	0	0	0	1	0	0	0	0	0	0	1
31	Pauetet	0	0	0	0	1	0	0	0	1	0	0	1
33	Pauetet	1	0	0	0	1	0	0	0	1	0	0	1
34	Segorbe	1	0	0	0	1	0	0	1	0	1	1	0
37	Segorbe	1	0	0	0	1	0	0	1	0	0	0	1
45	Segorbe	0	0	0	0	1	0	0	1	0	0	0	1
112	Tombul	0	0	0	0	1	0	1	0	0	0	0	1
111	Tombul	0	0	0	0	1	0	1	0	0	1	0	1
120	Tombul	0	0	0	0	1	0	1	0	0	1	0	1
F3	T. G. Romana	0	0	1	0	1	0	1	0	0	0	0	1
7	T. G. Romana	0	0	1	0	1	0	0	0	1	1	0	0
113	T. G. Romana	0	0	0	0	1	0	0	0	1	0	0	1
F4	Tonda di Giffoni	1	0	1	0	0	0	1	0	1	1	0	0
8	Tonda di Giffoni	0	0	1	0	0	0	1	0	0	1	0	1
22	Tonda di Giffoni	0	0	1	0	1	0	0	0	0	1	0	1

Tabelle 51: Fortsetzung von Tabelle 50.

Nr.	Sorte	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[13]	[32]
		950	900	850	750	700	650	600	550	500	450	350	2700
S2	Ata Baba	0	0	0	0	0	0	0	0	0	0	0	0
102	Ata Baba	0	0	0	0	0	0	0	0	0	0	0	0
104	Çakildak	0	0	0	0	0	0	0	0	0	0	0	0
103	Çakildak	0	0	0	0	1	0	0	0	0	0	0	0
105	Çakildak	0	0	0	0	0	0	0	0	0	0	0	0
107	Mortarella	0	0	0	0	1	0	0	0	0	0	0	0
106	Mortarella	0	0	0	0	0	0	0	0	0	0	0	0
S1	Mortarella	0	0	0	0	0	0	0	0	0	0	0	0
2	Negret	0	0	0	0	0	0	0	0	0	0	0	0
108	Negret	0	0	0	0	0	0	0	0	0	0	0	1
40	Butler	0	0	0	0	0	0	0	0	0	0	0	0
57	Butler	0	0	0	0	0	0	0	0	0	0	0	0
109	Butler	0	0	0	0	1	0	1	0	0	1	1	0
F2	Corabel	0	0	0	0	0	0	0	0	0	0	0	0
32	Corabel	1	0	0	0	0	0	0	0	0	0	0	0
35	Corabel	0	1	0	0	0	0	0	0	0	0	0	0
46	Ennis	0	0	0	0	1	0	0	0	0	0	0	0
77	Ennis	0	0	0	0	1	0	0	0	1	0	0	0
59	Ennis	0	0	0	0	1	0	0	0	1	0	0	0
43	Barcelona	0	1	0	0	0	0	0	0	0	0	0	0
52	Barcelona	0	1	0	1	0	0	0	0	0	0	0	0
82	Barcelona	0	0	1	0	0	0	0	0	0	0	0	0
F1	Pauetet	0	1	0	0	0	0	0	0	0	0	0	0
31	Pauetet	0	0	0	1	0	0	0	0	0	0	1	0
33	Pauetet	0	1	0	1	0	0	0	0	0	0	0	0
34	Segorbe	0	0	0	1	0	1	0	0	0	0	0	0
37	Segorbe	1	0	0	0	0	0	0	0	0	0	0	0
45	Segorbe	0	0	0	0	0	0	0	0	0	0	0	0
112	Tombul	0	0	1	0	0	0	0	0	0	0	0	0
111	Tombul	0	0	1	0	0	0	0	0	0	0	0	0
120	Tombul	0	1	0	0	0	1	0	0	0	0	0	0
F3	T. G. Romana	0	0	0	1	0	0	0	0	0	0	0	0
7	T. G. Romana	0	0	0	0	1	0	0	0	0	0	0	0
113	T. G. Romana	0	0	0	0	0	0	0	0	0	0	0	0
F4	Tonda di Giffoni	0	0	1	0	0	0	0	0	0	0	0	0
8	Tonda di Giffoni	0	0	0	1	0	0	0	0	1	0	0	0
22	Tonda di Giffoni	0	0	0	1	0	0	0	1	0	0	0	0

Tabelle 52: Fortsetzung von Tabelle 51.

Nr.	Sorte	[32]	[32]	[32]	[32]	[32]	[32]	[32]	[32]	[32]	[32]	[32]	[32]
		2200	2100	2000	1500	1350	1300	1250	1200	1100	1000	900	800
S2	Ata Baba	1	0	1	0	0	0	1	1	0	1	0	1
102	Ata Baba	1	0	1	0	0	0	0	1	0	0	0	1
104	Çakildak	1	0	1	1	0	1	0	1	0	0	0	1
103	Çakildak	1	0	0	0	0	0	0	0	0	0	0	0
105	Çakildak	1	0	0	0	0	0	1	1	1	0	0	1
107	Mortarella	1	0	0	0	0	0	0	1	1	0	0	1
106	Mortarella	1	1	0	0	0	0	0	1	1	0	0	1
S1	Mortarella	1	0	0	0	0	0	0	1	1	0	0	1
2	Negret	1	0	0	0	0	0	0	1	1	0	0	1
108	Negret	1	0	0	0	0	0	0	1	1	0	0	1
40	Butler	1	0	0	0	0	0	0	1	0	0	0	1
57	Butler	0	0	0	0	0	0	0	1	0	0	0	1
109	Butler	0	0	0	0	0	0	0	1	0	0	0	1
F2	Corabel	0	0	1	0	0	0	0	1	1	0	0	1
32	Corabel	0	0	1	0	0	0	0	1	0	0	0	1
35	Corabel	0	0	1	0	0	0	0	1	0	0	0	1
46	Ennis	1	0	1	0	0	0	0	1	0	0	0	1
77	Ennis	1	0	0	0	0	0	0	1	0	0	0	1
59	Ennis	1	0	0	0	0	0	0	1	1	0	0	1
43	Barcelona	1	0	0	0	0	0	0	1	0	0	0	1
52	Barcelona	1	0	1	0	0	0	0	1	0	0	0	1
82	Barcelona	1	0	0	0	0	0	0	1	0	0	0	1
F1	Pauetet	1	0	1	0	1	0	0	1	0	0	0	1
31	Pauetet	1	0	1	0	0	0	0	1	0	0	0	1
33	Pauetet	1	0	1	0	0	0	0	1	0	0	0	1
34	Segorbe	1	0	0	0	0	0	0	1	1	1	0	1
37	Segorbe	1	0	0	0	0	0	0	1	1	1	0	1
45	Segorbe	1	0	0	0	0	0	0	1	1	1	0	1
112	Tombul	1	0	1	1	0	1	0	1	1	1	0	1
111	Tombul	1	0	1	1	0	1	0	1	1	1	1	1
120	Tombul	1	0	1	1	0	0	0	1	1	1	1	1
F3	T. G. Romana	1	0	0	0	0	1	0	1	0	0	0	1
7	T. G. Romana	1	0	1	0	0	0	0	1	0	0	0	1
113	T. G. Romana	1	0	1	0	0	0	0	1	0	0	0	1
F4	Tonda di Giffoni	1	0	0	0	0	0	0	1	0	0	0	1
8	Tonda di Giffoni	1	0	0	0	0	0	0	1	0	0	0	1
22	Tonda di Giffoni	1	0	0	0	0	0	0	1	0	0	0	1

Tabelle 53: Fortsetzung von Tabelle 52.

Nr.	Sorte	[32]	[32]	[32]	[7]	[7]	[7]	[7]	[7]	[7]	[7]	[7]	[7]
		700	600	500	3500	3000	2500	2000	1900	1700	1600	1500	1300
S2	Ata Baba	0	0	0	0	1	1	0	0	0	1	1	0
102	Ata Baba	0	1	0	0	0	1	1	0	0	1	1	0
104	Çakildak	0	0	0	1	1	1	0	0	0	0	1	0
103	Çakildak	0	0	0	1	1	1	0	0	0	0	1	0
105	Çakildak	0	0	1	1	0	1	0	0	0	0	1	0
107	Mortarella	0	0	0	1	1	1	0	0	0	0	1	0
106	Mortarella	0	1	0	0	1	1	0	0	0	0	1	0
S1	Mortarella	0	0	0	1	1	1	1	0	0	0	1	1
2	Negret	0	1	0	0	1	1	1	0	1	1	1	0
108	Negret	1	0	0	0	1	1	1	1	0	0	1	0
40	Butler	0	1	0	0	0	1	1	0	0	0	1	0
57	Butler	0	1	0	0	0	1	0	0	0	0	1	0
109	Butler	0	1	0	0	0	0	0	0	0	0	0	0
F2	Corabel	0	0	0	0	0	1	0	0	0	0	1	0
32	Corabel	0	0	0	0	0	1	0	0	0	0	1	0
35	Corabel	0	0	0	0	0	1	0	0	0	0	1	0
46	Ennis	0	0	0	0	0	0	1	0	0	0	1	0
77	Ennis	0	0	0	0	0	0	1	0	0	0	1	0
59	Ennis	0	1	0	0	0	1	1	0	0	0	1	0
43	Barcelona	0	1	0	0	0	1	1	0	0	0	1	0
52	Barcelona	0	1	0	0	0	0	0	0	0	0	1	0
82	Barcelona	0	1	0	0	0	0	0	0	0	0	0	0
F1	Pauetet	0	0	0	0	0	0	0	0	0	0	0	0
31	Pauetet	0	0	0	0	0	0	0	0	0	0	0	0
33	Pauetet	0	0	0	0	0	0	0	0	0	0	0	0
34	Segorbe	1	0	0	0	0	0	0	0	0	0	0	0
37	Segorbe	0	0	0	0	0	0	0	0	0	0	0	0
45	Segorbe	0	0	0	0	0	0	0	0	0	0	0	0
112	Tombul	1	0	1	0	0	1	0	0	1	1	0	1
111	Tombul	1	0	1	0	0	1	0	0	1	1	0	1
120	Tombul	1	0	1	0	0	1	0	0	1	1	0	1
F3	T. G. Romana	0	1	0	0	0	1	0	0	0	1	0	0
7	T. G. Romana	0	1	0	0	0	1	0	0	0	1	0	1
113	T. G. Romana	0	0	0	0	0	0	0	0	0	0	0	0
F4	Tonda di Giffoni	0	1	0	0	1	1	0	0	1	1	0	1
8	Tonda di Giffoni	0	1	0	0	0	1	0	0	1	1	0	1
22	Tonda di Giffoni	0	1	0	0	1	1	0	0	1	1	0	1

Tabelle 54: Fortsetzung von Tabelle 53.

Nr.	Sorte	[7]	[7]	[7]	[7]	[7]	[7]	[17]	[17]	[17]	[17]	[17]	[17]
		1200	1100	900	700	500	300	3000	2500	2000	1700	1600	1500
S2	Ata Baba	1	1	0	0	0	0	1	0	1	0	1	1
102	Ata Baba	1	1	0	0	0	0	1	0	1	0	1	1
104	Çakildak	1	0	0	0	0	0	1	1	1	0	1	1
103	Çakildak	1	0	0	0	0	0	1	1	1	0	1	1
105	Çakildak	1	0	0	0	0	0	1	1	1	0	1	1
107	Mortarella	1	1	0	0	0	0	1	0	1	0	1	1
106	Mortarella	1	1	0	0	0	0	1	0	1	0	0	1
S1	Mortarella	1	1	0	0	0	0	1	0	1	0	1	1
2	Negret	1	0	0	0	0	0	1	0	1	0	1	1
108	Negret	1	1	0	0	0	0	1	0	1	0	1	1
40	Butler	1	1	0	0	0	0	1	0	1	1	0	1
57	Butler	1	1	0	0	0	0	1	0	1	0	0	1
109	Butler	1	1	1	0	0	1	1	0	0	0	0	0
F2	Corabel	1	1	0	0	0	0	1	0	1	1	0	1
32	Corabel	1	1	0	0	0	0	1	0	0	1	0	1
35	Corabel	1	1	0	0	0	0	1	0	1	0	0	1
46	Ennis	1	1	1	1	1	0	1	0	0	0	0	1
77	Ennis	1	1	1	1	1	0	1	0	0	0	0	0
59	Ennis	1	1	1	1	1	0	1	0	1	0	0	1
43	Barcelona	1	1	0	1	1	0	1	1	0	0	0	1
52	Barcelona	1	1	1	0	0	0	1	1	0	0	0	1
82	Barcelona	1	1	1	0	0	0	1	1	0	0	0	1
F1	Pauetet	1	1	0	0	0	0	1	1	1	0	0	1
31	Pauetet	1	1	0	0	0	0	1	0	1	0	0	1
33	Pauetet	1	1	1	0	0	0	1	1	1	0	0	1
34	Segorbe	1	1	1	1	0	0	1	0	0	0	0	0
37	Segorbe	1	1	1	0	0	0	1	1	1	0	0	1
45	Segorbe	1	1	1	0	0	0	1	0	0	0	0	1
112	Tombul	1	0	0	0	0	0	1	0	1	0	1	1
111	Tombul	1	0	0	0	0	0	1	0	1	0	1	1
120	Tombul	0	0	0	0	0	0	1	0	1	0	1	1
F3	T. G. Romana	0	0	0	0	0	0	1	0	1	0	0	1
7	T. G. Romana	0	0	0	0	0	0	1	0	1	0	0	1
113	T. G. Romana	0	0	0	0	0	0	1	0	0	0	1	0
F4	Tonda di Giffoni	1	0	0	0	0	0	1	0	0	0	0	1
8	Tonda di Giffoni	1	0	0	0	0	0	1	0	0	0	0	1
22	Tonda di Giffoni	0	0	0	0	0	0	1	1	1	0	0	1

Tabelle 55: Fortsetzung von Tabelle 54.

Nr.	Sorte	[17]	[17]	[17]	[17]	[17]	[17]	[17]	[21]	[21]	[21]	[21]	[21]
		1300	1200	1000	900	800	500	400	2000	1750	1600	1500	1300
S2	Ata Baba	1	1	0	1	0	0	0	0	1	1	0	1
102	Ata Baba	0	1	1	0	0	0	0	0	1	1	0	1
104	Çakildak	0	1	1	1	1	0	1	0	1	1	0	1
103	Çakildak	0	1	1	1	1	0	1	0	1	1	0	1
105	Çakildak	0	1	1	1	1	0	1	0	1	1	0	1
107	Mortarella	0	1	0	1	1	0	1	0	1	1	0	1
106	Mortarella	0	1	0	1	1	0	1	0	1	1	0	1
S1	Mortarella	0	1	0	1	1	0	1	0	1	1	0	1
2	Negret	0	1	0	0	0	0	0	0	0	1	0	1
108	Negret	0	1	0	0	0	0	0	0	0	1	0	1
40	Butler	0	1	0	1	1	0	1	0	0	0	0	1
57	Butler	0	1	0	1	1	0	1	0	0	0	0	1
109	Butler	0	0	0	1	1	0	1	0	0	0	0	1
F2	Corabel	0	1	1	1	1	0	1	0	0	1	0	1
32	Corabel	0	1	0	1	1	0	1	0	0	0	0	1
35	Corabel	0	1	1	1	1	0	0	0	0	0	0	1
46	Ennis	0	1	1	0	0	0	0	0	0	0	0	0
77	Ennis	0	1	1	0	0	0	0	0	0	0	0	0
59	Ennis	0	1	1	0	0	0	0	0	0	0	0	1
43	Barcelona	0	1	0	0	1	0	1	0	0	0	0	1
52	Barcelona	0	1	0	0	1	0	1	0	0	0	0	1
82	Barcelona	0	1	0	0	0	0	1	0	0	0	0	1
F1	Pauetet	0	1	0	0	1	0	1	0	0	0	0	1
31	Pauetet	0	1	0	0	1	0	1	0	0	0	0	1
33	Pauetet	0	1	1	0	1	0	1	0	0	0	0	1
34	Segorbe	0	1	1	0	1	1	1	0	0	0	0	1
37	Segorbe	0	1	1	0	1	0	1	0	0	0	0	1
45	Segorbe	0	1	0	0	0	0	1	0	0	0	0	1
112	Tombul	0	1	0	0	0	0	0	0	1	0	1	1
111	Tombul	1	1	0	1	0	0	0	0	1	0	1	1
120	Tombul	0	1	0	0	0	0	0	0	1	0	1	1
F3	T. G. Romana	0	1	1	0	0	0	0	1	0	0	1	1
7	T. G. Romana	0	1	1	1	0	0	0	0	0	0	1	1
113	T. G. Romana	0	1	0	0	0	0	0	0	0	0	1	1
F4	Tonda di Giffoni	0	1	0	0	0	0	0	1	1	1	1	1
8	Tonda di Giffoni	0	1	0	1	0	0	0	1	1	1	1	1
22	Tonda di Giffoni	0	1	0	0	0	0	0	1	1	1	1	1

Tabelle 56: Fortsetzung von Tabelle 55.

Nr.	Sorte	[21]	[21]	[21]	[21]	[21]	[21]	[21]	[21]	[24]	[24]	[24]	[24]
		1200	1000	800	650	550	500	350	250	4500	3500	3000	2500
S2	Ata Baba	1	1	1	1	1	1	0	0	0	0	0	0
102	Ata Baba	1	1	1	1	1	1	0	0	0	0	0	0
104	Çakildak	1	1	1	1	0	0	0	0	1	1	1	1
103	Çakildak	1	1	1	1	1	1	0	0	1	1	0	1
105	Çakildak	1	1	1	1	1	1	0	0	1	1	0	1
107	Mortarella	1	1	1	1	1	1	0	0	0	0	1	0
106	Mortarella	1	1	1	1	1	1	0	0	0	0	1	0
S1	Mortarella	1	1	1	1	1	1	0	0	0	0	1	0
2	Negret	1	1	1	1	1	1	0	1	0	0	0	0
108	Negret	1	1	1	1	1	1	0	0	0	0	0	1
40	Butler	1	1	1	1	1	0	1	1	0	0	0	0
57	Butler	1	1	1	1	1	0	1	1	0	0	0	0
109	Butler	1	1	1	1	1	0	1	1	0	0	0	0
F2	Corabel	1	1	1	1	1	1	0	1	0	0	0	0
32	Corabel	1	1	1	1	1	0	1	1	0	0	0	0
35	Corabel	1	1	1	1	1	0	1	1	0	0	0	0
46	Ennis	1	1	1	1	1	1	1	1	0	0	0	1
77	Ennis	1	1	1	1	1	1	1	1	0	0	0	1
59	Ennis	1	1	1	1	1	1	1	1	0	0	0	1
43	Barcelona	1	1	1	1	1	1	1	1	0	0	0	0
52	Barcelona	1	1	1	1	1	1	1	1	0	0	0	0
82	Barcelona	1	1	1	1	1	1	1	1	0	0	0	1
F1	Pauetet	1	1	1	1	1	1	1	0	0	0	0	0
31	Pauetet	1	1	1	1	1	1	1	1	0	0	0	0
33	Pauetet	1	1	1	1	1	1	1	1	0	0	0	0
34	Segorbe	1	1	1	1	1	1	1	1	0	0	0	1
37	Segorbe	1	1	1	1	1	1	1	1	0	0	0	1
45	Segorbe	1	1	1	1	1	1	1	1	0	0	0	1
112	Tombul	1	1	1	1	1	1	0	0	0	0	0	1
111	Tombul	1	1	1	1	1	1	0	0	0	0	0	1
120	Tombul	1	1	1	1	1	1	0	1	0	0	0	1
F3	T. G. Romana	1	1	1	1	1	1	0	1	0	0	0	1
7	T. G. Romana	1	1	1	1	1	1	0	1	0	0	0	1
113	T. G. Romana	1	1	1	1	1	1	0	1	0	0	0	1
F4	Tonda di Giffoni	1	1	1	0	1	0	0	0	0	0	0	1
8	Tonda di Giffoni	1	1	1	0	1	0	0	0	0	0	0	1
22	Tonda di Giffoni	1	1	1	0	1	0	0	0	0	0	0	1

Tabelle 57: Fortsetzung von Tabelle 56.

Nr.	Sorte	[24]	[24]	[24]	[24]	[24]	[24]	[24]	[24]	[24]	[24]	[24]	[24]
		2000	1900	1800	1700	1500	1400	1100	1000	800	700	600	500
S2	Ata Baba	0	0	1	0	1	0	1	1	1	0	0	1
102	Ata Baba	0	0	1	0	1	0	1	1	1	0	0	1
104	Çakildak	1	0	0	0	0	1	1	1	1	0	0	1
103	Çakildak	1	0	0	0	0	1	1	1	1	1	0	1
105	Çakildak	1	0	0	0	0	1	1	1	1	1	0	1
107	Mortarella	1	0	0	0	0	1	1	1	1	1	0	1
106	Mortarella	1	0	0	0	0	1	1	1	1	1	0	1
S1	Mortarella	1	0	0	0	0	1	1	1	1	1	0	1
2	Negret	0	0	1	0	1	0	1	1	1	0	0	1
108	Negret	0	0	1	0	1	0	1	1	1	0	1	1
40	Butler	0	1	0	0	0	1	1	1	1	0	1	1
57	Butler	0	1	0	0	0	1	1	1	1	0	1	1
109	Butler	0	0	0	0	0	0	0	0	0	0	0	0
F2	Corabel	0	1	0	0	0	1	1	1	1	0	1	1
32	Corabel	0	1	0	0	0	1	1	1	1	0	1	1
35	Corabel	0	1	0	0	0	1	1	0	0	0	0	0
46	Ennis	0	0	1	0	1	0	1	1	1	0	1	1
77	Ennis	0	0	1	0	1	0	1	0	0	0	1	1
59	Ennis	0	0	1	0	1	0	1	0	0	0	1	1
43	Barcelona	0	0	1	0	1	0	1	0	0	0	1	1
52	Barcelona	0	0	1	0	1	0	1	0	0	0	1	1
82	Barcelona	0	0	1	0	1	0	1	0	0	0	1	1
F1	Pauetet	0	0	1	1	0	1	1	1	1	0	1	1
31	Pauetet	0	0	1	1	0	1	1	1	1	0	1	1
33	Pauetet	0	0	1	1	0	1	1	1	1	0	1	1
34	Segorbe	0	0	1	0	0	1	1	1	1	0	1	1
37	Segorbe	0	0	1	0	0	1	1	1	1	0	1	1
45	Segorbe	0	0	1	0	0	1	1	1	1	0	1	1
112	Tombul	0	0	1	0	1	0	1	1	1	0	1	1
111	Tombul	0	0	1	0	1	0	1	1	1	0	1	1
120	Tombul	0	0	1	0	1	0	1	1	1	0	1	1
F3	T. G. Romana	0	0	1	0	0	0	1	1	1	0	1	1
7	T. G. Romana	0	0	1	0	1	0	1	1	1	0	1	1
113	T. G. Romana	0	0	1	0	1	0	1	1	1	0	1	1
F4	Tonda di Giffoni	0	0	1	0	1	0	1	1	1	0	1	1
8	Tonda di Giffoni	0	0	1	0	1	0	1	1	1	0	1	1
22	Tonda di Giffoni	0	0	1	0	1	0	1	1	1	0	1	1

Tabelle 58: Fortsetzung von Tabelle 57.

Nr.	Sorte	[35]	[35]	[35]	[35]	[35]	[35]	[35]	[35]	[35]	[36]	[36]	[36]
		2000	1600	1200	1100	900	800	700	550	400	3500	3000	2500
S2	Ata Baba	0	1	1	1	1	1	1	1	0	0	0	0
102	Ata Baba	0	1	1	1	1	1	1	1	0	0	0	1
104	Çakildak	1	1	1	1	1	1	1	1	0	1	1	1
103	Çakildak	1	1	1	1	1	1	1	1	0	1	0	1
105	Çakildak	1	1	1	1	1	1	1	1	0	1	0	1
107	Mortarella	1	1	1	1	1	1	1	1	0	1	0	1
106	Mortarella	1	1	1	1	1	1	1	1	0	1	0	1
S1	Mortarella	1	1	1	1	1	1	1	1	0	1	0	1
2	Negret	0	1	1	1	1	1	1	1	0	0	0	0
108	Negret	1	1	1	1	1	1	1	1	0	0	0	0
40	Butler	0	1	1	1	1	1	1	1	0	0	0	0
57	Butler	0	1	1	1	1	1	1	1	0	0	0	0
109	Butler	0	0	1	1	1	1	1	1	0	0	0	0
F2	Corabel	1	1	1	1	1	1	1	1	0	0	0	0
32	Corabel	1	1	1	1	1	1	1	1	0	0	0	0
35	Corabel	1	1	1	1	1	1	1	1	0	0	0	0
46	Ennis	0	1	1	1	1	1	1	1	0	0	0	1
77	Ennis	0	1	1	1	1	1	1	1	0	0	0	1
59	Ennis	0	1	1	1	1	1	1	1	0	0	0	1
43	Barcelona	0	1	1	1	1	1	1	1	1	0	0	1
52	Barcelona	0	1	1	1	1	1	1	1	0	0	0	1
82	Barcelona	0	1	1	1	1	1	1	1	0	0	1	1
F1	Pauetet	0	1	1	1	1	1	1	1	0	0	0	0
31	Pauetet	0	1	1	1	1	1	1	1	0	0	0	0
33	Pauetet	0	1	1	1	1	1	1	1	0	0	0	0
34	Segorbe	0	1	1	1	1	1	1	1	0	0	0	0
37	Segorbe	0	1	1	1	1	1	1	1	0	0	0	0
45	Segorbe	0	1	1	1	1	1	1	1	0	0	0	0
112	Tombul	1	1	1	1	1	1	1	1	0	0	0	0
111	Tombul	1	1	1	1	1	1	1	1	0	0	0	0
120	Tombul	1	1	1	1	1	1	1	1	0	0	0	0
F3	T. G. Romana	0	1	1	1	1	1	1	1	0	0	0	0
7	T. G. Romana	0	1	1	1	1	1	1	1	0	0	0	0
113	T. G. Romana	0	1	1	1	1	1	1	1	0	0	0	0
F4	Tonda di Giffoni	0	1	1	1	1	1	1	1	0	0	0	0
8	Tonda di Giffoni	0	1	1	1	1	1	1	1	0	0	0	0
22	Tonda di Giffoni	1	1	1	1	1	1	1	1	0	0	0	0

Tabelle 59: Fortsetzung von Tabelle 58.

Nr.	Sorte	[36]	[36]	[36]	[36]	[36]	[36]	[36]	[36]	[36]	[36]	[36]
		2200	2000	1900	1700	1500	1200	1100	1000	800	700	500
S2	Ata Baba	0	1	1	0	1	0	1	1	0	0	1
102	Ata Baba	0	1	1	0	1	0	1	1	0	0	1
104	Çakildak	0	1	1	1	1	0	1	0	0	0	1
103	Çakildak	0	1	1	1	1	0	1	1	0	1	1
105	Çakildak	0	1	1	1	1	0	1	1	0	1	1
107	Mortarella	0	1	1	0	1	0	1	1	0	1	1
106	Mortarella	0	1	1	0	1	0	1	1	0	1	1
S1	Mortarella	0	1	1	0	1	0	1	1	0	1	1
2	Negret	0	1	1	0	1	0	1	1	0	1	1
108	Negret	0	1	1	0	1	0	1	1	0	1	1
40	Butler	1	1	0	1	1	0	1	1	0	1	1
57	Butler	1	1	0	1	1	0	1	1	0	1	1
109	Butler	1	1	0	1	1	0	1	1	0	1	1
F2	Corabel	0	0	0	0	0	0	0	0	0	0	0
32	Corabel	0	0	0	0	0	0	0	0	0	0	0
35	Corabel	0	0	0	0	0	0	0	0	0	0	0
46	Ennis	0	1	0	0	1	0	1	0	0	1	1
77	Ennis	0	1	0	0	1	0	1	0	0	0	1
59	Ennis	0	1	0	0	1	0	1	0	0	1	1
43	Barcelona	0	1	0	1	1	0	1	1	0	0	1
52	Barcelona	0	1	0	1	1	0	1	1	0	0	1
82	Barcelona	0	1	0	1	1	0	1	1	0	0	1
F1	Pauetet	0	1	1	1	1	0	1	1	0	1	1
31	Pauetet	0	1	1	1	1	0	1	1	0	1	1
33	Pauetet	0	1	1	1	1	0	1	1	0	1	1
34	Segorbe	0	1	1	1	1	0	1	1	1	1	1
37	Segorbe	0	1	1	1	1	0	1	1	0	1	1
45	Segorbe	0	1	1	1	1	0	1	1	0	1	1
112	Tombul	0	1	1	1	1	0	1	1	0	1	1
111	Tombul	0	1	1	1	1	0	1	1	0	1	1
120	Tombul	0	1	1	1	1	0	1	1	0	1	1
F3	T. G. Romana	1	1	0	0	1	1	0	0	0	1	1
7	T. G. Romana	1	1	0	0	1	1	0	0	0	1	1
113	T. G. Romana	1	1	0	0	1	1	0	0	0	1	1
F4	Tonda di Giffoni	1	1	0	1	1	1	0	0	0	1	1
8	Tonda di Giffoni	1	1	0	0	1	1	0	0	0	1	1
22	Tonda di Giffoni	1	1	0	0	1	1	0	0	0	1	1

9.10.10 Entwicklung sortenspezifischer SCAR-Marker

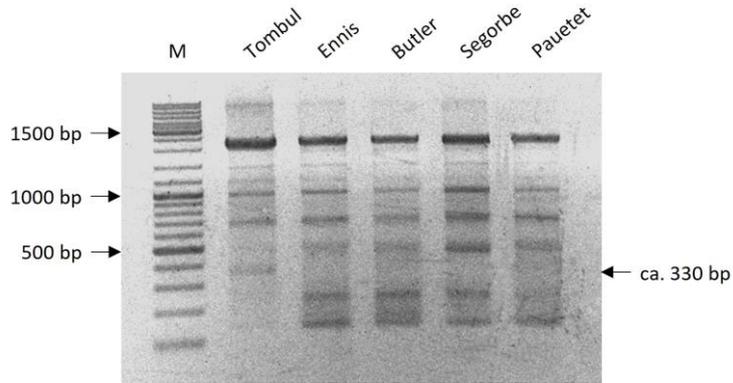


Abbildung 75: Gelbild der RAPD-PCR (Primer 17) für die Überprüfung der Bande bei 330 bp bei der Sorte Tombul.

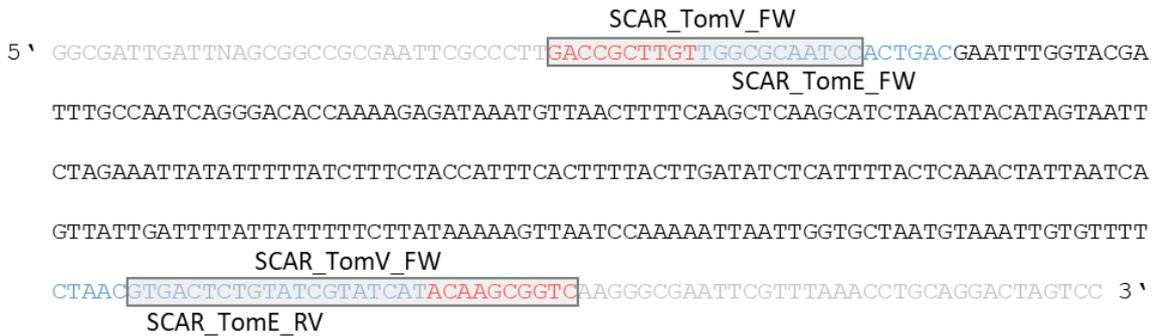


Abbildung 76: Sequenz der Kolonie 16 (Sorte: Tombul) inklusive der basierend darauf entwickelten Primer für den sortenspezifischen Nachweis. Rot: RAPD-Primer 17, Blau: eingeschobene Primer SCAR_TomE_FW/RV, Grau hinterlegt: Primerpaar SCAR_TomV_FW/RV, das auf einer Verlängerung der RAPD-Primer beruht.

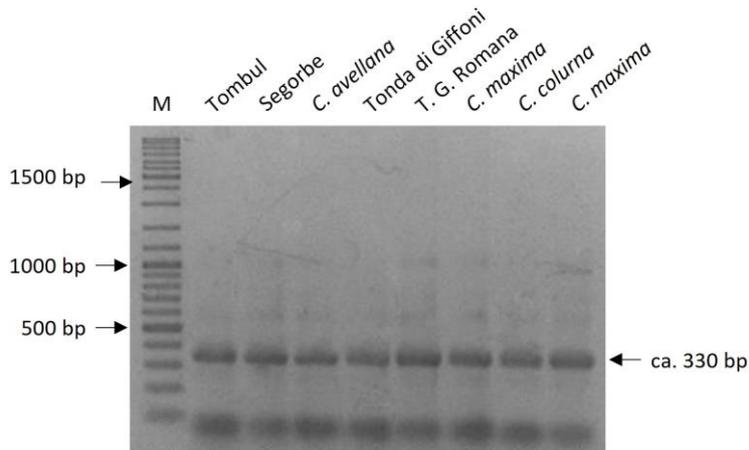


Abbildung 77: Gelbild der PCR zur Überprüfung des entwickelten Primerpaares SCAR-TomV_FW/RV.

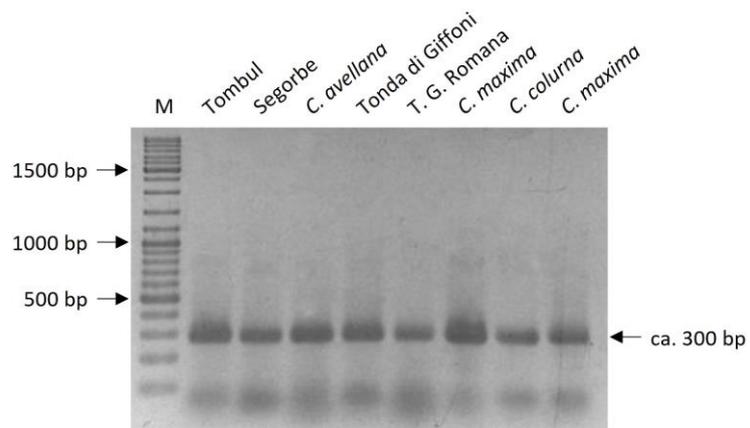


Abbildung 78: Gelbild der PCR zur Überprüfung des entwickelten Primerpaares SCAR-TomE_FW/RV.

10 Danksagung

Zu guter Letzt möchte ich noch die Gelegenheit nutzen allen Menschen in meinem Umfeld zu danken, die diese Arbeit erst ermöglicht und zu ihrem Gelingen beigetragen haben.

Ich bedanke mich bei meinem Doktorvater Prof. Dr. Fischer für die Überlassung dieses Themas und der Möglichkeit, von extern kommend Teil seiner Arbeitsgruppe zu werden und mich in ein für mich neues Forschungsgebiet einarbeiten zu können. Zudem möchte ich mich für das mir entgegengebrachte Vertrauen und die gewährten Freiheiten sowie die übertragene Verantwortung bedanken.

Ich bedanke mich bei meinen Diplomandinnen Charline Köster, Lisa Paul, Sarah Obermeyer, Lisa Schramm und Maya Möller, die mit ihren Ergebnissen zu dieser Arbeit beigetragen haben. Ich danke euch für die viele gute Laune, die ihr ins Labor gebracht habt und die freundschaftlichen Verhältnisse, die daraus entstanden sind.

Ich möchte mich auch bei allen Kooperationspartnern und Firmen bedanken, die mich und das von mir betreute FEI-Projekt im Laufe der Arbeit unterstützt haben.

Ich bedanke mich bei meiner Arbeitsgruppe für die vielen lustigen Momente, wie z.B. die Mittagspausen bei einer Runde Bäng, die dazu beigetragen haben, dass die Promotion eine besondere Zeit war. Ganz besonders danke ich den Kellerkindern. Ich finde wir haben es uns so weit abseits des Rests unserer Arbeitsgruppe wirklich schön gemacht, v.a. mit unseren flauschigen nussbegeisterten Mitbewohnern ;-). Mein besonderer Dank gilt hierbei Stefanie Schelm für die vielen gemeinsamen Momente und wissenschaftlichen Diskussionen und den großartigen Input. Das alles hat meine Promotionszeit sehr geprägt. Ich freue mich auch sehr, dass ihr, Anne und Nils, Teil der Kellerkinder wurdet. Ihr habt die Arbeitsgruppe sehr bereichert. Und mit dir, Nils, hatte ich endlich einen Kaffee-Buddy dort unten.

Ganz besonders möchte ich mich auch bei meiner Familie bedanken ohne die mir diese Arbeit nie so gut gelungen wäre und ohne die ich heute nicht diese Worte schreiben würde. Mama und Papa, ihr habt mir nicht nur finanziell ermöglicht meinen Lebensweg bis hierher so gehen zu können, sondern habt mir auch mit eurer Liebe und Unterstützung immer den Rücken gestärkt und seid mit mir durch alle Höhen und Tiefen gegangen und habt mich auch aus den Momenten des Zweifels befreit und mir immer neuen Mut gemacht. Dafür bin ich euch unendlich dankbar!

Und zuletzt danke ich dir René. Du bist das Beste, was mir in dieser Zeit passiert ist! Du hast mein Leben unendlich bereichert und ich bin dir für deine unerschütterliche Liebe und Unterstützung so dankbar. Du hast mir in der gesamten Zeit und in den vielen Monaten des Schreibens so unendlich viel Kraft gegeben. Dafür danke ich dir von ganzem Herzen! Ich liebe dich!

11 Eidesstattliche Versicherung

„Hiermit versichere ich an Eides statt, die vorliegende Dissertation selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt zu haben. Das gebundene Exemplar der Dissertation entspricht dem in elektronischer Form eingereichten Dissertationsexemplar. Ich versichere, dass diese Dissertation nicht in einem früheren Promotionsverfahren eingereicht wurde.“

Hamburg, den 22.07.2023



Christina Lang