

**Genomweite Sequenzierung des Fledermaus-
Malariaparasiten *Polychromophilus murinus* und
bioinformatische Analyse von Parasit-Wirt-
Interaktions-Proteinen**

DISSERTATION

zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)

am Fachbereichs Biologie, der Fakultät für Mathematik,
Informatik und Naturwissenschaften der Universität Hamburg

vorgelegt von

ELSA AUL
(geb. Ziegler)

aus Hamburg

Hamburg im März 2020

“According to Darwin’s Origin of Species, it is not the most intellectual of the species that survives; it is not the strongest that survives; but the species that survives is the one that is able best to adapt and adjust to the changing environment in which it finds itself”

Leon C. Megginson, 1963

Tag der Disputation: 03. Juli 2020

Gutachter der Dissertation:

1. Gutachterin: Prof. Dr. rer. nat. Iris Bruchhaus
Bernhard-Nocht-Institut für Tropenmedizin
Abteilung Zelluläre Parasitologie
Bernhard-Nocht-Straße 74, 20359 Hamburg

2. Gutachterin: Prof. Dr. rer. nat. Susanne Dobler
Universität Hamburg, Institut für Zoologie
Abteilung Molekulare Evolutionsbiologie
Martin-Luther-King-Platz 3, 20146 Hamburg

Inhalt

Abbildungsverzeichnis	7
Tabellenverzeichnis.....	9
Zusammenfassung	11
1 Einleitung.....	12
1.1 Das Phylum Apikomplexa	12
1.1.1 Malaria und malariaähnliche Parasiten.....	14
1.1.2 Phylogenetische Beziehungen der Haemosporida	15
1.2 <i>Polychromophilus</i> als Modelorganismus	17
1.2.1 Biologie und Ökologie der Wirte.....	19
1.2.2 Epidemiologische Aspekte	22
1.2.3 Morphologie und Lebenszyklus.....	22
1.3 Parasit-Wirt-Interaktionen auf Proteinebene	27
1.3.1 Export putativ wirtszell-interagierender Proteine.....	27
1.3.2 Genomanalyse zur Identifizierung relevanter Proteine und damit verbundene Hürden.....	28
1.4 Zielsetzung	30
2 Material.....	31
2.1 Arbeitsgeräte	31
2.2 Gebrauchsmaterialien	32
2.3 Chemikalien.....	32
2.4 Lösungen und Puffer.....	33
2.5 Molekularbiologische Kits	34
2.6 Enzyme und andere biologische Substanzen	34
2.7 Oligonukleotide	35
2.8 Datenbanken und Programme	37
3 Methoden.....	43
3.1 Gewinnung von Untersuchungsmaterial	43
3.2 Anfertigung von Blutaussstrichen und Bestimmung der Parasitämie.....	43
3.3 Anreicherung von Oozysten, Mikrogameten und Gametozyten	44
3.3.1 Mitteldarmpräparation der Fledermausfliegen	44
3.3.2 <i>In vitro</i> Exflagellation.....	44
3.3.3 Dichtegradientenzentrifugation in Ficoll.....	44
3.4 Isolierung genomischer DNA	45
3.4.1 Isolierung genomischer DNA aus Fledermausfliegen.....	45

3.4.2	Isolierung genomischer DNA aus Oozysten und Mikrogameten	45
3.4.3	Isolierung genomischer DNA aus Blut.....	45
3.5	Amplifizierung der genomischen DNA.....	46
3.6	Selektive Amplifizierung der genomischen DNA	47
3.6.1	Komplementärstrangsynthese.....	47
3.7	PCR-basierte Methoden	48
3.7.1	Nested-PCR.....	49
3.7.2	Stepdown-PCR.....	49
3.7.3	Banden-Reamplifizierung	49
3.7.4	Kolonie-PCR	49
3.7.5	Quantitative Real-Time-PCR	50
3.8	Analyse und Aufreinigung von Nukleinsäuren.....	52
3.8.1	Konzentrationsbestimmung von Nukleinsäuren	52
3.8.2	Elektrophoretische Auftrennung von Nukleinsäuren	52
3.8.3	Extraktion von DNA aus Agarosegelen.....	53
3.8.4	Enzymatische Aufreinigung der PCR-Produkte	53
3.8.5	Phenol/Chloroform-Extraktion	53
3.8.6	Ethanol- und Isopropanol-Präzipitation	54
3.9	Klonierungsmethoden.....	54
3.9.1	Ligation.....	54
3.9.2	Transformation und Selektion	55
3.9.3	Plasmidpräparation	55
3.10	Sequenzierung.....	55
3.10.1	Sanger-Sequenzierung	55
3.10.2	Paired-End-Sequenzierung.....	56
3.10.3	Mate-Pair-Sequenzierung.....	56
3.11	Allgemeine bioinformatische Methoden.....	56
3.11.1	DNA-Sequenzanalyse.....	56
3.11.2	BLAST.....	57
3.11.3	Primerdesign	57
3.11.4	Herstellung der SWGA-Primer.....	58
3.12	Allgemeine Methoden mit NGS-Daten	58
3.12.1	Bearbeitung von NGS-Daten.....	58
3.12.2	Aufbereitung der Rohdaten.....	59
3.12.3	<i>De novo</i> Assembly.....	59
3.12.4	Mapping.....	59
3.12.5	Mapping als Sortierprozess.....	60
3.13	Beurteilung des Anteils an Kontaminationen im Assembly	61
3.14	Auftrennung der wirts- und parasitenspezifischen Reads	61
3.14.1	Readextraktion über verwandte Referenzgenome und GC-Gehalt	61
3.14.2	Readextraktion basierend auf genomischen Sequenzen der Wirtsfledermaus	62

3.15	Referenzbasiertes Assembly	64
3.15.1	Herstellung der genomischen Sequenzen von <i>P. murinus</i>	64
3.15.2	Herstellung der genomischen Sequenzen von <i>M. daubentonii</i>	64
3.16	Annotation	65
3.16.1	Annotation des Genoms von <i>P. murinus</i>	65
3.16.2	Annotation der mitochondrialen und apikoplastiden DNA von <i>P. murinus</i>	66
3.16.3	Annotation des Genoms von <i>M. daubentonii</i>	66
3.16.4	Annotation der mitochondrialen DNA <i>M. daubentonii</i>	67
3.17	Evaluierung der Annotation	67
3.18	Funktionelle Annotation	67
3.18.1	Vorhersage von Domänen und Signalsequenzen.....	68
3.18.2	Gene Ontology (GO) Annotation.....	68
3.19	Identifizierung orthologer Proteine	68
3.20	Phylogenetische Analyse der Sequenzdaten	69
4	Ergebnisse	71
4.1	Vorexperimente zur Anreicherung des Parasiten.....	71
4.1.1	Prävalenz und Parasitenintensität von <i>P. murinus</i>	71
4.1.2	Qualitätsbeurteilung der WGA-DNA.....	73
4.1.3	Quantifizierung des Anteils der Parasiten-DNA nach WGA	74
4.2	Sequenzierung der WGA-DNA mittels Illumina	75
4.2.1	<i>De novo</i> Assembly und Identifizierung parasitenspezifischer Contigs	75
4.2.2	Extraktion der Sequenzen parasitären Ursprungs.....	78
4.2.3	Phylogenetische Position von <i>P. murinus</i> als Schwesterngruppe zu <i>Plasmodium</i>	80
4.3	Sequenzierung des Wirtsgenoms.....	81
4.3.1	Die mitochondriale DNA von <i>M. daubentonii</i>	84
4.3.2	Phylogenetische Position von <i>M. daubentonii</i>	85
4.3.3	Das Genom der Wirtsfledermaus.....	87
4.4	Genomanalyse von <i>Polychromophilus murinus</i>	89
4.4.1	Selektive Anreicherung und Genomresequenzierung	89
4.4.2	Das mitochondriale Genom von <i>P. murinus</i> und darauf basierende phylogenetische Analysen	92
4.4.3	Konvergente Reduktion des Apikoplasten innerhalb der Haemosporida.....	94
4.4.4	Das Kern-Genom von <i>P. murinus</i>	98
4.5	Identifikation wirtzell-interagierender Proteine	104
4.5.1	Konservierte Proteine der Aconoidasida.....	105
4.5.2	Konservierte Proteine der Haemosporida	107
4.5.3	Konservierte Proteine der Plasmodiidae	109
4.5.4	Plastizität des Exportsignals	110
4.5.5	Putative Protein-Protein-Interaktionen zwischen Parasit und Wirt.....	112
5	Diskussion	113

5.1	<i>Polychromophilus murinus</i> – Ein herausforderndes Genom-Sequenzierungsprojekt	113
5.1.1	Das unzertrennliche Parasiten-Wirt-System	114
5.1.2	Facetten eines Metagenoms: Ko-Infektionen und andere Kontaminationen.....	116
5.1.3	Die zielgerichtete Anreicherung des Parasiten-Genoms.....	118
5.2	Evolution charakteristischer Merkmale im Genom der Haemosporiden.....	121
5.2.1	Genomreduktion und GC-Gehalt als Zeichen der Evolution.....	121
5.2.2	Das konservierte Genrepertoire der Haemosporiden	123
5.2.3	Proteine mit Relevanz für die Parasit-Wirt-Interaktion	125
5.3	Ausblick: Parasiten prägen das Genom ihres Wirts — und umgekehrt.....	128
6	Literaturverzeichnis.....	130
7	Anhang	149
7.1	Abkürzungsverzeichnis.....	149
7.2	Ergänzende Tabellen	152
7.3	Ergänzende Abbildungen.....	165
	Eidesstattliche Versicherung	171
	Danke	172

Abbildungsverzeichnis

Abb. 1.1: Schematischer Aufbau (a) und phylogenetische Stellung (b) der Apikomplexa.....	13
Abb. 1.2: Phylogenetische Beziehungen innerhalb der Haemosporida.....	17
Abb. 1.3: End- und Zwischenwirt von <i>P. murinus</i>	19
Abb. 1.4: Schematische Darstellung des Entwicklungszyklus von <i>P. murinus</i>	23
Abb. 1.5: Mikroskopische Aufnahmen unterschiedlicher Stadien in der Fledermausfliege.	24
Abb. 1.6: Mikroskopische Aufnahmen unterschiedlicher Stadien in der Fledermaus.....	26
Abb. 3.1: Schematische Darstellung der Whole Genome Amplifizierung.....	46
Abb. 3.2: Schematische Darstellung von Positionierungsmöglichkeiten beim Mapping.....	60
Abb. 3.3: Übersichtsschema zur Identifikation der Paired-End-Reads von <i>P. murinus</i>	62
Abb. 3.4: Herstellung von Contigs parasitären Ursprungs auf hohem Zuverlässigkeitsniveau.	63
Abb. 4.1: Experimentelle Strategie zur Sequenzierung des Genoms von <i>P. murinus</i>	71
Abb. 4.2: Nachweis der parasitären DNA in den Wasserfledermäusen.....	72
Abb. 4.3: Repräsentative lichtmikroskopische Aufnahmen infizierter Erythrozyten.....	72
Abb. 4.4: Whole-Genome-Amplifikation (WGA) des Untersuchungsmaterials.....	73
Abb. 4.6: Analyse der DNA-Zusammensetzung des Untersuchungsmaterials.....	74
Abb. 4.7: Verteilung des GC-Gehalts im Rohdaten-Assembly.....	76
Abb. 4.8: Grad der „Kontamination“ im Assembly der Paired-End-Sequenzierung.....	77
Abb. 4.9: Taxonspezifische Einordnung der Rohdatensequenzen (GC-Gehalt < 30 %).	78
Abb. 4.10: Verteilung identifizierter Gene nach unterschiedlichen Readextraktionsmethoden. ...	79
Abb. 4.11: Gemeinsame Gene nach <i>Myotis</i> - (a), Apikomplexa- (b), und GC-Filter (c).....	80
Abb. 4.12: Phylogenetische Beziehungen der Haemosporida basierend auf Genomdaten.....	80
Abb. 4.13: Grad der Kontamination bei der Sequenzierung der nicht infizierten Fledermaus.....	83
Abb. 4.15: Organisation der mtDNA von <i>M. daubentonii</i>	84
Abb. 4.16: Phylogenie von Fledermäusen der Gattung <i>Myotis</i>	86
Abb. 4.17: Verteilung repetitiver DNA im Genom von <i>M. daubentonii</i>	89
Abb. 4.19: Abdeckung der Sequenzen von <i>P. murinus</i> im Überblick.....	90
Abb. 4.20: Verteilung des GC-Gehalts im Mate-Pair-Assembly.....	91
Abb. 4.21: Taxonspezifische Einordnung des Mate-Pair-Rohdatenassemblys.....	92
Abb. 4.22: Organisation der mtDNA von <i>P. murinus</i>	93
Abb. 4.23: Phylogenetische Stellung von <i>P. murinus</i> basierend auf Daten der mtDNA.....	94
Abb. 4.24: rRNA/tRNA-Gencluster des Apikoplast-Genoms von <i>P. murinus</i>	95

Abb. 4.25: Organisation des Apikoplast-Genoms von <i>P. murinus</i> .	97
Abb. 4.26: Rearrangement des Apikoplasmins im phylogenetischen Kontext.	98
Abb. 4.27: Durchschnittliche Abdeckung der genomischen Sequenzen von <i>P. murinus</i> .	99
Abb. 4.28: Bewertung der Vollständigkeit des Genoms von <i>P. murinus</i> im Vergleich.	100
Abb. 4.29: Verwendung der Aminosäuren im Vergleich.	101
Abb. 4.30: Analyse der putativen Lokalisation annotierter Proteine von <i>P. murinus</i> .	103
Abb. 4.31: Homologe Proteine innerhalb der Haemosporida.	104
Abb. 4.32: Einteilung der konservierten Proteine der Aconoidasida in Proteinklassen.	105
Abb. 4.33: Molekulare Funktion membranassoziierter konservierter Proteine der Haemosporida.	108
Abb. 4.34: Plastizität und Evolution des PEXEL innerhalb der Familie der Plasmodiidae.	111
Abb. 4.35: Protein-Protein-Netzwerk zwischen parasitären, exportierten Proteinen und dem Wirt.	112
Abb. 5.1: Expression der 6-cys Proteine im Lebenszyklus von <i>Plasmodium</i> .	127
Abb. 7.1: ML-Phylogenie der Haemosporida basierend auf genomischen Sequenzen.	165
Abb. 7.2: Phylogenie der Gattung <i>Myotis</i> basierend auf BI-Analysen.	165
Abb. 7.3: Phylogenie der Gattung <i>Myotis</i> basierend auf ML-Analysen.	166
Abb. 7.4: Phylogenie der Gattung <i>Myotis</i> basierend auf MP-Analysen.	166
Abb. 7.5: Phylogenie der Haemosporida basierend auf BI-Analysen der mtDNA.	167
Abb. 7.6: Phylogenie der Haemosporida basierend auf ML-Analysen der mtDNA.	167
Abb. 7.7: Phylogenie der Haemosporida basierend auf MP-Analysen der mtDNA.	168
Abb. 7.8: BI-Analyse mitochondrialer Proteine der Haemosporida.	168
Abb. 7.9: MP-Analyse mitochondrialer Proteine der Haemosporida.	169
Abb. 7.11: Phylogenetische Stellung von <i>P. murinus</i> basierend auf Daten der mtDNA.	169
Abb. 7.12: Vergleich der repetitiven Region der Apikoplasten von <i>P. murinus</i> und <i>P. chabaudi</i> .	170
Abb. 7.13: Konservierte Syntänie-Blöcke von <i>P. murinus</i> im Vergleich.	170

Tabellenverzeichnis

Tab. 2.1: Verwendete Primer zur Amplifizierung und Sequenzierung.	35
Tab. 2.2: Vektorspezifische Primer für Kolonie-PCR und Sequenzierung.	36
Tab. 2.3: Primer zur Quantifizierung der relativen Anteile von parasitärer und Wirt-DNA.....	36
Tab. 2.4: Primerzusammensetzung zur selektiven Anreicherung der parasitären DNA.....	37
Tab. 3.1: Zusammensetzung und Temperaturprofil der SWGA-Reaktion.	47
Tab. 3.2: Zusammensetzung und Temperaturprofil der Komplementärstrangsynthese.	47
Tab. 3.3: Zusammensetzung und Temperaturprofil der Standard-PCR.....	48
Tab. 3.4: Zusammensetzung und Temperaturprofil der Kolonie-PCR.....	50
Tab. 3.5: Zusammensetzung und Temperaturprofil der quantitativen Real-Time PCR.	50
Tab. 3.6: Eigenschaften der Paired-End-Sequenzierung.....	56
Tab. 3.7: Eigenschaften der Mate-Pair-Sequenzierung.....	56
Tab. 3.8: Übersicht der lokalen BLAST-Datenbanken.	57
Tab. 3.9: Übersicht der Datensätze für die phylogenetischen Analysen.....	69
Tab. 3.10: Verwendete Parameter der phylogenetischen Analysen.....	70
Tab. 4.1: Parasitämie der infizierten Fledermäuse.....	72
Tab. 4.2: Zusammenfassende Statistik über Roh- und gefilterte Sequenzierungsdaten.....	75
Tab. 4.3: Statistiken des Rohdaten-Assemblys.	75
Tab. 4.4: Vergleich der Readextraktionsmethoden.....	78
Tab. 4.5: Statistik der Sequenzierungs- und Mappingdaten von <i>M. daubentonii</i>	81
Tab. 4.6: Statistiken zur Herstellung des ersten Genom-Entwurfs von <i>M. daubentonii</i>	87
Tab. 4.7: Statistiken zum Entwurf des Genoms von <i>M. daubentonii</i> im Vergleich.....	88
Tab. 4.8: Zusammenfassende Statistik der Mate-pair-Sequenzierung.....	90
Tab. 4.9: Statistiken des Mate-Pair-Assemblys.	91
Tab. 4.10: Statistiken zur genomischen Sequenz von <i>P. murinus</i> im Vergleich.	99
Tab. 4.11: Genom-Eigenschaften von <i>P. murinus</i> im Vergleich zu zwei anderen Haemosporiden.	101
Tab. 4.12: Der GC-Gehalt der Kodonpositionen.....	101
Tab. 4.13: Putative Invasions- und Transmigrationsabhängige Proteine.	109
Tab. 7.1: Iupac-Abkürzungskode für Nukleotide.	152
Tab. 7.2: Iupac-Abkürzungskodes für Aminosäuren.	152
Tab. 7.3: Übersicht der gefangenen Fledermäuse.....	153

Tab. 7.4: GC-Gehalt genomischer Loci von <i>Polychromophilus</i> spp.	154
Tab. 7.5: Verwendete Genomdaten zur Rekonstruktion der Phylogenie der Haemosporida.	154
Tab. 7.6: Verwendete OrthoMCL IDs zur Rekonstruktion der Phylogenie der Haemosporida.	155
Tab. 7.7: Annotation der mitochondrialen DNA von <i>Myotis daubentonii</i>	156
Tab. 7.8: MtDNA Sequenzdaten zur Rekonstruktion der Phylogenie der Gattung <i>Myotis</i>	156
Tab. 7.9: Anteil repetitiver DNA im Genom von <i>M. daubentonii</i>	157
Tab. 7.10: Vergleich der transponierbaren Elemente im Säugetiergenom.	157
Tab. 7.11: Kodonverwendung in der mtDNA von <i>P. murinus</i>	157
Tab. 7.12: Eigenschaften der Primersets zur selektiven Amplifizierung von <i>P. murinus</i>	158
Tab. 7.13: MtDNA Sequenzdaten zur Rekonstruktion der Phylogenie der Apikomplexa.	158
Tab. 7.14: Annotation der Apikoplast-DNA von <i>Polychromobilus murinus</i>	159
Tab. 7.15: Kodonverwendung im Apikoplast-Genom von <i>P. murinus</i>	160
Tab. 7.16: PIDNA Sequenzdaten zur Rekonstruktion der Phylogenie der Apikomplexa.	160
Tab. 7.17: Kodonverwendung in der genomischen DNA von <i>P. murinus</i>	161
Tab. 7.18: Exportierte Proteine von <i>Polychromophilus murinus</i>	161
Tab. 7.19: Konservierte Proteine der Plasmodiidae, die nicht in fledermausinfizierenden Gattungen enthalten sind	163

Zusammenfassung

Eine parasitische Lebensweise ist durch eine stetige wechselseitige Interaktion zwischen Parasit und Wirt gekennzeichnet. Haemosporida sind extrem gut an die Interaktionen mit ihrem Wirt und dessen Immunsystem angepasst. Sie sind in der Lage intrazellulär zu leben, das Immunsystem zu umgehen und teilweise zu unterdrücken, exponierte Antigene schnell zu verändern oder sogar einen Verlust des immunologischen Gedächtnisses zu induzieren. Als Erreger der Malaria spielen humanpathogene Haemosporida der Gattung *Plasmodium* eine besondere medizinische Bedeutung. Neben diesen Plasmodien gibt es jedoch auch eine Reihe weiterer, teils weniger gut untersuchte, Parasiten, die auch in anderen Wirbeltieren malariaähnliche Krankheiten hervorrufen und für erhebliche Dezimierungen in Wildtierpopulationen verantwortlich sind. *Polychromophilus* spp. sind Malariaparasiten, die sich auf Fledermäuse als Wirt spezialisiert haben und von ektoparasitischen Fledermausfliegen der Familie Nycteribiidae übertragen werden. Der Lebenszyklus von *Polychromophilus* spp. gleicht in vielen Aspekten denen der nahe verwandten Plasmodienarten. Allerdings dringen die Vertreter der Gattung *Polychromophilus* nur für die Gametozytenentwicklung in Erythrozyten ein und weisen keine asexuelle Vermehrung in den Blutzellen auf. Damit fehlt den Parasiten genau der Teil des Lebenszyklus, der mit den typischen Symptomen diverser Malariaerreger in Verbindung steht. Die biologischen Unterschiede und Gemeinsamkeiten zwischen den fledermausinfizierenden Parasiten und *Plasmodium* machen Vertreter der Gattung *Polychromophilus* zu hervorragenden Organismen, um das Verständnis wichtiger Aspekte der Malariaparasiten zu vertiefen. Im Zuge dieser Arbeit wurde daher der Fokus auf die Genomsequenzierung des Erregers *Polychromophilus murinus* gelegt, der die heimische Wasserfledermaus (*Myotis daubentonii*) infiziert. Es wurden Methoden etabliert, um die genomische DNA des Malariaparasiten in ausreichender Menge verfügbar zu machen und einen ersten Genomentwurf aus dem Blut einer natürlich infizierten Wasserfledermaus zu erstellen. Die durchgeführten Analysen unterstützen die paraphyletische Abstammung der Gattung *Plasmodium* in Hinblick auf fledermausinfizierende Malariaerreger und zeigen, dass *P. murinus* anderen Malariaparasiten in Genanordnung und -inhalt ähnelt. Die Genome zeigen dabei einen hohen Grad der Erhaltung von Genorganisation und -komplexität. Proteine, die eine besondere Rolle bei der Interaktion mit dem Wirt spielen sind auf der Ebene der Haemosporida weitestgehend konserviert. Gleichzeitig zeigen fledermausinfizierende Parasiten einen konsekutiven Verlust von Proteinen, die an der Invasion und Replikation der Erythrozyten beteiligt sind und somit potenziell mit der Pathogenität anderer Malariaerreger in Verbindung stehen. Die Fertigstellung des Genoms von *P. murinus* stellt eine wertvolle Ressource zur Verfügung, um die Untersuchung dieser Spezies voranzutreiben und gleichzeitig wesentliche Aspekte im Lebenszyklus verwandter Parasiten weiter aufzuklären.

1 Einleitung

Parasitismus ist eine in der Natur allgegenwärtige Lebensform, die eine zentrale Rolle in sämtlichen Ökosystemen spielt. Definitionsgemäß handelt es sich bei Parasiten um Organismen, die in einem oder mehreren Entwicklungsstadien physiologisch von einem andersartigen Organismus, dem Wirt, abhängig sind (Castro et al., 1996). Zur Unterscheidung von anderen Lebensformen ist der Parasitismus zusätzlich dadurch charakterisiert, dass Wirte durch pathogene Eigenschaften der Parasiten erheblich beeinträchtigt, bei geringer Parasitenlast (Befallsintensität) allerdings nicht getötet werden (Crofton, 1971). Parasit-Wirt-Beziehungen werden so auf eine antagonistische Wechselwirkung zwischen den beteiligten Organismen reduziert. Tatsächlich ist eine deutliche Abgrenzung von anderen Lebensformen mit einseitig vorteilhaften nichtschädlichen (Kommensalismus) oder gegenseitig fakultativ vorteilhaften (Mutualismus) Wechselbeziehungen oftmals nur schwer möglich (Miller et al., 2007). In der Literatur finden sich zahllose Beispiele für unterschiedliche Ausprägungen des Parasitismus, die einen fließenden Übergang zum Kommensalismus darstellen, und zumeist das Resultat einer ausgeprägten Koevolution zwischen Parasit und Wirt sind (vgl. Mideo, 2009; Cézilly et al., 2010; Joop et al., 2016). Ob und inwiefern ein Wirt durch eine Parasitose also zu Schaden kommt, hängt neben der Befallsintensität und der Virulenz des Parasiten auch von der Anpassung des Wirtes an diesen ab.

Parasitäre Lebensweisen sind bei Organismen aller taxonomischen Gruppen zu finden. Aus Infektionsbiologischer Sicht werden alle Prionen, Viren, Bakterien, Pilze, Protozoa und Metazoa, die Wirte infizieren können, als Parasiten gewertet. Zur besseren Abgrenzung von Bakteriologie, Mykologie und Virologie wird der Begriff Parasit in der medizinischen Parasitologie auf Arthropoden, Helminthen und Protozoen beschränkt (Castro et al., 1996). Aufgrund ihrer enormen Diversität nehmen Infektionen mit Protozoen einen hohen medizinischen Stellenwert ein (Lozano et al., 2012).

1.1 Das Phylum Apikomplexa

Protozoen repräsentieren eine diverse Gruppe einzelliger Eukaryoten, die weltweit in nahezu allen terrestrischen und aquatischen Ökosystemen vertreten sind. Über 200.000 unterschiedliche Spezies wurden bislang beschrieben, die als Protozoen klassifiziert werden können (Cox, 2002). Die Mehrheit dieser sind freilebende Organismen. Etwa 10.000 beschriebene Vertreter weisen jedoch eine parasitische Lebensweise auf (Cox, 2002).

Parasitische Protozoen sind, und waren schon immer, schwer zu klassifizieren (Cox, 2002). Traditionell wurden die meisten parasitären Protozoen taxonomisch in die Klasse der Sporozoen eingeordnet. Die Einordnung erfolgte weitgehend auf dem Morphospezies-Konzept. Heute weiß man allerdings, dass es sich bei Sporozoen um keine phylogenetisch verwandte Gruppe handelt, sondern dass dies viel mehr ein Kollektivname für einzellige heterotrophe und sporenbildende Eukaryoten ist, die eine parasitische Lebensweise aufweisen (Alters, 1999). Viele ehemalige Sporozoen werden heute im Phylum der Apikomplexa zusammengefasst. Dazu zählen ungefähr 6.000 beschriebene Arten (Adl et al., 2007). Sequenzierungsdaten von Umweltproben deuten

jedoch darauf hin, dass diese nur einen kleinen Teil der Diversität der Apikomplexa repräsentieren (Adl et al., 2007). Schätzungen zufolge weisen alle Invertebraten und Vertebraten mindestens einen spezifischen Parasiten aus dem Phylum der Apikomplexa auf (Morrison, 2009).

Den meisten beschriebenen Arten dieses Phylums wird eine große human- und veterinärmedizinische Bedeutung beigemessen. Hierzu zählen unter anderem Erreger der Toxoplasmose (*Toxoplasma gondii*), mit dem rund 25-30 % der Bevölkerung infiziert sind (Montoya et al., 2004), die Erreger der Kryptosporidiose (*Cryptosporidium parvum* und *C. hominis*), die für rund 20% aller Durchfallerkrankungen bei Kindern in Entwicklungsländern verantwortlich sind (Shrivastava et al., 2017), sowie die Erreger der Gattungen *Eimeria*, *Theileria* und *Babesia*, die diarrhöische und fieberhafte Erkrankungen bei Geflügelarten, Schafen und Rindern hervorrufen können und enorme wirtschaftliche Schäden verursachen (Kaltungo et al., 2013). Mit über 210 Millionen Infektionen pro Jahr, von denen schätzungsweise etwa 450.000 tödlich verlaufen, sind jedoch die Erreger der Malaria aus der Gattung *Plasmodium* zweifelsohne die bedeutendsten Vertreter der Apikomplexa (WHO, 2018).

Namensgebend für die Apikomplexa ist die charakteristische Ansammlung struktureller und sekretorischer Organellen am apikalen Ende invasiver, motiler Stadien (Aikawa, 1971). Dieser sogenannte apikale Komplex ist maßgeblich an den Invasionsprozessen der Wirtszelle beteiligt (Morrissette et al., 2002; Gubbels et al., 2012). Zusammengesetzt wird dieser Komplex in der Regel aus den sekretorischen Mikronemen und Rhoptrien sowie aus zytoskelettalen Strukturen wie dem Polring und dem Konoid (Abb. 1.1 a). Zusätzlich werden auch eine variierende Anzahl an elektronendichten Vesikeln (*dense granules*) zum apikalen Komplex gezählt, die im Zytoplasma des Parasiten verteilt sind und ihren Inhalt am Ende der Invasion in die Wirtszelle abgeben (Mercier et al., 2005).

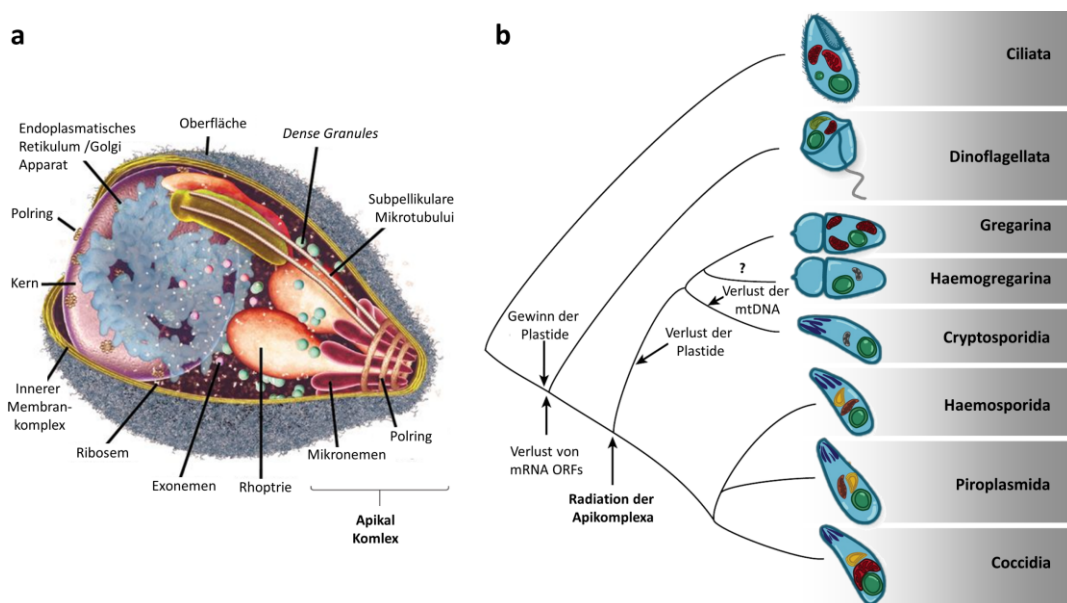


Abb. 1.1: Schematischer Aufbau (a) und phylogenetische Stellung (b) der Apikomplexa. (Abb. nach Cowman et al., 2006; Vaidya et al., 2009; modifiziert)

Ausgehend von den Polringen erstreckt sich das minimalistische Zytoskelett der Parasiten, bestehend aus einem Netzwerk von subpellikularen Mikrotubuli, einem proteinhaltigen Membranskelett und abgeflachten Vesikeln, den sogenannten Alveoli (Bannister et al., 2000). Membranskelett und Alveoli setzen sich zum inneren Membrankomplex zusammen, der gemeinsam mit der äußeren Plasmamembran die Pellikula bildet (Morrisette et al., 2002). Dieses dreischichtige Membransystem ist ein charakteristisches Merkmal, das die Apikomplexa zusammen mit Ziliaten (wie das Wimperntierchen) und Dinoflagellaten (Hauptbestandteil des Phytoplanktons) in die phylogenetische Übergruppe der Alveolata stellt (Adl et al., 2012).

Die Apikomplexa umfassen, neben einigen einzelnen Gattungen und Arten (wie die Kryptosporiden), fünf Hauptgruppen (Haemogregarinen, Gregarinen, Haemosporiden, Piroplasmiden und Kokzidien, siehe Abb. 1.1 b) (Adl et al., 2012). Die genaue Abfolge der Aufspaltungen ist bisher noch nicht vollständig erfasst. Phylogenetische Studien basierend auf rDNA-Analysen, einzelnen Genen oder Genomdaten fassen Haemosporiden und Piroplasmiden jedoch als Aconoidasida zusammen, die eine enge Verwandtschaft zu den Kokzidien zeigen (Escalante et al., 1995; Douzery et al., 2004; Kuo et al., 2008; Borner et al., 2017). Auch die engere Verwandtschaft zwischen Haemogregarinen und Gregarinen, sowie der Gattung *Cryptosporidium* ist inzwischen weitgehend akzeptiert (Ryan et al., 2016).

Auch wenn viele Beziehungen einzelner Gattungen innerhalb dieser Gruppen molekularphylogenetisch fundiert sind, erfolgt die Einteilung einzelner Parasitenarten immer noch primär anhand morphologischer Merkmale, der Art des Wirtes, des Gewebetropismus, sowie anderer Charakteristika im Lebenszyklus (Frénal et al., 2009). In der Regel weisen alle Vertreter der Apikomplexa einen komplizierten Lebenszyklus mit mehreren Differenzierungsstadien und haplohomophasischem Generationswechsel auf, wobei einige Parasiten direkt zwischen ihren Vertebraten-Wirten übertragen werden können (Morrisette et al., 2002). Im Gegensatz dazu gehören die Erreger der Malaria zu der Unterordnung der Haemosporida, deren Vertreter sich durch einen obligaten Wirtswechsel zwischen blutsaugenden Arthropoden und Wirbeltieren auszeichnen (O'Donoghue, 2017).

1.1.1 Malaria und malariaähnliche Parasiten

Haemosporida parasitieren eine Vielzahl von terrestrischen Wirbeltieren. Morphologisch beschrieben wurden dabei hunderte Arten, die Vögel (> 200), Reptilien (> 150), Primaten (> 30) und andere Säugetiere, wie Nager und Fledermäuse, (> 20) infizieren (Outlaw et al., 2014; Faust et al., 2015). Eingeteilt werden diese in rund 20 Gattungen (O'Donoghue, 2017), wobei lediglich die acht Gattungen *Haemocystidium*, *Haemoproteus*, *Hepaticystis*, *Leucocytozoon*, *Nycteria*, *Parahaemoproteus*, *Plasmodium* und *Polychromophilus* sowohl morphologisch als auch molekularbiologisch beschrieben worden sind (z. B. Martinsen et al., 2008; Pineda-Catalan et al., 2013; Borner et al., 2016; Perkins et al., 2016). Parasiten der Gattung *Plasmodium* weisen das breiteste Wirtsspektrum auf und infizieren nahezu alle Vögel, Reptilien, Primaten, Nager, Huftiere und Fledermäuse (Martinsen et al., 2013; Perkins et al., 2016). Arten der Gattung *Haemocystidium* sind dagegen auf Reptilien als Wirt angewiesen (Pineda-Catalan et al., 2013), während Vertreter der Gattungen *Leucocytozoon*,

Haemoproteus sowie *Parahaemoproteus* in Vögeln und Reptilien beschrieben worden sind (z. B. Garnham, 1966; Martinsen et al. 2008; Outlaw & Ricklefs, 2014). Innerhalb der Säugetiere beherbergen Fledermäuse die umfangreichste Anzahl an Haemosporidien. Neben einigen bislang noch nicht weiter molekularbiologisch beschriebenen Gattungen (z. B. *Biguetiella*, *Dionisia*, *Johnsprentia*, *Spratiella*) und *Plasmodium*, wurden in Fledermäusen vor allem Vertreter der Arten *Hepaticocystis*, *Nycteria* sowie *Polychromophilus* identifiziert, wobei *Plasmodium* und *Hepaticocystis* auch andere Säuger infizieren (Garnham, 1973a; Schaer et al., 2013; Ramasindrazana et al., 2018).

Gemeinsam haben diese Parasiten, dass sie in einem bestimmten Stadium ihres Lebenszyklus intrazellulär in Blutzellen der Wirbeltiere leben und von blutsaugenden Insekten aufgenommen und verbreitet werden können. Obwohl sich die intrazellulären Blutstadien der einzelnen Arten anhand nur weniger Merkmale unterscheiden, werden diese neben der Taxonomie ihrer Wirte und Zwischenwirte, klassischerweise als Identifizierungskriterien herangezogen (Valkiūnas, 2004a). Die Einteilung in eine der vier Familien Garniidae, Haemoproteidae, Leucocytozoidae und Plasmodiidae basiert vorrangig auf der Kombination von An- oder Abwesenheit zweier Hauptmerkmale: die asexuelle Vermehrung in Erythrozyten (erythrozytäre Schizogonie) und die Entstehung des Hämozoin, einer kristallinen Substanz, die aus Häm-Molekülen des verstoffwechselten Hämoglobins zusammengesetzt ist (Galen et al., 2018). Gebildet wird das Hämozoin von Arten, die den Familien Haemoproteidae oder Plasmodiidae zugordnet werden (Galen et al., 2018). Die erythrozytäre Schizogonie findet dagegen in den Gattungen der Familien Plasmodiidae und Garniidae statt (Valkiūnas et al., 2017). Vertreter der Leucocytozoidae teilen keine der beiden Eigenschaften (Valkiūnas, 2004a). Definitionsgemäß werden lediglich Parasiten, die sowohl das Hämozoin als auch die erythrozytäre Schizogonie aufweisen, und damit zu den Plasmodiidae gezählt werden, als Malaria Parasiten bezeichnet (Pérez-Tris et al., 2005). Interessanterweise besteht inzwischen jedoch Einigkeit über die paraphyletische Abstammung der Malaria Parasiten in Hinblick auf Parasiten, die malariaähnliche Krankheiten vor allem bei Fledermäusen hervorrufen (z. B. Perkins et al., 2002; Martinsen et al., 2008; Witsenburg et al., 2012; Borner et al., 2016; Galen et al., 2018).

1.1.2 Phylogenetische Beziehungen der Haemosporida

Erste Studien zur phylogenetischen Verwandtschaft der Haemosporida beruhten klassischerweise auf morphologischen Charakteristika der Parasiten. Vor allem Kriterien wie Größe, Form und Art der infizierten Wirtszellen in giemsa-gefärbten Blutaussstrichen der Zwischenwirte, sowie die Anzahl der gebildeten Merozoiten, Bildung des Hämozoin oder Grad der Parasitämie wurden dabei berücksichtigt (Carreno et al., 1997; Valkiūnas, 2004a). Diese Eigenschaften, die auch heute noch hauptsächlich zur Speziesidentifizierung herangezogen werden, deuteten auf einen monophyletischen Ursprung der Gattung *Plasmodium*, während die Gattung *Leucocytozoon* dagegen an der Basis der Haemosporida vermutet wurde (Carreno et al., 1997; Perkins, 2014).

Teilweise unterscheiden sich die Parasiten nur bedingt in ihrer äußeren Form (Perkins, 2014). Zudem wird diese Vorgehensweise zur phylogenetischen Rekonstruktion auch dadurch erschwert, dass sich viele Merkmale durch ähnliche ökologische Selektionsdrücke konvergent entwickeln. Die

Signifikanz dieser Charakteristika bei der phylogenetischen Rekonstruktion wurde daher bereits vielfach in Frage gestellt (z. B. Kissinger et al., 2002; Outlaw et al., 2011; Witsenburg et al., 2012).

Aufgrund ihrer medizinischen Bedeutung waren die ersten molekularen phylogenetischen Analysen auf primaten- und nagetierinfizierende Plasmodienarten begrenzt (Perkins, 2014). Die ersten Ansätze erfolgten dabei mit einer kleinen Taxonanzahl auf der Basis des 18 SSU rRNA-Gens (z. B. Waters et al., 1991; Escalante et al., 1994; Kissinger et al., 2002). Mit der Veröffentlichung der ersten mitochondrialen Nukleotidsequenzdaten setzte sich jedoch die phylogenetische Analyse auf Basis des Cytochrom b-Gens (*cytb*) durch, die auch heute noch in vielen Publikationen Anwendung findet (z. B. Escalante et al., 1995; Boundenga et al., 2018). Auch die umfangreiche Analyse von Perkins und Schall aus dem Jahre 2002 basierte auf *cytb* (Perkins et al., 2002). Neben diversen Plasmodienarten, aus Säugetieren, Reptilien und Vögeln, enthielt die Analyse Parasiten der Gattungen *Haemoproteus*, *Hepatocystis* und *Leucocytozoon*. Mit dem Piroplasmiden *Theileria annulata* als Außengruppe konnte dabei eine basale Position der *Leucocytozoon* gestützt werden. Gleichzeitig konnte erstmals ein polyphyletischer Ursprung der Gattung *Plasmodium* gezeigt werden, bei dem reptilien- und vogelinfizierende Arten eine Schwesterngruppe mit der Gattung *Haemoproteus* bildeten, während eine nähere Verwandtschaft zwischen *Hepatocystis* und den säugetierinfizierenden *Plasmodien* unterstützt wurde (Perkins et al., 2002). Eine weitestgehend übereinstimmende Topologie mit einer monophyletischen Gruppe, die alle Vertreter der Gattung *Plasmodium* und *Hepatocystis* einschließt, resultierte auch aus der ersten Multigenanalyse (Martinsen et al., 2008). Neben dem *cytb* wurden dabei drei zusätzliche Loci aus dem Kerngenom, dem mitochondrialen Genom sowie dem Genom des Apikoplasten verwendet. Eine besondere Erkenntnis dieser Analyse war, dass die Diversifizierung der Parasiten mit einem Wechsel der Vektorfamilie einherging (Martinsen et al., 2008). Bestätigt wird diese Hypothese auch durch neuere Analysen (Galen et al., 2018; Pacheco et al., 2018). Dabei wird im Allgemeinen ein genereller Trend von reptilien- und vogelinfizierenden Gattungen zu säugetierinfizierenden Parasiten nahegelegt, der mit einem Übergang zu Vertretern der Stechmücken (*Culicidae*) als Vektor verbunden ist.

Weitere Multigenanalysen, die den Datensatz von Martinsen et al. (2008) um molekulare Sequenzen von weiteren Gattungen und teilweise auch wiederentdeckten einzelnen Arten erweiterten, brachten eine Reihe unterschiedlicher Topologien zutage (Witsenburg et al., 2012; Schaer et al., 2013; Boundenga et al., 2016; Lutz et al., 2016; Martinsen et al., 2016). Insbesondere fledermausinfizierende Haemosporida warfen in diesem Zusammenhang immer wieder neue Fragen auf. Im Gegensatz zu Plasmodien weisen die Vertreter der Gattungen *Hepatocystis*, *Nycteria* und *Polychromophilus* keine asexuelle Vermehrung in Erythrozyten auf (erythrozytäre Schizogonie). Dennoch wurde immer wieder eine nahe Verwandtschaft zwischen *Hepatocystis* und Nagetierinfizierenden Plasmodien nahe gelegt (Boundenga et al., 2018), während *Nycteria* und *Polychromophilus* teilweise zu vogel- und reptilieninfizierenden Plasmodien (Witsenburg et al., 2012; Schaer et al., 2015; Boundenga et al., 2016) oder als Basis zu den säugetierinfizierenden Plasmodien gruppiert wird (Schaer et al., 2013; Lutz et al., 2016). In Hinblick auf die fledermausinfizierenden Haemosporida belegt die Mehrheit dieser Studien einen paraphyletischen Ursprung der Gattung

Plasmodium, aber auch ein polyphyletischer Ursprung wird in der Literatur diskutiert (Outlaw et al., 2011).

Interessanterweise unterstützt auch eine aktuelle Multigenanalyse den polyphyletischen Ursprung von *Plasmodium* (Galen et al., 2018). Im Gegensatz zu den vorausgegangenen Analysen basiert diese auf 21 Gen-Sequenzen und unterstützt die Einteilung von *Plasmodium* in mindestens fünf unabhängige Gruppen: die Untergattungen *Plasmodium*, *Vinckeia* und *Laverania* sowie zwei Gruppen, die huftierinfizierende bzw. vögel- und reptilieninfizierende Gattungen einschließt (Abb. 1.2). Die Ergebnisse dieser Arbeit implizieren einen einfachen Ursprung der säugetierinfizierenden Haemosporida aus einem vogel- und reptilieninfizierenden Vorfahren, der mit dem Vektorwechsel zur Stechmücke einherging, und einen konsekutiven Wirtswechsel zurück zu Sauropsiden (Galen et al., 2018). Zugleich zeigt die Analyse, dass die Untersuchung von fledermausinfizierenden Haemosporiden für ein besseres Verständnis der Evolutionsgeschichte der Gattung *Plasmodium* wesentlich ist. Der Wechsel der Parasiten zur oder von der Fledermaus, wird dabei sogar als Schlüsselfaktor zur Diversität der Haemosporiden angesehen (Galen et al., 2018).

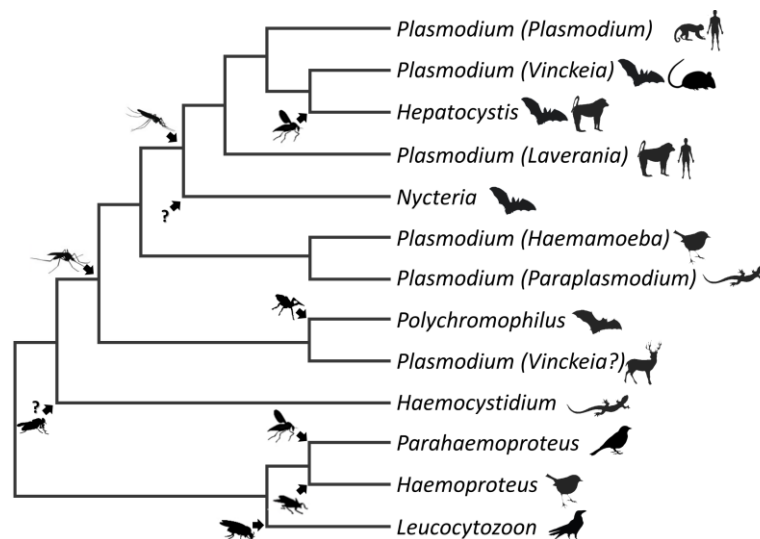


Abb. 1.2: Phylogenetische Beziehungen innerhalb der Haemosporida.

Beispielhafte Wirtsarten sind als Silhouette neben den entsprechenden Haemosporida Taxa dargestellt. Vektorenguppen sind dagegen mit einem Pfeil der entsprechenden Verzweigung markiert (Topologie entspricht den Ergebnissen von Galen et al., 2018.)

1.2 *Polychromophilus* als Modelorganismus

Erste Beschreibungen von Haemosporiden in Fledermäusen gehen bis in das 19. Jahrhundert zurück (vgl. Megali et al., 2011; Schaer et al., 2013). Basierend auf morphologischen Einzelberichten aus dem letzten Jahrhundert wurden bisher mindestens acht fledermausinfizierende Gattungen identifiziert: *Biguetiella*, *Dionisia*, *Hepatocystis*, *Johnspretia*, *Nycteria*, *Plasmodium*, *Polychromophilus* und *Sprattiella* (Garnham et al., 1953; Landau et al., 1978; Landau et al., 1980a; Landau et al., 2012a; Landau et al., 2012b; Schaer et al., 2013; Ramasindrazana et al., 2018). Mit Ausnahme der Gattungen *Plasmodium* und *Hepatocystis* handelt es sich dabei um Gattungen, die ausschließlich Fledermäuse infizieren.

Diese außergewöhnliche Vielfalt der fledermausinfizierenden Haemosporiden liegt mit großer Wahrscheinlichkeit auch in der großen phylogenetischen Diversität der Fledermäuse begründet. Nach den Nagetieren bilden Fledermäuse mit über 1300 beschriebenen Spezies die artenreichste Ordnung der Säugetiere mit einer nahezu weltweiten Verbreitung (Burgin et al., 2018). Basierend auf molekularen Daten wurden Haemosporiden bislang in sieben Fledermausfamilien nachgewiesen: Hipposideridae, Megadermatidae, Miniopteridae, Pteropodidae, Rhinolophidae, Rhinonycteridae und Vespertilionidae (Duval et al., 2007; Lutz et al., 2016; Ramasindrazana et al., 2018). Die Gattungen *Plasmodium*, *Hepatocystis* und *Nycteria* umfassen Malariaparasiten, die bislang vor allem in Fledermäusen der Unterordnung Yinpterochiroptera der tropischen Bereiche Asiens und Afrikas identifiziert wurden. Im Gegensatz dazu infizieren Parasiten der Gattung *Polychromophilus* Vertreter der Familien Miniopteridae und Vespertilionidae (Unterordnung Yangochiroptera) und sind bislang die einzigen fledermausinfizierenden Haemosporiden, die eine weite Verbreitung sowohl in tropischen als auch in gemäßigten Regionen zeigen (Schaer et al., 2018).

Innerhalb der Gattung *Polychromophilus* wurden fünf Arten morphologisch unterschieden. *P. adami* und *P. corradetti* wurden ursprünglich in der Kronen-Langflügelfledermaus (*Miniopterus inflatus*) und der zentralafrikanischen Langflügelfledermaus (*M. minor*) beschrieben, die in weiten Teilen Afrikas südlich der Sahara beheimatet sind (Landau et al., 1980b). Die Art *P. deanei* wurde dagegen lediglich vereinzelt in Mittel- und Südamerika nachgewiesen und ursprünglich in der gemeinen Blütenfledermaus (*Glossiphaga sorcinia*) entdeckt (Deane et al., 1961). Als Typuswirt des Parasiten wird jedoch das schwarze Mausohr (*Myotis nigriscans*) angesehen, in der *P. deani* zehn Jahre später wiederentdeckt wurde (Garnham et al., 1971). Die anderen beiden Arten, *P. melanipherus* und *P. murinus*, die erstmals in der Langflügelfledermaus (*Miniopterus schreibersii*) bzw. in der Zweifarbfledermaus (*Vespertilio murinus*) identifiziert wurden, zeigen jeweils wesentlich größere Verbreitungsgebiete in Afrika, Asien, Australien und Europa (Dionisi, 1898a, 1898b; Garnham, 1973a). Während *P. melanipherus* fast ausschließlich in Wirten der Gattung *Miniopterus* nachgewiesen wurde (Duval et al., 2012; Ramasindrazana et al., 2018), wurde *P. murinus* in unterschiedlichen Vertretern aus der Familie Vespertilionidae identifiziert, scheint aber besonders häufig in Arten der Gattung *Myotis* aufzutreten (Megali et al., 2011; Schaer et al., 2013).

Übertragen werden die Parasiten von Fledermausfliegen (Nycteribiidae) der Gattungen *Nycteribida*, *Penicillida* und *Basilia* (Mer et al., 1947; Gardner et al., 1988; Megali et al., 2011). Dabei handelt es sich um flügellose Ektoparasiten der Ordnung Diptera, die obligat mit ihrem Wirt, der Fledermaus, verbunden sind. Da Fledermausfliegen jedoch ausschließlich auf Fledermäuse spezialisiert sind, wird vermutet, dass das erste Auftreten des *Polychromophilus*-Vorfahren bei Fledermäusen entweder auf Moskitos (Culicidae spp.) oder auf Lausfliegen (Hippoboscidae spp.) zurückzuführen ist (Witsenburg et al., 2012). Die klare genetische Differenzierung der Arten *P. melanipherus* und *P. murinus* und ihre Spezifität auf der Ebene der Wirtsgattung impliziert für diesen Wirtswechsel einen Zeitraum um die Diversifizierung der Miniopteridae und Vespertilionidae vor 49-38 Millionen Jahren (Miller-Butterworth et al., 2007; Duval et al., 2012). Diese Gegebenheit macht die Vertreter der Gattung *Polychromophilus*, insbesondere *P. melanipherus* und *P. murinus*, zu ganz besonders interessanten Modelorganismen: Es handelt sich um Parasiten mit einer universellen

Verbreitung, einer phylogenetischen Position vermutlich an der Basis der Plasmodien und mit einer natürlichen Assoziation zwischen Wirt und Vektor, die einen bemerkenswerten Fall von Koevolution darstellt.

1.2.1 Biologie und Ökologie der Wirte

Infektionen von Fledermäusen sind aufgrund ihrer medizinischen Bedeutung als Reservoir von hochpathogenen Krankheitserregern in den letzten Jahren immer mehr zu einem Schwerpunkt der Forschung geworden (Calisher et al., 2006; Brook et al., 2015). Studien zu fledermausspezifischen Erregern sind dagegen jedoch immer noch eher dürftig. Tatsächlich wurde der Vektor von *P. melanipherus* beispielsweise bis heute noch nicht nachgewiesen, obwohl davon ausgegangen wird, dass der Parasit von *Nycteribia schmidlii*, der wirtsspezifischen Fledermausfliege der Typusart *Miniopterus schreibersii*, übertragen wird (Witsenburg et al., 2015a; Ramasindrazana et al., 2018). Im Gegensatz dazu ist der Vektor von *P. murinus*, die Fledermausfliege *Nycteribia kolanatii* (Theodor et al., 1954), bereits vor über 30 Jahren beschrieben worden (Gardner et al., 1988).

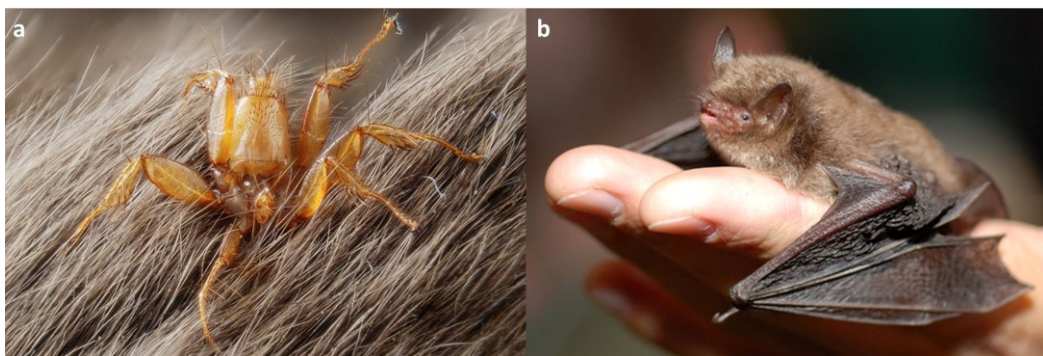


Abb. 1.3: End- und Zwischenwirt von *P. murinus*.

Weibliche Fledermausfliege der Familie Nycteribiidae (a) und *Myotis daubentonii* (b). (San Martin, 2008, 2010.)

1.2.1.1 Die Fledermausfliege *Nycteribia kolanatii*

Wie auch andere Fledermausfliegen zeigt *N. kolanatii* extreme morphologische Anpassungen an einen parasitären Lebensstil (Abb. 1.3 a): die Augen und Flügel der gelb-braunen Insekten sind vollständig reduziert, die Flügelmuskulatur ist verkümmert und der Körper dorsoventral abgeflacht (Hutson, 1984). Die Extremitäten der Fledermausfliege sind die im Verhältnis zum Körper extrem lang und liegen dem Thorax dorsal auf. Dadurch sind die Beine weit gespreizt, was den 2-2,5 mm großen Insekten ein insgesamt sehr spinnenähnliches Erscheinungsbild verleiht (Petersen et al., 2007). Die Beinanordnung ermöglicht der außergewöhnlich schnellen Fledermausfliege sich auf dem Wirt in jede Richtung zu bewegen und so der Fellpflege des Wirtes zu entgehen (Scheffler, 2008; Dick et al., 2015). Darüber hinaus verfügt die Fledermausfliege über eine Reihe weiterer morphologischer Eigenschaften, die einen dauerhaften Aufenthalt auf ihren Wirten ermöglichen. Durch Tarsallkrallen, Borstenkämme an Kopf, Thorax, Abdomen und Beinen sowie kammartigen Auswüchse des Chitinpanzers, den Ctenidien, ist *N. kolanatii* stark an der Fortbewegung und Verankerung im Fell ihrer fliegenden Wirte angepasst (Dick et al., 2015). Die Fledermausfliegen

sind, unabhängig vom Geschlecht, obligate Haematophagen, die beim Wirtswechsel in der Regel auf den direkten Körperkontakt ihrer Wirte angewiesen sind. Lediglich weibliche Fledermausfliegen verlassen ihre Wirte, um ihre verpuppungsreifen Larven im Quartier der Wirtsfledermäuse abzulegen (Szentiványi et al., 2016). Im Vergleich zu anderen vielen anderen Insekten weisen die Vertreter der *Nycteribidae* damit eine außergewöhnliche Reproduktionsstrategie auf: alle Larvenstadien entwickeln sich innerhalb des Weibchens (adenotrophe Viviparie) (Dick et al., 2006). Die abgelagerten Larven bilden innerhalb kürzester Zeit eine rot-braune Puppe, aus der sich die adulte Fledermausfliege entwickelt, die wiederum eine neue Wirtsfledermaus zur Kolonisierung aufsucht (Hutson, 1984).

Der Reproduktionszyklus der langlebigen Fledermausfliegen ist eng an den ihrer Wirte gekoppelt (Lourenço et al., 2008). Obwohl *N. kolenatii* gegenüber Umwelteinflüssen sehr widerstandsfähig ist und sich auch in den Wintermonaten vom Blut winterschlafender Fledermäuse ernährt, erreicht die Populationsgröße der Ektoparasiten im Sommer zur Zeit der Wochenstuben ihr Maximum (Hutson, 1984; Gardner et al., 1988; Witsenburg et al., 2015b). Da *N. kolenatii* hochgradig wirtsspezifisch ist, entspricht auch die geographische Verbreitung der Fledermausfliege weitgehend der ihrer Wirtsfledermäuse. Als Hauptwirt des Ektoparasiten fungiert vor allem die Wasserfledermaus *Myotis daubentonii* (Kuhl, 1817) (Abb. 1.3 b), deren Lebensraum sich von Westeuropa bis Japan erstreckt (Hutson, 1984; Bogdanowicz, 1994; Witsenburg et al., 2014). Allerdings existieren auch vereinzelte Aufzeichnungen von anderen parasitierten Fledermausarten der Familie Vespertilionidae. Interessanterweise wurde *N. kolenatii* unter anderem auch auf den Fledermäusen *Eptesicus serotinus*, *Myotis myotis*, *Nyctalus noctula*, *Rhinolophus hipposideros* und *Vespertilio murinus* beschrieben (Szentiványi et al., 2016). Dabei handelt es sich um Arten, in denen auch bereits vereinzelnd *P. murinus* nachgewiesen wurde (Dionisi, 1898a; Megali et al., 2011; Borner et al., 2016).

1.2.1.2 Die Wasserfledermaus *Myotis daubentonii*

Obwohl es sich bei *M. daubentonii* nicht um den Typuswirt von *P. murinus* handelt, kann inzwischen davon ausgegangen werden, dass es sich dabei um den primären Wirt in Europa handelt (Gardner et al., 1987; Megali et al., 2011). Im Vergleich zu anderen europäischen Fledermäusen weist die Wasserfledermaus auch einen besonders starken Befall mit Fledermausfliegen auf (Haelewaters et al., 2017). In britischen Populationen waren rund 73 % aller Individuen mit durchschnittlich zwei Parasiten der Art *N. kolenatii* infiziert (Hutson, 1984). Eine deutlich höhere Befallsintensität mit durchschnittlich sechs Fledermausfliegen pro Fledermaus, konnte dagegen in einer Population in der Schweiz nachgewiesen werden (Witsenburg et al., 2015b). Die hohe Infektionsrate der Wasserfledermäuse mit *N. kolenatii* spiegelt sich auch bei *P. murinus* wider: hier wurden in unterschiedlichen Wasserfledermauspopulationen Prävalenzen von etwa 50 % bzw. 75 % nachgewiesen (Megali et al., 2011; Witsenburg et al., 2014). Die hohe Prävalenz bei Fledermäusen wird in der Literatur zu einem gewissen Grad auch auf ihre hohen Populationsdichten innerhalb von Fledermauskolonien zurückgeführt, die eine regelmäßige horizontale Übertragung von Parasiten begünstigen (Calisher et al., 2006; Lourenço et al., 2008).

Wie andere Fledermausarten sind Wasserfledermäuse hochsoziale Tiere, die den Großteil ihres Lebens in Gruppen verbringen (August et al., 2014). Im Sommer lassen sich diese Gruppen weitgehend durch eine temporäre, räumliche Trennung der Geschlechter charakterisieren. Dafür sammeln sich rund 20-30 Weibchen im Frühjahr (Ende März/Anfang April) an Quartieren, die im Sommer auch zur Aufzucht der Jungtiere als Wochenstuben genutzt werden (Lučan et al., 2011). Männchenkolonien umfassen dagegen im Durchschnitt nur etwa 10 Tiere (Encarnação et al., 2005). Als Waldfledermaus ist *M. daubentonii* vor allem in strukturreichen Habitaten mit einem hohen Waldanteil anzutreffen. Hier nutzen die relativ kleinen Fledermäuse (4-6 cm und 7-15 g) vor allem Baumhöhlen oder -spalten, seltener Felsspalten, Gebäude oder Nistkästen als Tagesquartiere, in denen sie meist engen Körperkontakt mit anderen Tieren suchen. Da die Wasserfledermaus zur Futtersuche auf stehende und langsam fließende Gewässer angewiesen ist, korreliert die Anzahl der Fledermäuse mit dem Anteil der Wasseroberflächen im Habitat (Dietz et al., 2006). Mit dem kurzhaarigen dunkelgrau-braunen Fell, das sich stark von der grauweißen Unterseite des Körpers abhebt, ist die nachtaktive Fledermaus daher vor allem an Teichen und Seen gut zu beobachten.

M. daubentonii ernährt sich fast ausschließlich von Wasserinsekten, bevorzugt werden dabei Zuckmücken, die mit den auffallend großen, borstenbehaarten Füßen oder der Schwanzflughaut von der Wasseroberfläche erbeutet werden (Vaughan, 1997). Insgesamt ist die Lebensweise der Wasserfledermäuse an die witterungsbedingten Schwankungen der Nahrungsverfügbarkeit der gemäßigten Zone angepasst. Das reichhaltige Nahrungsangebot im Frühjahr und Sommer korreliert mit dem erhöhten Nahrungsbedarf der Wasserfledermaus während der energieintensiveren Reproduktionsphasen Schwangerschaft, Laktation und Spermatogenese. Dabei können die Tiere täglich nahezu die Hälfte des eigenen Gewichts an Insekten konsumieren (Encarnação et al., 2006). Nahrungsengpässe in der Sommerperiode können mit Tagestorpor überbrückt werden, bei dem die Stoffwechselrate der Wasserfledermaus reduziert wird (Dietz & Kalko, 2006). Die insektenarme, kalte Jahreszeit verbringt *M. daubentonii* im Winterschlaf. Als Winterquartiere dienen vor allem frostgeschützte Höhlen oder Stollen, die meistens nicht mehr als 100 km von den Sommerquartieren entfernt sind (Rogowska et al., 2007). Bereits im Spätsommer kann die Wasserfledermaus an den Winterquartieren beobachtet werden (Dragu et al., 2019). Dabei zeigen die Tiere eine erhöhte Flugaktivität in Gruppen. Das ausgeprägte Schwärmverhalten von *M. daubentonii* spielt einerseits eine Rolle bei der Lokalisierung geeigneter Überwinterungsgebiete und ist andererseits eine wichtige Komponente im Paarungsverhalten der Wasserfledermaus (Angell et al., 2013; van Schaik et al., 2015). Wie bei den meisten europäischen Fledermäusen wird der Reproduktionszyklus der Wasserfledermaus durch den Winterschlaf unterbrochen. Da Spermatogenese und Ovulation mit einem größeren zeitlichen Abstand erfolgen, werden die Spermien nach der Kopulation im Genitaltrakt der Weibchen gespeichert (Orr et al., 2014). Damit die Geburt der Jungtiere mit der maximalen Nahrungsverfügbarkeit in den Sommermonaten zusammenfällt, finden Eisprung und Befruchtung erst nach dem Winterschlaf statt. In der Regel bringen die Weibchen nur einmal im Jahr ein einzelnes Jungtier zur Welt. Die Tragzeit der Tiere beträgt rund zwei Monate, sodass die Populationsgröße in den Sommerquartieren der Wasserfledermaus etwa Anfang Juli ihr Maximum erreicht (Encarnação et al., 2010).

1.2.2 Epidemiologische Aspekte

Die hohen Populationsdichten von Wirt und Vektor in den Sommermonaten wirken sich auch auf die Prävalenz des intrazellulären Parasiten *P. murinus* aus. Dabei steigen die Infektionsraten der Fledermäuse bis in den späten September hinein stetig an (Witsenburg et al., 2014). Die Fellpflege der Fledermäuse ist eine effiziente Maßnahme um den Befall mit Ektoparasiten in Grenzen zu halten (Christe et al., 2007). Jungtiere weisen dagegen eine nur unzureichende Fellpflege auf und sind daher auch häufiger mit *N. kolenatii* und *P. murinus* infiziert (Witsenburg et al., 2014). Aufgrund ihres noch nicht vollständig ausgebildeten Immunsystems weisen die juvenilen Fledermäuse gleichzeitig auch Infektionen mit der höchsten Parasitenlast auf (Witsenburg et al., 2014). Einen Unterschied in der Infektionsrate zwischen Weibchen und Männchen, wie zuvor auch bereits bei *P. melanipherus* beschrieben, scheint es dagegen bei *M. daubentonii* nicht zu geben (Megali et al., 2011; Ramasindrazana et al., 2018). Im Gegensatz zu weiblichen Fledermäusen wechseln männliche Wasserfledermäuse häufiger ihre Quartiere und spielen eine zentrale Rolle bei der sozialen Vernetzung unterschiedlicher Kolonien und Wochenstuben (August et al., 2014). Gleichzeitig tragen die Männchen damit wahrscheinlich auch entscheidend zur Ausbreitung des Parasiten bei.

Über die Auswirkungen einer Infektion mit *P. murinus* ist bislang nur wenig bekannt, insbesondere die physiologischen Folgen für die Fledermaus wurden bisher recht wenig untersucht. Grippeähnliche Symptome, wie sie bei einer Malariainfektion charakteristisch sind, oder Anämien wurden bei mit *P. murinus* infizierten Wasserfledermäusen bislang nicht nachgewiesen (Witsenburg et al., 2014). Allerdings scheint die Infektion einen negativen Effekt auf die körperliche Verfassung vor allem bei den subadulten Tieren zu haben (Witsenburg et al., 2014). Zurückzuführen ist dieses Phänomen vor allem auf die hohen energetischen Kosten, die mit einer Immunreaktion einhergehen (Lochmiller et al., 2000). Wesentlich bedrohlicher ist die Infektion mit *P. murinus* dagegen für *N. kolenatii*. Hier wirkt sich die Infektion negativ auf die Lebensdauer der infizierten Fledermausfliegen aus (Witsenburg et al., 2015b). Weibliche Fledermausfliegen haben insgesamt einen längeren Lebensspanne und bauen während der Tragzeit deutlich mehr Fettreserven als männlichen Fledermausfliegen auf (Haelewaters et al., 2017). Dennoch scheinen die Auswirkungen der Infektion unabhängig vom Geschlecht des Vektors zu sein. Im Vergleich zur Fledermauspopulation ist die Prävalenz von *P. murinus* bei *N. kolenatii* mit rund 16 % deutlich geringer (Witsenburg et al., 2015b). Zurückzuführen ist dies einerseits auf die gesteigerte Mortalitätsrate durch die Infektion und andererseits durch ihr adaptiven Vorzug von Fledermäusen mit geringeren Infektionsraten (Witsenburg et al., 2015b).

1.2.3 Morphologie und Lebenszyklus

Der Lebenszyklus von *P. murinus* ist sehr gut beschrieben. Wie bei anderen Haemosporiden verläuft die Entwicklung von *P. murinus* in einem komplexen Vermehrungszyklus mit einer Vielzahl unterschiedlicher Stadien, einem haplohomophasischen Generationswechsel und einem obligatorischen Wirtswechsel (Gardner et al., 1987). Der Lebenszyklus lässt sich dabei grob in drei

Abschnitte unterteilen. Die geschlechtliche Vermehrung, die Gamogonie, mit anschließender Reduktionsteilung, der Sporogonie, findet definitionsgemäß im Endwirt, der Fledermausfliege, statt. Im Zwischenwirt, der Fledermaus, vermehrt sich *P. murinus* zunächst ungeschlechtlich in unterschiedlichen Organen (Schizogonie) und differenziert sich anschließend in Erythrozyten zu Gametozyten, die sich erst in der Fledermausfliege weiterentwickeln. Einen allgemeinen Überblick über die wesentlichen Entwicklungsstadien ist in Abb. 1.4 dargestellt.

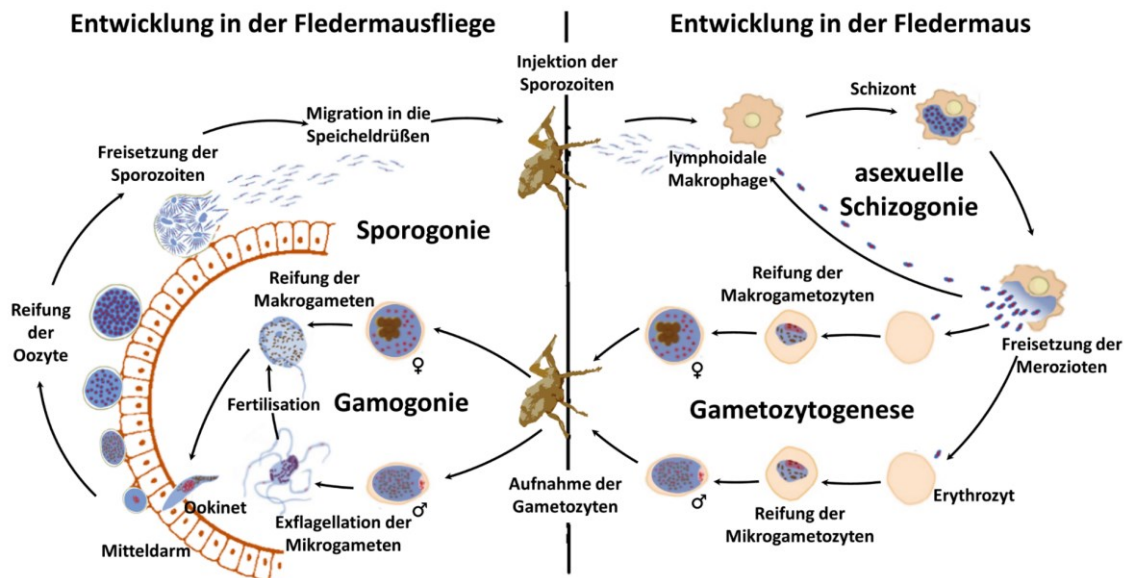


Abb. 1.4: Schematische Darstellung des Entwicklungszyklus von *P. murinus*.

Die geschlechtliche Vermehrung (Gamogonie) des Parasiten, beginnt mit der Ausreifung von Gametozyten in den Erythrozyten der Fledermaus (Gametozytogenese). Diese werden von der Fledermausfliege aufgenommen, wo die Fertilisation dieser im Mitteldarm stattfindet. Der daraus hervorgegangene Ookinet reift in der Darmwand der Fledermausfliege zur Oozyste aus und setzt Sporozoiten frei (Sporogonie), die über die Speicheldrüsen der Fledermausfliege die Fledermaus infizieren. Über die Blutbahn gelangen die Sporozoiten zu phagozytierenden Zellen diverser Organe und reifen zu Schizonten aus, die Merozoiten in die Blutbahn entlassen. Diese entwickeln sich zu erneuten Gewebsschizonten (asexuelle Schizogonie) oder differenzieren sich in den Erythrozyten der Fledermaus zu den für die Fledermausfliege infektiösen Gametozyten. (Abb. nach Gardner et al., 1987; Centers for Disease Control and Prevention, 2018.)

1.2.3.1 Geschlechtliche Vermehrung in der Fledermausfliege

Die Transmission des Parasiten von der Fledermaus zur Fledermausfliege erfolgt bei der Blutmahlzeit in Form von intraerythrozytären Gametozyten. Im Verdauungstrakt der Fledermausfliege reifen die aufgenommenen Gametozyten innerhalb kürzester Zeit zu haploiden Gameten heran. Induziert werden kann dieser Vorgang, die Gamogonie, bereits allein durch den abrupten Temperaturabfall zwischen warmblütigem Zwischenwirt und wechselwarmem Vektor (z. B. Garnham et al., 1967; Billker et al., 1997; Palinauskas et al., 2013). Dabei runden die intraerythrozytären Gametozyten innerhalb von vier bis fünfzehn Minuten ab und permeabilisieren die umgebende Erythrozytenmembran (Gardner et al., 1988) (Abb. 1.5 a). Während sich weibliche Makrogametozyten anschließend zu kugelförmigen Makrogameten differenzieren, durchläuft der Mikrogametozyt drei mitotische Kernteilungen und formt flagellenartige Mikrogameten (Sinden et al., 1976). Die Mikrogameten werden in einem rapiden, lichtmikroskopisch leicht

identifizierbaren Vorgang, der als Exflagellation bezeichnet wird, vom Residualkörper des Mikrogametozyten ausgestoßen (Abb. 1.5 b). Obwohl in den ursprünglichen Beschreibungen lediglich die Freisetzung von maximal sechs 10 µm großen, motilen Mikrogameten aus einem Mikrogametozyten beobachtet werden konnte, erscheint die Bildung von acht Mikrogameten pro Zelle, wie auch bei anderen Haemosporiden konserviert, wahrscheinlicher (Gardner et al., 1988; Carreno et al., 1997; Valkiūnas, 2004b).

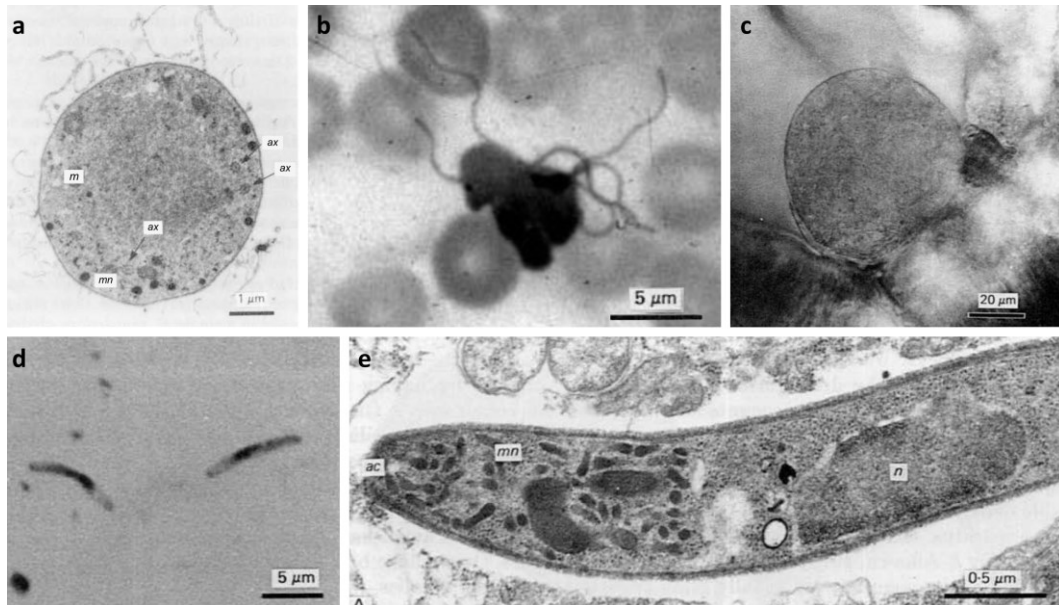


Abb. 1.5: Mikroskopische Aufnahmen unterschiedlicher Stadien in der Fledermausfliege.

Ruptur der Erythrozytenmembran vor der Exflagellation (a), Exflagellation der Mikrogameten (b), Reifer Oozyt an der äußeren Mitteldarmwand (c), freigesetzte Sporoziten aus der Speicheldrüse von *N. kolenatii* als Giemsa gefärbtes Präparat (d) und elektronenmikroskopische Aufnahme. (Abb. nach Gardner et al., 1988; Molyneux, 2012; modifiziert.)

Nach der Fertilisation von Mikrogamet und Makrogamet entwickelt sich aus der Zygote ein länglicher Ookinet. Dabei handelt es sich um eine Zellform, die bei *P. murinus* noch nicht beschrieben wurde, in näher verwandten Arten jedoch das Darmlumen verlässt und in das Mitteldarmepithel dringt (Valkiūnas, 2004b). Im extrazellulären Raum zwischen Epithel und Basalmembran differenziert sich der Ookinet schließlich zur Oozyste und leitet damit die Sporogonie innerhalb des Endwirtes ein (Abb. 1.5 c). Dabei durchläuft die Oozyste, geschützt von einer äußeren zweiphasigen Kapsel, die 0,3-1 µm dick werden kann, mehrere Kernteilungen (Gardner et al., 1988). Hat die reife Oozyste, die dann auch als Sporoblast bezeichnet wird, eine Größe von etwa 60-70 µm erreicht, werden mehr als 1.000 Sporozoiten in die Hämolymphe der Fledermausfliege freigesetzt (Gardner et al., 1988; Carreno et al., 1997) (Abb. 1.5 d, e). Über die Basalmembran gelangen die geraden bis leicht halbmondförmigen Sporozoiten schließlich in das Lumen der Speicheldrüsen. Dort persistiert der Parasit, bis er mit dem abgesonderten Speichel bei einer weiteren Blutmahlzeit der Fledermausfliege auf die Fledermaus übertragen wird.

1.2.3.2 Ungeschlechtliche Vermehrung in der Fledermaus

Mit der Blutmahlzeit einer infizierten Fledermausfliege wird die Entwicklung des Parasiten in der Fledermaus eingeleitet. Durch den Speichel, den die Fledermausfliege vor der eigentlichen Nahrungsaufnahme zur Verhinderung der Blutgerinnung injiziert, gelangt *P. murinus* in Form von Sporozoiten in die Haut der Fledermaus. Nah verwandte Parasiten, allen voran die Säugetier-infizierenden Plasmodien, bewegen sich in der Dermis zufällig in dem Bereich der Einstichstelle bis sie in dermale Blutgefäße eindringen und über das Blutgefäßsystem im Wirt verteilt werden (Amino et al., 2006; Yamauchi et al., 2007). Andere Haemosporiden vermehren sich dagegen zunächst auch in Makrophagen oder Fibroblasten in der direkten Umgebung der Einstichstelle (Valkiūnas et al., 2017). Die initiale Infektion der Fledermaus durch *P. murinus* ist noch weitgehend ungeklärt. Exoerythrozytäre Formen des Parasiten treten lediglich in geringer Anzahl auf und sind recht schwer zu identifizieren (Garnham, 1973; Landau et al., 1980b). Beschrieben wurden Schizonten in Knochenmark, Lunge, Niere, Milz und Leber (Mer et al., 1947). Dabei werden scheinbar lediglich gewebsständige, lymphoidale Makrophagen (Phagozyten des retikuloendothelialen Systems) infiziert, wobei auch bereits Schizonten in Makrophagen des peripheren Bluts beschrieben worden sind (Manwell, 1946; Garnham, 1973; Landau et al., 1980 und darin enthaltene Referenzen). Obwohl die Zusammenhänge zwischen den einzelnen Stadien noch unklar sind, lassen die beschriebenen Schizonten auf eine zweiphasige asexuelle Schizogonie schließen. Teilweise konnten in diversen Zellen unterschiedlicher Organe bis zu 30 µm große, nicht eingekapselte Schizonten identifiziert werden, die über hundert Merozoiten enthielten (Mer et al., 1947; Landau et al., 1980b) (Abb. 1.6 a, b). In einer der ersten Beschreibungen werden dagegen Schizonten mit lediglich 20 bis 22 Merozoiten nachgewiesen (Manwell, 1946). Parallel dazu konnten auch deutlich kleinere kugelförmige Schizonten in der Leber identifiziert werden, die nur bis zu 10 µm groß werden (Landau et al., 1980b). Auch die nah verwandte Art *P. daenii* bildet in der Leber deutlich kleinere Schizonten mit bis zu 40 Merozoiten (Marinkelle, 1995) (Abb. 1.6 c).

Nach einer unbekannten Anzahl an asexuellen Vermehrungen werden, wahrscheinlich ausgehend von den kleineren Leber-Schizonten der Kupfferzellen, Merozoiten in das Blut freigesetzt, die sich in Erythrozyten zu weiblichen oder männlichen Gametozyten differenzieren (Garnham, 1973a). Begünstigt wird die Gametozytogenese, zumindest bei den nah verwandten Plasmodienarten, durch eine erhöhte Parasitämie oder eine verstärkte Immunantwort des Wirtes sowie die damit verbundene Stresszunahme des Parasiten (Talman et al., 2004). Auch der Invasionsprozess der Merozoiten in die Erythrozyten, ist mit den nah verwandten Plasmodienarten wahrscheinlich vergleichbar. Hier kommt es zu einer ersten, reversiblen Bindung an Oberflächenrezeptoren, einer Reorientierung, bei der das apikale Ende des Parasiten zur Erythrozytenmembran gerichtet wird, der Freisetzung der Substanzen aus Rhoptrien und Mikronemen bis schließlich zur Einstülpung der Erythrozytenmembran (Koch et al., 2016).

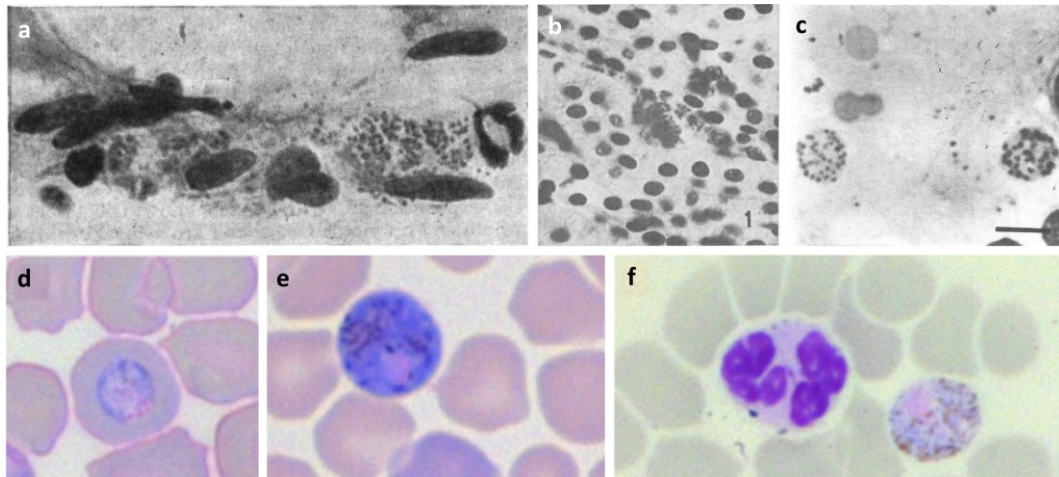


Abb. 1.6: Mikroskopische Aufnahmen unterschiedlicher Stadien in der Fledermaus.

Reife Schizonten im Knochenmark (a), in der Niere (b) und in der Leber (c), unreifer Gametozyt (d), Makro- (e) und Mikrogametozyt (f). (Abb. nach Mer et al., 1947; Landau et al., 1980; Marinkelle, 1995; Witsenburg et al., 2014; modifiziert.)

Nach Abschluss des Invasionsprozesses ist der intrazelluläre Parasit von der parasitophoren Vakuole umgeben, deren Membran sich einerseits aus der parasitären Plasmamembran und andererseits aus der Membran des Erythrozyten zusammensetzt. Die parasitophore Vakuole konnte in diversen Stadien beschrieben werden und bleibt während der gesamten Entwicklung der Gametozyten von *P. murinus* erhalten (Gardner et al., 1988). Junge Gametozyten erscheinen im Blutausschlag dabei als typische Ringstadien, die sich durch zentrale Nahrungsvakuolen und randständige Nukleoli charakterisieren lassen (Garnham, 1973a) (Abb. 1.6 d). Mit voranschreitender Entwicklung der Gametozyten wird das vom Hämoglobinabbau freigesetzte Häm aus der Wirtszelle immer mehr zum Hämozoïn polymerisiert. Dadurch erscheinen zunehmend mehr lichtmikroskopisch sichtbare Aggregate in unterschiedlicher Größe (Pigmente), die zufällig im Zytoplasma von *P. murinus* verteilt sind. Der ausgewachsene Gametozyt füllt die gesamte Wirtszelle aus und ist etwas größer als ein uninfizierter Erythrozyt (Gardner et al., 1988). Weibliche Makrogametozyten und männliche Mikrogametozyten unterscheiden sich in Form und Größe und lassen sich aufgrund unterschiedlicher Farbintensitäten in Giemsa-gefärbten Blutausschlägen leicht voneinander differenzieren (Gardner et al., 1988; Witsenburg et al., 2014) (Abb. 1.5 e, f). Mikrogametozyten sind in der Regel etwas größer als die weiblichen Formen, und eher oval (Witsenburg et al., 2014). Nach Färbung erscheint das Zytoplasma der Mikrogametozyten hellblau mit einem großen, rosa-gefärbten, eher netzförmig diffusem Zellkern mit wenig kondensiertem Chromatin (Gardner et al., 1988; Witsenburg et al., 2014). Die sphärischen Makrogametozyten weisen nach Färbung ein intensiv, blau-violettes, wahrscheinlich besonders ribosomreiches Zytoplasma auf, und lassen sich zudem durch einen kleinen, kompakten exzentrischen Zellkern charakterisieren (Gardner et al., 1988). Beide Gametozytenformen enthalten nach vollständiger Reifung 10-30 Pigmente, ein geschlechtsspezifischen Unterschied in der Anzahl und Lokalisierung dieser, wie er bei nah verwandten Arten beschrieben wurde, scheint es nicht zu geben (Valkiūnas, 2004c; Witsenburg et al., 2014). Die intraerythrozytären Gametozyten zirkulieren im inaktiven Zustand im Blut, bis sie wiederum von der Fledermausfliege über eine

Blutmahlzeit aufgenommen werden. Makrogametozyten sind im Blut der Fledermaus dabei häufiger vertreten (Foster, 1979). Zurückzuführen ist dieses Phänomen intuitiv darauf, dass der männliche Gametozyt bis zu acht Mikrogameten ausbildet, sodass im Mitteldarm der Fledermausfliege ein ausgewogenes Geschlechterverhältnis für die Fertilisation vorliegt (Henry et al., 2019).

1.3 Parasit-Wirt-Interaktionen auf Proteinebene

Der komplexe Lebenszyklus des Parasiten, der sich durch einen stetigen Wechsel von Wirtszellen charakterisieren lässt, erfordert eine permanente Interaktion mit den unterschiedlichsten Oberflächenstrukturen seiner Wirte. Bereits beim Eindringen führt die bloße Anwesenheit des Parasiten zur Stimulierung des Immunsystems des Wirtes (Clayton et al., 2014; Gomes et al., 2016; Belachew, 2018). Sowohl bei der Fledermausfliege als auch bei der Fledermaus sind spezifische Zellrezeptoren (pattern recognition receptors, PRRs) in der Lage gut konservierte molekulare Strukturen der Haemosporiden (pathogen-associated molecular patterns, PAMPs) zu erkennen und Signalkaskaden zu aktivieren, die in einer direkten jedoch unspezifischen Abwehr der Parasiten münden (vgl. Buchmann, 2014). Auf der anderen Seite haben pathogene Erreger Wege gefunden, die generelle Abwehrreaktionen ihrer Wirte zu umgehen, zu unterdrücken oder sogar zu ihren Gunsten zu manipulieren (z. B. Schmid-Hempel, 2008; Belachew, 2018; Thakur et al., 2019). Auch die intrazelluläre Lebensweise der Haemosporiden ist eine Adaption, der die Parasiten vor dem Einfluss des Immunsystems zu schützen vermag. Gleichzeitig erfordert diese jedoch eine gut auf den Wirt abgestimmte Interaktion, bei der viele Wechselwirkungen zwischen den Proteinen von Wirt und Parasit involviert sind. Die Mechanismen der Invasion und anschließenden Modifikation der Wirtszellen spielen in Hinblick auf Parasiten-Wirt-Interaktionen daher eine ganz besondere Schlüsselrolle. Viele der Mechanismen sind über alle Vertreter der Gattung *Plasmodium* hinweg konserviert, teilweise handelt es sich dabei jedoch auch um linien-spezifische Innovationen, die auch in nah verwandten Parasiten nicht zu finden sind (z. B. Baum et al., 2006a; Reid, 2015; Koch et al., 2016).

1.3.1 Export putativ wirtszell-interagierender Proteine

Die meisten Proteine, die direkt oder indirekt in die Invasionsprozesse der infektiösen Haemosporiden beteiligt sind, sind entweder GPI-verankerte Proteine, Membranproteine, oder peripher-assoziierte Proteine, die an solchen gebunden sind (Beeson et al., 2016). Bei der Mehrheit dieser Proteine handelt es sich jedoch um Typ-1-Transmembranproteine, die neben einer C-terminalen Transmembrandomäne ein hydrophobes Signalpeptid innerhalb der ersten hundert Aminosäuren aufweisen (Sargeant et al., 2006). Durch das Signalpeptid werden die Proteine über das endoplasmatische Retikulum (ER) in den sekretorischen Transportweg eingeleitet bevor sie basierend auf ihrer spezifischen Signalsequenz in eine der sekretorischen Organellen gelangen (Kats et al., 2006). Obwohl noch nicht genauer charakterisiert, spielen in diesem Zusammenhang

auch unterschiedliche Proteasen, die als Reifungsproteasen während des Transports fungieren, eine wichtige Rolle (Kim, 2004; Besteiro et al., 2009; Parussini et al., 2010; Jimenez-Ruiz et al., 2016).

Die proteolytische Prozessierung als Regulationsmechanismus zur Lokalisierung von Proteinen ist auch wesentlicher Bestandteil der intrazellulären Stadien der Parasiten (Chang et al., 2008). Zur Etablierung und Erhaltung aller für den Parasiten essenziellen Kompartimente und Strukturen an und in einer infizierten Wirtszelle bedarf es einer Vielfalt von Proteinen, die mit den wirtseigenen Proteinen wechselwirken. Der intrazelluläre Proteintransport über die parasitophore Vakuolenmembran ist in der Gattung *Plasmodium* konserviert, jedoch am ausführlichsten in *P. falciparum* untersucht (de Koning-Ward et al., 2009; Russo et al., 2010; Matthews et al., 2013). Die exportierten Proteine können dabei in zwei Gruppeneingeteilt werden, abhängig davon, ob sie das pentamere Signalmotiv (RxLxE/Q/D), das *Plasmodium* Export Element (PEXEL), aufweisen oder nicht (Hiller et al., 2004; Marti et al., 2004). PEXEL-Proteine werden im endoplasmatischen Retikulum durch die Protease Plasmepsin V gespalten, N-acetyliert und schließlich in die PV sezerniert, wo sie über das Translokon PTEX (*Plasmodium* translocon of exported proteins) in die Wirtszelle transportiert werden (Boddey et al., 2009; de Koning-Ward et al., 2009; Russo et al., 2010). Proteine mit PEXEL-Motiv sind in *P. falciparum* besonders häufig, andere Arten der Gattung *Plasmodium* weisen dagegen vermehrt Proteine auf, die alternative PEXEL-ähnliche Sequenzen aufweisen, oder keine konservierten Motive besitzen (PEXEL-negative exported Proteins, PNEPs) (Sargeant et al., 2006; Pick et al., 2011; Schulze et al., 2015; Spielmann et al., 2015; Ngatho et al., 2019).

1.3.2 Genomanalyse zur Identifizierung relevanter Proteine und damit verbundene Hürden

Wesentliche Aspekte der zugrundeliegenden molekularen Interaktionen zwischen Haemosporida und ihren Wirten konnten erst durch die Sequenzierung einiger Vertreter der Gattung *Plasmodium* identifiziert werden (Carlton et al., 2002; Gardner et al., 2002; Carlton et al., 2005). Der erste Entwurf des humanpathogenen Malariaerregers *P. falciparum* wurde nach sieben Jahren internationaler Bemühungen im Jahr 2002 erfolgreich abgeschlossen (Gardner et al., 2002; Le Roch et al., 2012). Dabei handelt es sich um ein haploides Genom, dessen Größe ursprünglich auf 22,8 Mbp geschätzt wurde. Das Kerngenom verteilt sich auf 14 Chromosomen, die sich über 0,6-3,3 Mbp erstrecken. Zusätzlich zu seinem Kerngenom weist *P. falciparum* ein lineares Element aus tandemförmig wiederholten 6kbp Einheiten, das mitochondriale Gene enthält, und das zirkuläre 35 kbp große Apikoplasten-Genom auf (Wilson et al., 1991).

Mit der Etablierung neuer Sequenziermethoden (Next generation sequencing, NGS) wurde die Genomsequenzierung seit dem Jahr 2005 kontinuierlich erleichtert (Margulies et al., 2005; Gužvić, 2013). Seither ist auch die Anzahl der veröffentlichten Genome aus dem Taxon der Haemosporida rapide angestiegen. Neben den Genomen aller humanpathogenen Arten findet man inzwischen auch die Sequenzen von primaten-, nager- und vogelinfizierenden Plasmodienarten, darunter auch die Genome der Modellorganismen *P. berbei*, *P. yoelii*, *P. relictum* und *P. gallinaceum* (Gardner et al., 2002;

Carlton et al., 2002; Carlton, 2003; Carlton et al., 2008; Pain et al., 2008; Bright et al., 2012; Tachibana et al., 2012; Otto et al., 2014a; Otto et al., 2014b; Böhme et al., 2018; Su et al., 2019). Auch das Genom einer *Haemoproteus*-Art wurde inzwischen sequenziert (Bensch et al., 2016). Ein Vergleich zeigt, dass die einzelnen Vertreter kompakte Genome in der Größe von 18-30 Mbp besitzen, die, auf 14 Chromosomen verteilt, etwa 4.600-4.700 konservierte protein-kodierende und eine variierende Anzahl an artspezifischen akzessorischen Genen aufweisen (Otto et al., 2014a; Bushell et al., 2017).

Abgesehen von den umfassenden Informationen, die aus der Genomanalyse über die einzelnen Arten gewonnen werden können, ermöglichen vergleichende Analyse über die verschiedenen Arten hinaus, ein breites Spektrum an Forschungsschwerpunkten. Damit ist es möglich sowohl artspezifische Anpassungen als auch wesentliche konservierte Funktionen während des komplexen Lebenszyklus der Parasiten zu identifizieren, die auf neue Angriffsziele für Interventionen hinweisen können (Galinski, 2019). Viele konservierte Mechanismen bei der Invasion und der Zellmodifikation, aber beispielsweise auch der PEXEL-abhängige Proteintransport, konnten bereits durch Sequenzanalysen bei *P. falciparum* beschrieben werden (Sargeant et al., 2006). Gleichzeitig sind aber noch viele Fragen zur Biologie der Parasiten, wie z. B. auch zu konservierten Immunevasionsstrategien, zur Transmission, zu Wirtswechseln oder zu konservierten Adaptationen offen.

Die biologischen Unterschiede aber auch Gemeinsamkeiten zwischen fledermausinfizierenden Gattungen und *Plasmodium* und die phylogenetische Verwandtschaft zwischen diesen machen *Polychromophilus* zu einer potenziell interessanten Gattung für das Verständnis wichtiger Aspekte der Biologie von Malariaparasiten. Bislang blieb eine Sequenzierung auch aufgrund vieler technischer Hürden noch aus. Das größte Hindernis für eine direkte Sequenzierung von Haemosporiden ist die geringe Menge an Parasiten-DNA im Vergleich zur Wirts-DNA (Auburn et al., 2011; Venkatesan et al., 2012). Dies hat inzwischen eine Reihe an Methoden hervorgebracht, die eine effiziente Aufreinigung der DNA parasitären Ursprungs versprechen. Aktuelle Lösungen Wirtskontamination zu umgehen, umfassen die Anreicherung durch hybride Selektionstechniken, bei denen DNA-Sonden, die komplementär zur DNA des Parasiten sind, an die Probe hybridisiert werden (Melnikov et al., 2011; Oyola et al., 2013). In anderen Projekten erfolgt eine Auftrennung auf Zellebene, bei dem Leukozyten von Erythrozyten getrennt, infizierte Zellen in Kulturen angelegt oder Parasiten aus ihren Wirtszellen herausisoliert werden (Shermann et al., 1960; Schuster, 2002; Trang et al., 2004; Auburn et al., 2011; Venkatesan et al., 2012; Palinauskas et al., 2013). Andererseits werden aber auch Methoden zur DNA-Trennung genutzt, die auf einer CsCl-Gradienten-Ultrazentrifugation oder auf der spezifischen Depletion der Wirtszell-DNA beruhen (Dame et al., 1987; Fechery et al., 2013; Oyola et al., 2013). Die Charakteristika der einzelnen Parasiten-Genome aber auch Besonderheiten im Lebenszyklus können die Möglichkeiten zur Isolierung von Parasit und Wirt deutlich einschränken (Johnston et al., 1999). Daher erfordern Genomprojekte bislang individualisierte Lösungen, die von der zu untersuchenden Art abhängig sind.

1.4 Zielsetzung

Die vollständige Sequenzierung der Genome einiger Vertreter der Gattung *Plasmodium* eröffnete neue Möglichkeiten zur Identifizierung adaptiver Parasit-Wirt-Interaktionen auf molekularer Ebene (Carlton et al., 2002; Gardner et al., 2002; Pain et al., 2008; Liu et al., 2016; Böhme et al., 2018). Viele identifizierte Mechanismen sind dabei auf Proteine und Proteinfamilien zurückzuführen, die bislang nur in der Gattung *Plasmodium* identifiziert werden konnten (Sargeant et al., 2006; Wasmuth et al., 2009; Otto, Böhme, et al., 2014; Reid, 2015). In Hinblick auf die paraphyletische Abstammung der Gattung *Plasmodium* kann ein besseres Verständnis evolutiver Parasit-Wirt-Interaktionen jedoch nur gewonnen werden, wenn auch nah verwandte Parasiten berücksichtigt werden. Im Zuge der vorliegenden Arbeit wird deshalb ein Fokus auf das Genom des heimischen Fledermausparasiten *Polychromophilus murinus* gelegt. Im Gegensatz zu Plasmodien weisen Parasiten der Gattung *Polychromophilus* keine erythrozytäre Schizogonie auf, sodass in Blutproben von Fledermäusen nur eine geringe Parasitämie vorliegt und nur wenig genomisches Material gewonnen werden kann. Die Sequenzierung des kompletten Genoms eines intrazellulären Parasiten, der nicht kultiviert werden kann und aufgrund der geringen Parasitenlast nur in begrenzter Menge aus seinem Wirt isoliert werden kann, stellt methodisch eine besondere Herausforderung dar. Daher sollten in einem ersten Schritt Methoden etabliert werden, die genomische DNA des Malariaparasiten in ausreichender Menge für eine Genomsequenzierung bereitstellen.

Ziel dieser Arbeit war die Sequenzierung und die Annotation des Genoms sowie eine erste molekulare Charakterisierung eines fledermausinfizierenden Parasiten der Gattung *Polychromophilus*. Der Fokus sollte dabei auf Proteinen liegen, die eine putative Rolle bei der Interaktion mit dem Wirt haben. Dies soll dazu beitragen, die Gattung *Polychromophilus* phylogenetisch einzuordnen und konservierte sowie spezifische Merkmale als Schlüsselfaktoren der Parasit-Wirt-Interaktionen im Taxon der Haemosporida zu identifizieren. Weiterhin soll dadurch das Verständnis wichtiger evolutiver Vorgänge, wie die Parasitenresistenz, Wirtswechsel und die Entstehung von Pathogenität verbessert werden

2 Material

2.1 Arbeitsgeräte

Im Rahmen dieser Arbeit wurden folgende Geräte verwendet:

Agarosegelkammern:	Owl EasyCast B1A Minigel-Elektrophoresesystem (ThermoFischer Scientific, St. Leon-Rot), Owl EasyCast B1A B2 Gel-Elektrophoresesystem (ThermoFischer Scientific, St. Leon-Rot), Horizontal-Elektrophoresekammer (A. Hartenstein, Würzburg)
Autoklav:	Systec VE-55 (Systec, Wettenberg), Tuttnauer 3850 ELV (Biomedis, Gießen)
Brutschränke:	Innova 4200 Incubator Shaker (Brunswick Scientific, Nürtingen), Innova 4200 Incubator Shaker (Brunswick Scientific, Nürtingen)
Binokular:	SZ61 (Olympus, Hamburg)
Eismaschine:	GB-1555 (Brema Ice Makers, Villa Cortese, (IT))
Geldokumentation:	Fusion Fx7 (Peqlab Biotechnologies, Erlangen)
Heizblöcke:	Bio TDB-100 (A. Hartenstein, Würzburg)
Homogenisator:	MINILys Personal Homogenisator, (Bertin Technologies, Montigny-le-Bretonneux (F))
Horizontal-Schüttler:	Stuart See-saw Rocker SSL4 (Barloworld Scientific, Staffordshire), Unimax 2010 (Heidolph Instruments, Schwabach)
Magnetrührer:	RSM-10HS (Phoenix Instruments, Garbsen), RCT basic (IKA Labortechnik, Staufen)
Mikroskop:	Olympus BX51 (Olympus, Hamburg)
Mikrowelle:	Grill (LG, Seoul, (KR))
Photometer:	Nanodrop 1000 (Peqlab Biotechnologies, Erlangen) Qubit 2.0 Fluorometer (Invitrogen, Karlsruhe)
pH-Messgerät:	InoLab pH 720 (WTW, Weilheim)
Reinstwasseranlage:	Mili-Q (Milipore, Bedford, (US))
Rotator:	Intelli Mixer RM2 (Elmi, Riga, (LV))
Spannungsgeber:	Power supplies EV 243 (Consort, Turnhout, (BE)) Power supplies EV 245 (Consort, Turnhout, (BE))
Sicherheitswerkbank:	LaminAir HV2448 (Heraeus, Hanau)
Thermocycler:	ABI 7500 Real-Time PCR System (Thermo Fisher Scientific, St. Leon-Rot) T1 Thermocycler (Biometra, Göttingen) LabCycler Basic (SensoQuest GmbH, Göttingen) LabCycler Gradient (SensoQuest GmbH, Göttingen)
Thermomixer:	Thermomixer 5436 (Eppendorf, Hamburg)
Ultrazentrifuge:	Beckmann L-80 Ultrazentrifuge (Beckmann Coulter, Krefeld)
Vakuumpzentrifuge:	Speed Vac SC110 (Thermo/Savant, Dreieich),

Vortexer:	REAX2000 (Heidolph Instruments, Schwabach) Vortex-Genie 2 (Scientific Industries, Bohemia)
Waagen:	Präzisionswaage SE 622 (VWR, Darmstadt), Sartorius BP210S (Sartorius AG, Göttingen), Sartorius Quintinx (Sartorius AG, Göttingen)
Wasserbad:	Wasserbad 1002 (GFL, Burgwedel) Wasserbad 1083 (GFL, Burgwedel)
Zentrifugen:	Sigma 1-14 Mikrozentrifuge (Sigma, Osterode am Harz), Heraeus Labofuge 400R (Heraeus, Hanau), Plate Centrifuge (Starlab, Hamburg), Zentrifuge 5415R (Eppendorf, Hamburg), Zentrifuge 5417R (Eppendorf, Hamburg)

2.2 Gebrauchsmaterialien

Alle Gebrauchsmaterialien wie Reaktionsgefäße (0,2/0,5/1,5/2 ml), 96-well-Platten, Zentrifugenröhrchen (15/50 ml), serologische Kunststoffpipetten (5/10/25 ml) sowie Pipettenspitzen mit und ohne Filter (10/200/1000 µl) wurden, wenn nicht anders angegeben, von den Firmen Eppendorf (Hamburg), Greiner Bio-One (Frickenhäusen), Hirschmann (Eberstadt), Sarstedt (Nümbrecht-Rommelsdorf) und Starlab (Ahrensburg) bezogen. Verwendete chirurgische Einmal-Skalpelle, Injektionskanülen und Einmalspritzen stammten vom Hersteller B. Braun (Melsungen), Objektträger, Deckgläser von ThermoFischer Scientific (St. Leon-Rot). In der Regel wurden nichtsterile Materialien für 20 min bei 121 °C autoklaviert.

2.3 Chemikalien

Die verwendeten Chemikalien wurden von den Firmen in der Regel im Reinheitsgrad p.a. (pro analysi, analysenrein) bezogen und ohne vorherige Aufreinigung eingesetzt.

Agar-Agar, Kobe I	Carl Roth, Karlsruhe
Agarose Low Melt	Carl Roth, Karlsruhe
Agarose NEEO Ultra-Qualität	Carl Roth, Karlsruhe
Ampicillin Natriumsalz	Carl Roth, Karlsruhe
Borsäure	Carl Roth, Karlsruhe
Bromphenolblau	Serva, Heidelberg
Cäsiumchlorid	Carl Roth, Karlsruhe
Cellulose (medium fibers)	Sigma-Aldrich, Steinheim
Citronensäure	Merck, Darmstadt
Chloroform	Carl Roth, Karlsruhe
di-Kaliumhydrogenphosphat (K_2HPO_4)	Carl Roth, Karlsruhe
di-Natriummethylenediamintetraessigsäure (Na_2EDTA)	Carl Roth, Karlsruhe
di-Natriumhydrogenphosphat Dihydrat ($Na_2HPO_4 \cdot 2 H_2O$)	Carl Roth, Karlsruhe

DPBS (engl. Dulbecco's Phosphate Buffered Saline)	Sigma-Aldrich, Steinheim
Essigsäure	Carl Roth, Karlsruhe
Ethanol >99,8 % (v/v)	Carl Roth, Karlsruhe
Ethidiumbromidlösung 1%	Carl Roth, Karlsruhe
Ethylendiamintetraessigsäure (EDTA)	Carl Roth, Karlsruhe
Ficoll PM 400 (Polysuccrose 400)	Sigma-Aldrich, Steinheim
Giemsa-Stammlösung (Azur-Eosin-Lösung)	Carl Roth, Karlsruhe
Harnstoff	Carl Roth, Karlsruhe
Hoechst 33258, Pentahydrate 10 mg/ml	Life Technologies, Darmstadt
Hefeextrakt	Carl Roth, Karlsruhe
Isoamylalkohol	Carl Roth, Karlsruhe
Isopropanol 99 % (v/v)	Carl Roth, Karlsruhe
Isopropyl- β -D-thiogalactosid (IPTG)	Carl Roth, Karlsruhe
Kaliumchlorid (KCl)	Carl Roth, Karlsruhe
Kaliumdihydrogenphosphat (KH_2PO_4)	Merck, Darmstadt
Natriumacetat	Carl Roth, Karlsruhe
Natriumchlorid (NaCl)	Carl Roth, Karlsruhe
Natriumdodecylsulfat (SDS Pellets)	Carl Roth, Karlsruhe
Natriumhydroxid (NaOH)	Carl Roth, Karlsruhe
Roti-Phenol/Chloroform/Isoamylalkohol	Carl Roth, Karlsruhe
Saccharose	Carl Roth, Karlsruhe
Salzsäure (HCl)	Carl Roth, Karlsruhe
Tri-Natriumcitrat Dihydrat	Carl Roth, Karlsruhe
Tris	Carl Roth, Karlsruhe
Trypton/Penton aus Casein	Carl Roth, Karlsruhe
X- β -Gal (5-Brom-4-chlor-3-indoxyl- β -D-Galactosid)	Carl Roth, Karlsruhe

2.4 Lösungen und Puffer

Wenn nicht anders angegeben wurden alle verwendeten Lösungen und Puffer mit deionisiertem Reinstwasser angesetzt. Angegebene pH-Werte beziehen sich auf Raumtemperatur und wurden mit HCl bzw. NaOH eingestellt. Die Sterilisation erfolgte entweder durch Autoklavieren (20 min bei 121°C) oder durch Sterilfiltration (sterile Spritzenfilter, Porengröße: 0,22 μm , Carl Roth, Karlsruhe).

Ampicillin-Stammlösung:	1 % (w/v) Ampicillin Natriumsalz.
Cäsiumchlorid-Lösung:	40,6 % (w/v) Cäsiumchlorid.
Ethidiumbromidlösung:	1 ml/l Ethidiumbromidlösung in TBE- oder TAE-Puffer.
Ficolllösung:	5,7 % (w/v) Ficoll in DPBS.
Giemsa-Färbelösung:	5 % (v/v) Giemsa-Stammlösung in Phosphatpuffer.
IPTG/X-Gal-Platten:	1,5 % (w/v) Agar-Agar, 0,01 % (w/v) Ampicillin, 0,5 mM IPTG, 0,01 % (w/v) X- β -Gal in LB-Medium.

LB-Amp-Platten:	1,5 % (w/v) Agar-Agar, 0,01 % (w/v) Ampicillin in LB-Medium.
LB-Medium:	0,5 % (w/v) Hefeextrakt, 1 % (w/v) Trypton, 1 % (w/v) NaCl, pH 7,5.
Loading-Dye (5x):	50 % (w/v) Saccharose, 0,1 % (w/v) Bromphenolblau, 50 mM Na ₂ EDTA, 4 M Harnstoff.
Lysepuffer:	40 mM Tris, 80 mM EDTA, 2 % (w/v) SDS, pH 8,0.
Natriumcitratlösung:	3,96 % (w/v) Tri-Natriumcitrat Dihydrat, 0,026 % (w/v) Citronensäure, pH 7.
PBS-Puffer (10x):	1,4 M NaCl, 27 mM KCl, 81 mM Na ₂ HPO ₄ , 15 mM KH ₂ PO ₄ , pH 7,4.
Phosphatpuffer:	6 mM Na ₂ HPO ₄ , 8 mM KH ₂ PO ₄ , pH 7,2.
SET-Puffer:	0,01 M NaCl, 50 mM Tris, 10 mM EDTA, pH 8.
TAE-Puffer (50x):	2 M Tris, 1 M Essigsäure, 50 mM EDTA, pH 8,5.
TBE-Puffer (10x):	890 mM Tris, 890 mM Borsäure, 500 mM Na ₂ EDTA, pH 8,0.
TE-Puffer:	10 mM Tris, 1 mM EDTA, pH 8,0.
Tris/HCl-Puffer:	40 mM Tris, pH 8,0.

2.5 Molekularbiologische Kits

AccuPrime Taq DNA Polymerase High Fidelity	Life technologies, Darmstadt
CloneJET PCR Cloning Kit	Thermo Fisher Scientific, St. Leon-Rot
Crystal Gel-Extraktionskit	Biolab Products, Bebensee
GeneJET Genomic DNA Purification Kit	Thermo Fisher Scientific, St. Leon-Rot
Illustra GenomiPhi V2 DNA Amplification Kit	GE Healthcare, Solingen
pGEM-T/ -T Easy Vector Systems	Promega, Mannheim
peqGOLD Plasmid Miniprep Kit I	Peqlab Biotechnologies, Erlangen
DNeasy Blood & Tissue Kit	Qiagen, Hilden
Maxima SYBR Green qPCR Master Mix	Fermentas, St. Leon-Rot
NEBNext Microbiome DNA Enrichment Kit	New England BioLabs, Frankfurt am Main
Precellys Keramik-Kit 1.4 mm	Peqlab Biotechnologies, Erlangen
Qubit dsDNA HS Assay Kit	Invitrogen, Karlsruhe
Qubit dsDNA BR Assay Kit	Invitrogen, Karlsruhe

2.6 Enzyme und andere biologische Substanzen

AccuPrime Taq DNA Polymerase System	Thermo Fisher Scientific, St. Leon-Rot
Carrier RNA	ThermoFischer Scientific, St. Leon-Rot
Desoxynucleotide (dNTPs)	Invitrogen, Karlsruhe
DNA Loading Dye (6x)	ThermoFischer Scientific, St. Leon-Rot
Exonuclease I	New England BioLabs, Frankfurt am Main
FspEI (5'...C ^m C(N) ₁₂ '...3')	New England BioLabs, Frankfurt am Main
GeneRuler 1 kb Ladder	ThermoFischer Scientific, St. Leon-Rot

GeneRuler DNA Ladder Mix	ThermoFischer Scientific, St. Leon-Rot
GeneRuler High Range DNA Ladder	ThermoFischer Scientific, St. Leon-Rot
LongAmp Taq Polymerase	New England BioLabs, Frankfurt am Main
MolTaq	Molzym, Bremen
MspJI (5'... ^m CNNR(N) ₉ '...3')	New England BioLabs, Frankfurt am Main
Pfu DNA Polymerase	Promega, Mannheim
phi29 DNA Polymerase	New England BioLabs, Frankfurt am Main
Phusion High-Fidelity DNA Polymerase	New England BioLabs, Frankfurt am Main
Proteinase K	Peqlab Biotechnologies, Erlangen
RNase A	Carl Roth, Karlsruhe
Shrimp Alkaline Phosphatase	Roche, Mannheim
T4 DNA-Ligase	Promega, Mannheim
T4 DNA-Polymerase	Promega, Mannheim

2.7 Oligonukleotide

Die verwendeten Oligonukleotide, im Folgenden lediglich als Primer bezeichnet, wurden von den Firmen Eurofins MWG (Ebersberg) und Sigma-Aldrich (Steinheim) synthetisiert. PCR- und Sequenzierprimer wurden salzfrei und lyophilisiert bezogen. Die modifizierten Primer für die SWGA (vgl. 2.7) wurden mit 1-2 Phosphorothioate-Bindungen am 3'-Ende in HPSF-Qualität (High Purity Salt Free) bestellt. Die Primer wurden nach Herstellerangaben jeweils auf eine Konzentration 100 pmol/μl eingestellt. Zum Einsatz in der PCR oder Sequenzierung (vgl. 2.11.1) wurden Arbeitslösungen von 10 pmol/μl hergestellt. Primersets für die SWGA wurden dagegen in einer Endkonzentration von 50 pmol/μl weiterverwendet. Die Lagerung erfolgte bei -20 °C. Die in dieser Arbeit verwendeten Primer sind in den Tab. 2.1 bis Tab. 2.4 aufgelistet.

Tab. 2.1: Verwendete Primer zur Amplifizierung und Sequenzierung.

Primer	Sequenz [5' – 3']	Produkt (bp)	Referenz
<i>asl (pmu)</i>			
asl_outer_F	GSK AAR TTT AAT GSK GCT GTW GG	~300	(Martinsen et al., 2008)
asl_outer_R	GGA TTA AYT TTA TGA GGC ATT G		
asl_nested_F	GCT GAT MAA AAT RTT GAT TGG	~250	
asl_nested_R	GAG GCA TTG TAC TAC TWC C		
<i>coxI (pmu)</i>			
coxI_outer_F	CTA TTT ATG GTT TTC ATT TTT ATT TGG TA	~1290	(Martinsen et al., 2008)
coxI_outer_R	AGG AAT ACG TCT AGG CAT TAC ATT AAA TCC		
coxI_nested_F	ATG ATA TTT ACA RTT CAY GGW ATT ATT ARG	~1070	
coxI_nested_R	GTA TTT TCT CGT AAT GTT TTA CCA AAG AA		
<i>clpc (pmu)</i>			
clpc_outer_F	AAA CTG AAT TAG CAA AAA TAT TA	~640	(Martinsen et al., 2008)
clpc_outer_R	CGW GCW CCA TAT AAA GGA T		
clpc_nested_F	GAT TTG ATA TGA GTG AAT ATA TGG	~580	
clpc_nested_R	CCA TAT AAA GGA TTA TAW G		
<i>cytb (pmu)</i>			
cytb_outer_F	TAA TGC CTA GAC GTA TTC CTG ATT ATC CAG	~1260	(Martinsen et al., 2008)
cytb_outer_R	TGT TTG CTT GGG AGC TGT AAT CAT AAT GTG		
cytb_nested_F	TCA ACA ATG ACT TTA TTT GG	~680	
cytb_nested_R	TGC TGT ATC ATA CCC TAA AG		

Primer	Sequenz [5' – 3']	Produkt (bp)	Referenz
HaemoScrF1 HaemoScrR1 HaemoScrF2	AAH TAT GGA GYG GWT GGT G TTA RRY TTC TYT GTT CDG C TAA TAC GAC TCA CTA TAG GGA CCW TGG GGW CAA ATG AG	~700 ~350	(Borner et al., 2016)
HaemoScrR2	ATT TAG GTG ACA CTA TAG AAG CAT TAT CWG GAT GWG MTA		
Pmu-cytb_F Pmu-cytb_R	GCT ACT GGT GCT TCA TTT GTA ATC TAC TGT AAT TGC ATT ATC TGG	~520	Diese Arbeit
sf3b1 (pmu)			
Sf3b1_outer_F Sf3b1_outer_R Sf3b1_nested_F Sf3b1_nested_R	GCH ACD ATG ATW GGW AAT ATG AG ANG TYT GDA TWG TWG CTC TYC T CAT GGR TTR VAT GAY GAR CA TCH CCT CCY TTR TCT GCW AT	~1430 ~1070	(Borner et al., 2016)
ef2a (pmu)			
Ef2a_F Ef2a_R	GTT CGT GAG ATC ATG AAC AAA AC CCT TGT AAA CCA GAA CCA AA	~600	(Schaer et al., 2013)
sf3b1 (mda)			
Mda-Sf3b1_F Mda-Sf3b1_R	GCA GCA TTT GAC TAA AGA AGA GA CCA GGA GTC TGA TCA GCT GTT TGA	~150	Diese Arbeit
Apikoplast (pmu)			
Pmu_SSU_F1 Pmu_SSU_F2 Pmu_LSU_F1 Pmu_LSU_F2 Pmu_Api3_R1 Pmu_Api3_R1 Pmu_Api47_F1 Pmu_Api47_F2	AGC AGC AGC GGT AAT ACA GAA TAG GTA GTG CAT GGC TGT CG AGA TAG TTT GAC TGG GGC GG CCA CAT TGA CGA AAA GGT TTG G GGA CTT GAA CCC TTA CAA CTT CCA TTG TGT ATA CCA AAT TTC ACC A ACC TCC ATT TCC TAA ATA ATC ACC T ACG TTC AAA TTG AGC AAA ATC A	vgl. 4.4.3	Diese Arbeit

Api, Apikoplast; asl, Adenylosuccinate Lyase; clpc, Caseinolytic Protease; coxI, Cytochrom Oxidase I; cytb, Cytochrom b, ef2a, Elongation Factor 2; F, forward Primer; mda, *Myotis daubentonii*; LSU, Large subunit ribosomal ribonucleic acid; pmu, *Polychromophilus murinus*; R, reverse Primer; sf3b1, Splicing factor 3B subunit 1; SSU, small subunit ribosomal ribonucleic acid.

Tab. 2.2: Vektorspezifische Primer für Kolonie-PCR und Sequenzierung.

Primer	Sequenz [5' – 3']	Referenz
pGEM-T*		
pGEM-T T7 pGEM-T SP6	TAA TAC GAC TCA CTA TAG GG ATT TAG GTG ACA CTA TAG	Promega, Mannheim
pJet1.2/ blunt*		
pJET1.2_F pJET1.2_R	CGA CTC ACT ATA GGG AGA GCG GC AAG AAC ATC GAT TTT CCA TGG CAG	Thermo Fisher Scientific, St. Leon-Rot

* Die entsprechenden Vektorkarten können aus dem jeweiligen Handbuch des Herstellers entnommen werden.

Tab. 2.3: Primer zur Quantifizierung der relativen Anteile von parasitärer und Wirt-DNA.

Primer	Gen	Sequenz	Produkt (bp)	Referenz
<i>Polychromophilus murinus</i>				
Pmu-Sf3b1_F Pmu-Sf3b1_R	sf3b1	TTG CAT GAC GAG CAA CAA AA AAG CAG CTA ACA CTT TCC CT	~150	Diese Arbeit
Pmuscr1_F Pmuscr1_R	conserved gene	CGC CAT CAA CAG TAC GGT CA AGT TGT TTT CTT GCT TGC TTC TCA	~390	Diese Arbeit
Pmuscr2_F Pmuscr2_R	adh	TTC GCT CGA CCA GTT CTT CC CCG CTG AAG GAT CGT TTA AAT CA	~230	Diese Arbeit
<i>Myotis daubentonii</i>				
Mda-Sf3b1_F Mda-Sf3b1_R	sf3b1	GCA GCA TTT GAC TAA AGA AGA GA CCA GGA GTC TGA TCA GCT GTT TGA	~150	Diese Arbeit

Primer	Gen	Sequenz	Produkt (bp)	Referenz
Mdascr1_F Mdascr1_R	prlr	GAA CGG CAA CCA GAC AGA GA CCA TGC CAA AGA TCA AGC CG	~300	Diese Arbeit
Mdascr2_F Mdascr2_R	pdcd6	AGC CTT CAT ACC CCT TCC CT GGG GCA CAG CTT TCA GTT TG	~220	Diese Arbeit

adh, ATP-dependent helicase; conserved gene, PF3D7_1456500 homolog; F, forward Primer; mda, *Myotis daubentonii*; pdcd6, Programmed cell death protein 6; pmu, *Polychromophilus murinus*; prlr, Prolactin Receptor, R, reverse Primer; sf3b1, Splicing factor 3B subunit 1.

Tab. 2.4: Primerzusammensetzung zur selektiven Anreicherung der parasitären DNA.

Set_ID	Primer	Sequenz [5' – 3']	Referenz
Set1	SWGA02, SWGA03	TAT TAT TAT T*A*G, ATA TAT ATT AT*T* A	Diese Arbeit
Set2	SWGA01, SWGA04	CTA ATA ATA A*T*A, TAA TAA TAT AT*A* T	Diese Arbeit
Set3	SWGA04, SWGA06	TAA TAA TAT AT*A* T, TAT TAG ATA* A*T	Diese Arbeit
Set4	SWGA04, SWGA05	TAA TAA TAT AT*A* T, ATT ATC TAA* T*A	Diese Arbeit
Set5	SWGA01, SWGA04, SWGA06	CTA ATA ATA A*T*A, TAA TAA TAT AT*A* T, TAT TAG ATA* A*T	Diese Arbeit
Set6	SWGA07, SWGA08, SWGA09 SWGA10	ATA TAT AAT AA*T, ATA TGA TAA T*A, TAA TAT AAA TA*A, TAG ATA AAT AA*T	Diese Arbeit
Set7	SWGA11, SWGA12, SWGA13, SWGA14	ATA TTA TTA G*A, ATT ATA TTA TT*A, TAT TAT TAT TA*G, TTA TAT TAT C*A	Diese Arbeit
Set8	SWGA01, SWGA15, SWGA16, SWGA17	CTA ATA ATA AT*A, TAA TAA TAT AA*T, TAA TAG ATA A*T, TAT AAT AGA* T	Diese Arbeit

* Phosphorothioat-Bindung; SWGA Selective Whole Genome Amplification.

2.8 Datenbanken und Programme

Ein Teil der Arbeit wurde auf Standardrechnern mit Windows Betriebssystem durchgeführt. Die meisten Berechnungen erfolgten jedoch auf einem Mehrprozessorsystem mit zwei 8-Kern Intel Xeon CPU ES-2650 Prozessoren (2,6 Hz, 64 GB RAM) und Linux Ubuntu 16.04 LTS Betriebssystem. Einige Analysen erfolgten zusätzlich auf dem High Performance Computing Cluster des regionalen Rechenzentrums der Universität Hamburg, ausgestattet mit jeweils zwei Intel Xeon E5-2630v3 CPUs.

Folgende Programme waren wesentlicher Bestandteil dieser Arbeit (letzter Zugriff: 12. Mai 2019):

Alignmentvisualisierung und -editierung:

Seaview v4.2.3	http://pbil.univ-lyon1.fr/software/seaview (Gouy et al., 2010)
Gblocks v0.91b	http://molevol.cmima.csic.es/castresana/Gblocks.html (Castresana, 2000)
Genedoc v2.7	http://genedoc.software.informer.com/ (Nicholas et al., 1997);
BioEdit v7.0.5	http://www.mbio.ncsu.edu/bioedit/page2.html (Hall, 1999)

Aufnahme und Bearbeitung von Mikroskopiebildern:

Cell Imaging Software <https://www.olympus-sis.com/>
(Olympus, Münster)

Bildbearbeitung:

Adobe Photoshop CS5 <https://www.adobe.com/de/products/photoshop.html>
(Adobe Systems GmbH; München)

CorelDraw X5 <https://www.coreldraw.com/>
(Corel Corporation, München)

ImageJ v1.48 <http://imagej.nih.gov/ij/>
(Schneider et al., 2012)

Daten-Ressource für Gen- und Organismen-Informationen und -Publikationen:

EuPathDB <https://eupathdb.org/eupathdb/>
(Aurrecoechea et al., 2017);

NCBI <http://www.ncbi.nlm.nih.gov/>
(National Center for Biotechnology Information; Bethesda (US))

PlasmoDB v12-34 <http://plasmodb.org/plasmo/>
(Aurrecoechea et al., 2009)

Darstellung und Auswertung von Sanger-Sequenzdaten:

FinchTV v1.4.0 <https://finchtv.software.informer.com/1.4/>
(Geospiza Inc, Seattle (US))

SnapGene v4.1 <http://www.snapgene.com/>
(GSL Biotech LLC; Chicago (US))

Darstellung von Kontaminationen im Assembly:

BlobTools v1.0.1 <https://blobtools.readme.io/docs>
(Laetsch et al., 2017);

***De novo* Assembly, Scaffolding und Mapping:**

BWA v0.7.17 <http://bio-bwa.sourceforge.net/>
(H. Li & Durbin, 2009)

CLC Genomics Workbench v7.5.1 <https://www.qiagenbioinformatics.com/products/clc-genomics-workbench/>
(QIAGEN Bioinformatics; Aarhus (DK))

SMALT v0.7.4 <https://www.sanger.ac.uk/science/tools/smalt-0>
(Genome Research Ltd.; Cambridgeshire (UK))

Velvet v1.2.10 <https://www.ebi.ac.uk/~zerbino/velvet/>
(Zerbino et al., 2008)

Entwurf der SWGA-Primer:

SWGA v0.4.3 <https://github.com/eclarke/swga>
(Clarke et al., 2017)

Entwurf und Überprüfung von Oligonukleotid-Primern:

genefisher2 <https://bibiserv.cebitec.uni-bielefeld.de/genefisher2/>
(Giegerich et al., 1996)

Oligo Analyzer v3.1	https://eu.idtdna.com/calc/analyzer (Integrated DNA Technologies; Illinois (US))
OligoCalc v.327	http://biotools.nubic.northwestern.edu/OligoCalc.html (Kibbe, 2007)
Primer3 v4.1.0	http://primer3.ut.ee/ (Untergasser et al., 2012)

Ermittlung geeigneter Substitutionsmodelle für phylogenetische Analysen:

jModelTest2 v2.1.10	https://github.com/ddarriba/jmodeltest2 (Darriba et al., 2012)
ProtTest3 v3.4.2	https://github.com/ddarriba/prottest3 (Darriba et al., 2011)

Gen-Annotation:

Augustus v3.2.2	http://bioinf.uni-greifswald.de/augustus/ (Stanke et al., 2003)
Companion v1.0.1	https://companion.sanger.ac.uk/ (Steinbiss et al., 2016)
SNAP [2013]	http://korflab.ucdavis.edu/software.html (Korf, 2004)
MAKER2 v2.31.9	http://www.yandell-lab.org/software/maker.html (Cantarel et al., 2007)

Genome Finishing:

GapFiller v1.10	https://www.baseclear.com/services/bioinformatics/basetools/gapfiller/ (Boetzer et al., 2012)
PAGIT v1.64	https://www.sanger.ac.uk/science/tools/pagit/ (Swain et al., 2012)
REAPR v1.0.18	http://www.sanger.ac.uk/science/tools/reapr (Hunt et al., 2013)
RaGOO v1.1	https://github.com/malonge/RaGOO (Alonge et al., 2019)

Grafische Darstellung phylogenetischer Stammbäume:

Treeview v1.6.6	http://taxonomy.zoology.gla.ac.uk/rod/treeview.html (Page, 1996)
FigTree v1.4.3	http://tree.bio.ed.ac.uk/software/figtree/ (Rambaut, 2004)

Identifizierung orthologer Gene und GO-Annotation:

Gene Ontology [2020-1]	http://geneontology.org/ (Ashburner et al., 2000)
OrthoMCL v5	http://orthomcl.org/orthomcl/ (Li, 2003)
PANTHER v14.0	http://pantherdb.org/ (Mi et al., 2019)

Identifizierung und Maskierung von Sequenz-Wiederholungen:

Repbases22.08	https://www.girinst.org/server/RepBase/index.php (Bao et al., 2015)
RepeatMasker v4.0.7	http://www.repeatmasker.org/ (Smit et al., 2013)
TRF v4.09	https://tandem.bu.edu/trf/trf.html (Benson, 1999)

Paarweise Sequenzvergleiche und Visualisierung:

Artemis Comparison Tool v13.0.0	https://www.sanger.ac.uk/science/tools/artemis-comparison-tool-act (Carver et al., 2005)
Exonerate v2.4.0	http://www.ebi.ac.uk/~guy/exonerate/ (Slater et al., 2005)
MUMmer v3.23	http://mummer.sourceforge.net/ (Kurtz et al., 2004)

Phylogenetische Analysen:

DendroPy v4.3.0	http://dendropy.org/ (Sukumaran et al., 2010)
Garli v2.0.1	https://github.com/zwickl/garli/ (Zwickl, 2006)
MrBayes v3.2.6	http://mrbayes.sourceforge.net/ (Huelsenbeck et al., 2001)
PHYMLIP v3.696	http://evolution.genetics.washington.edu/phymlip.html (Felsenstein, 1989)
Phyml	http://atgc.lirmm.fr/phyml/ (Guindon et al., 2005)
RAxML v8.2.10	https://sco.h-its.org/exelixis/software.html (Stamatakis, 2014)

Protein- und CDS-Analyse

InterProScan v.5.18-57.0	https://github.com/ebi-pf-team/interproscan (Jones et al., 2014)
MitoProt II 1.101	ftp://ftp.biologie.ens.fr/pub/molbio (Claros et al., 1996)
PATS [30/04/2018]	http://modlabcadd.ethz.ch/software/pats/ (Zuegge et al., 2001)
Phobius v1.01	http://phobius.sbc.su.se/ (Käll et al., 2004)
SignalP v4.1	http://www.cbs.dtu.dk/services/SignalP-4.1/ (Petersen et al., 2011)
TMHMM v2.0	http://www.cbs.dtu.dk/services/TMHMM/ (Sonnhammer et al., 1998)

Multiple Sequenzalignments:

MAFFT v7.3.13 <https://mafft.cbrc.jp/alignment/software/>
(Katoh, 2002)

Qualitätsbeurteilung genomischer Sequenzen

BUSCO v3.0.2 <https://busco.ezlab.org/>
(Waterhouse et al., 2018)

Qualitätskontrolle von Illumina-Reads:

FastQC v0.11.5 <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
(Andrews, 2010)

Quantifizierung von genomischer DNA:

7500 SDS v1.4.1 <https://www.thermofisher.com/de/de/home/technical-resources/software-downloads/applied-biosystems-7500-real-time-pcr-system.html>
Thermo Fisher Scientific, St. Leon-Rot;

Suche homologer Sequenzdaten:

GHOSTX v1.3.7 <http://www.bi.cs.titech.ac.jp/ghostx/>
(Suzuki et al., 2014)

HMMER v3.1b1 <http://hmmer.org/>
(Eddy, 2011)

ncbi-blast v2.3.0 <https://blast.ncbi.nlm.nih.gov/Blast.cgi/>
(Altschul et al., 1990)

Text- und Tabellenverarbeitung:

Adobe Acrobat X Pro <https://www.adobe.com/>
(Adobe Systems GmbH, München)

LibreOffice 5 <https://de.libreoffice.org/>
(The Document Foundation, Berlin)

Microsoft Office 2016 <https://www.microsoft.com/>
(Microsoft Deutschland GmbH, Unterschleißheim)

tRNA und rRNA-Gen Vorhersage:

Aragorn v1.2.38 <http://mbio-serv2.mbioekol.lu.se/ARAGORN/>
(Laslett, 2004)

RNAmmer v1.2 <http://www.cbs.dtu.dk/services/RNAmmer/>
(Lagesen et al., 2007)

tRNAscan-SE v2.0 <http://lowelab.ucsc.edu/tRNAscan-SE/>
(Lowe et al., 2016)

Übersetzung von DNA-Sequenzen in Proteinsequenzen:

ExPASy Translate Tool <http://web.expasy.org/translate/>
(Gasteiger, 2003)

Verarbeitung und Trimming der Mate-Pair Reads:

NxTrim v0.4.2 <https://github.com/sequencing/NxTrim/>
(O'Connell et al., 2015)

Verarbeitung von NGS-Datensätzen:

BamTools v2.4.0	https://github.com/pezmaster31/bamtools/ (Barnett et al., 2011)
Fastx-Toolkit v0.0.14	http://hannonlab.cshl.edu/fastx_toolkit/ (Gordon et al., 2010)
Galaxy v18.05	https://usegalaxy.org/ (Afgan et al., 2016)
SAMtools v1.3.1	http://samtools.sourceforge.net (Li et al., 2009)
SRAtoolkit v.2.6.2	https://www.ncbi.nlm.nih.gov/sra/docs/toolkitsoft/ (Team, 2015)

Andere vorausgesetzte Pakete und Programme:

Nextflow	https://www.nextflow.io/index.html (Centre for Genomic Regulation (CRG); Barcelona (SP))
perl v5.22.1	https://www.perl.org/get.html (L. Wall, Perl porter; Sebastopol, California (US))
python v2.7.12	https://www.python.org/ (Python Software Foundation; Wilmington (US))
R v3.2.3	https://cran.r-project.org/ (The R Foundation for Statistical Computing; Wien (AT))
RStudio v1.0	https://www.rstudio.com/ (RStudio; Boston (US))
ruby v2.3.1	https://www.ruby-lang.org/de/ (Y. Matsumoto, Ruby MRI; Heroku (JP))

3 Methoden

3.1 Gewinnung von Untersuchungsmaterial

Der Fang der Fledermäuse erfolgte in Kooperation mit dem Fledermauszentrum Noctalis an der Kalkberghöhle in Bad Segeberg. Die Kalkberghöhle ist eines der größten Fledermauswinterquartiere in Nordeuropa, in dem jährlich über 20.000 Fledermäuse acht verschiedener Arten überwintern. Durchgeführt wurde der Netzfang zur Hauptschwärmphase der Fledermäuse am 05. Oktober 2015. Zum Fang wurde ein feinmaschiges Stellnetz (Japannetz) am Noctalis-Gebäude in den Flugkorridor zu den drei Haupteinflugsöffnungen der Höhle aufgestellt. Die Netzlänge betrug 10 m, bei einer Höhe von ca. 6,5 m. Gefangene Tiere der Gattung *Myotis* wurden nach der Befreiung aus dem Netz für einige Minuten in Stoffsäckchen aufbewahrt. Anschließend wurden Art und Geschlecht bestimmt und die biometrischen Daten der Fledermaus (Körpermaße, Unterarmlänge und Gewicht) erfasst. Zusätzlich wurden die Tiere auf Ektoparasiten untersucht. Dafür wurden die Flughäute entfaltet sowie das Fell durchgepusht und die Ektoparasiten mit einer Pinzette in 70%igen (v/v) Alkohol überführt. Um ein austrocknen zu verhindern, wurden Fledermausfliegen separat in einem Eppendorfgefäß mit angefeuchteter Watte gesammelt und bis zur Mitteldarmpräparation (vgl. 2.3.1) bei 4 °C Raumtemperatur gelagert. Eine Blutabnahme erfolgte nur bei den gefangenen Wasserfledermäusen. Dafür wurde die Flügelvene mit einer sterilen 22 G Einmal-Kanüle punktiert. Der heraustretende Blutstropfen (zwischen 5 µl und 30 µl) wurde direkt in 15 µl einer isotonen Natriumcitratlösung überführt und wie in Abschnitt 2.3.2 dargestellt weiterbearbeitet. Mit jeweils einem Tropfen Vollblut wurden Blutausrichie wie in 2.2 beschrieben hergestellt. Die Fledermäuse wurden nach der Blutabnahme wieder an der Fangstelle freigelassen.

3.2 Anfertigung von Blutausrichien und Bestimmung der Parasitämie

Zur Bestimmung der Parasitämie im Blut der Fledermäuse wurden dünne Blutausrichie auf Objektträgern angefertigt. Dafür wurde jeweils ein Tropfen Vollblut auf einen fettfreien Objektträger aufgetragen und mithilfe eines zweiten Objektträgers gleichmäßig ausgestrichen. Die Ausrichie wurden anschließend an der Luft getrocknet und für 1 min in absolutem Methanol fixiert. Die Färbung der Ausrichie erfolgte für eine Stunde in Giemsa-Färbelösung bei Raumtemperatur. Im Anschluss wurden die Objektträger vorsichtig mit Wasser abgespült und getrocknet.

Eine erste Einschätzung der Parasitämie erfolgte durch die lichtmikroskopische Untersuchung der Blutausrichie bei 600-facher Vergrößerung. In jedem Blutausrich wurden dafür dreimal 100 zufällig ausgewählte Blickfelder durchgemustert und infizierte Erythrozyten ausgezählt. Lediglich Blutausrichie, in der mindestens ein Parasit gezählt wurde, wurden bei 1000-facher Vergrößerung ausgewertet. Die Bestimmung der Parasitämie erfolgte als prozentualen Anteil der infizierten

Erythrozyten an der Gesamterythrozytenzahl. Ausgezählt wurden jeweils 30.000 Erythrozyten bei 1000-facher Vergrößerung.

3.3 Anreicherung von Oozysten, Mikrogameten und Gametozyten

3.3.1 Mitteldarmpräparation der Fledermausfliegen

Zur Isolierung von Oozysten und eingekapselten Ookineten wurde die Mitteldärme der Fledermausfliegen herauspräpariert. Dafür wurden die Fledermausfliegen durch einen Kälteschock bei -20 °C für 3 min betäubt und für 10 s in 70 % (v/v) Ethanol getaucht. Die Präparation erfolgte unter dem Binokular bei 100-facher Vergrößerung in einem Tropfen 1x PBS-Puffer. Die Fledermausfliegen wurden am Thorax mit einer sterilen Seziernadel fixiert. Mit einer zweiten Nadel wurde das Abdomen vorsichtig vom Thorax getrennt und mit einer feinen Sezierpinzette langsam abgezogen, bis der Verdauungstrakt frei lag. Der am Thorax verbleibende Mitteldarm wurde mit einer Nadel von den anderen Organen der Fledermausfliege abgeschnitten, in 30 µl SET-Puffer überführt und bis zum weiteren Gebrauch bei -20 °C gelagert. Alle anderen akzessorischen Gewebe und Trümmer wurden zum Nachweis von Parasiten wie in 3.4.1 beschrieben homogenisiert.

3.3.2 *In vitro* Exflagellation

Die Isolierung von Mikrogameten erfolgte nach der Methode von Palinauskas et al. und basiert im Prinzip auf der Exflagellation von Mikrogametozyten und einem anschließenden Zentrifugationsschritt, der entstandene Mikrogameten von Blutzellen trennt (Palinauskas et al., 2013). Für die *in vitro* Exflagellation wurde das Natriumcitrat-Blut noch im Feld für 4 min bis 15 min an der Luft exponiert. Die Umgebungstemperatur lag bei 9 ± 1 °C. Anschließend wurden die Proben für 5 min bei 7.000 x g zentrifugiert. Der Überstand wurde jeweils in 75 µl SET-Puffer überführt und bis zur DNA-Isolierung (vgl. 3.4.2) bei -20 °C eingefroren. Aus dem verbleibenden Zentrifugat wurden Gametozyten angereichert (vgl. 3.3.3).

3.3.3 Dichtegradientenzentrifugation in Ficoll

Zur Anreicherung von Gametozyten aus dem Fledermausblut wurde eine Dichtegradientenzentrifugation mit Ficoll durchgeführt. Diese Methode ermöglicht eine Trennung aller Blutbestandteile nach ihrer Zelldichte. Zusätzlich bewirkt Ficoll die Agglutination von Erythrozyten und führt damit zu einer Sedimentation der Zellen unter dem Trennmittel. Für die Dichtegradientenzentrifugation wurde das Volumen der Blutproben nach der *in vitro* Exflagellation (vgl. 3.3.2) mit DPBS-Puffer (ThermoFischer Scientific, St. Leon-Rot) auf 300 µl aufgefüllt. Die Zellsuspension wurde vorsichtig über 450 µl vorgelegter Ficolllösung geschichtet und 30 min lang bei 600 x g und 20 °C zentrifugiert. Der Überstand wurde samt Ficollfraktion abgenommen und

bis zum weiteren Gebrauch bei -20 °C gelagert. Das Erythrozytenpellet wurde mit 50 µl DPBS-Puffer gewaschen und bei 120 x g und 20 °C für 10 min zentrifugiert. Im Anschluss wurden die Zellen mit 15 µL SET-Puffer versetzt und bis zur DNA-Isolierung (vgl. 3.4.3) bei 4 °C gelagert.

3.4 Isolierung genomischer DNA

3.4.1 Isolierung genomischer DNA aus Fledermausfliegen

Für die Isolierung von genomischer DNA aus den soliden Geweben der Fledermausfliegen wurden die Proben zunächst homogenisiert. Dafür wurden die Gewebe in flüssigem Stickstoff eingefroren und mit dem Precellys Keramik-Kit 1,4 mm (Peqlab Biotechnologies, Erlangen) auf höchster Intensität aufgeschlossen (dreimal 30 s in Intervallen mit mindestens 15 s Pause auf Eis). Die Proben wurden mit 180 µl PBS-Puffer und erneut dreimal bei höchster Intensität aufgeschlossen. Die homogenisierten Gewebe wurden mit 200 µl ATL-Puffer (Qiagen, Hilden) und 20 µl Proteinase K versetzt und über Nacht bei 56 °C inkubiert. Zur Aufreinigung der DNA wurde die Probe anschließend über Phenol/Chloroform extrahiert (vgl. 3.8.5) und über Ethanol gewaschen (vgl. 3.8.6). Die Probe wurde dabei vollständig getrocknet und in 100 µl nukleasefreiem Wasser resuspendiert. Die Lagerung erfolgte bei 20 °C.

3.4.2 Isolierung genomischer DNA aus Oozysten und Mikrogameten

Sowohl die herauspräparierten Mitteldärme der Fledermausfliegen als auch die Proben nach der In-vitro Exflagellation wurden zur DNA-Isolierung mit jeweils 250 µl Lysepuffer und 20 µl Proteinase K versetzt und über Nacht bei 56 °C inkubiert. Zusätzlich erfolgte ein RNA-Verdau. Dafür wurden die Proben jeweils mit 4 µl RNase A versetzt und 10 min bei Raumtemperatur inkubiert. Die Aufreinigung erfolgte anschließend über Phenol/Chloroform und über Ethanol (vgl. 3.8.5 und 3.8.6). Die getrockneten Pellets wurden in je 20 µl HPLC-H₂O aufgenommen (20 min bei 50 °C) und bis zum weiteren Gebrauch bei -20°C gelagert.

3.4.3 Isolierung genomischer DNA aus Blut

Zur Isolierung der genomischen DNA (gDNA) aus Gametozyten des Parasiten oder Blutzellen des Wirtes wurde das GeneJET Genomic DNA Purification Kit der Firma ThermoFischer Scientific (St. Leon-Rot) verwendet. Die Extraktion erfolgte dabei nach dem Herstellerprotokoll für Blutzellen aus Säugetieren. Um eine vollständige Wirkung der Proteinase K zu gewährleisten, wurde die Inkubationszeit des Lysates allerdings auf 1 h verlängert. Die Elution erfolgte in zwei Schritten mit jeweils 50 µl nukleasefreiem Wasser. Die Eluate wurden anschließend mit der Vakuumzentrifuge auf etwa 30 µl eingengt und bis zur weiteren Verwendung bei -20 °C gelagert.

3.5 Amplifizierung der genomischen DNA

Aufgrund der niedrigen Konzentrationen wurde die gDNA mittels Whole Genome Amplifizierung (WGA) vervielfältigt. Die in dieser Arbeit durchgeführte WGA basiert auf der isothermalen Multiple Displacement Amplifizierungsmethode (MDA). Das Funktionsprinzip der MDA beruht auf der kontinuierlichen Strangverdrängungsaktivität der ϕ 29-DNA-Polymerase, durch die fortlaufend neue *Template*-Bereiche für Zufallsoligomere entstehen (Abb. 3.1).

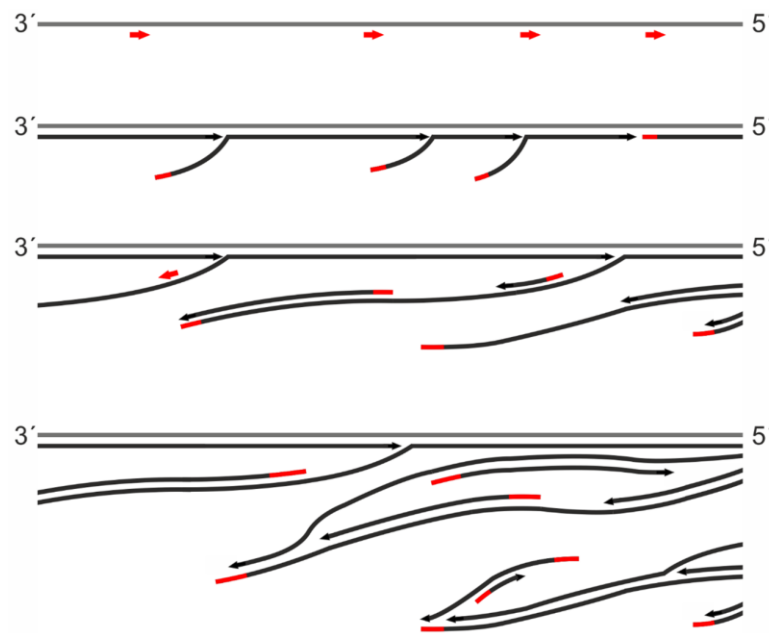


Abb. 3.1: Schematische Darstellung der Whole Genome Amplifizierung.

Genomische Einzelstrang-DNA (graue Linie) dient als Template zum Binden von Zufallsoligomeren (rote Pfeile). Die ϕ 29-DNA-Polymerase lagert sich wiederum an die Zufallsoligomere und synthetisiert einen komplementären Strang in 5'→3'-Richtung (schwarze Linie). Sobald die Polymerase während der Replikation auf einen doppelsträngigen Bereich trifft, wird das 5'-Ende des vorderen Strangs durch die Strangverdrängungsaktivität der Polymerase vom ursprünglichen Template abgelöst. Abgelöste DNA-Stränge liegen als Einzelstrang vor und dienen anschließend weiteren Zufallsoligomeren als Bindungsstelle. Verändert nach (Lage et al., 2003).

Für die WGA-Reaktion wurde das kommerzielle Illustr GenomiPhi V2 DNA Amplification Kit (GE Healthcare, Solingen) nach Angaben des Herstellers verwendet. Dafür wurde jeweils 1 μ l der isolierten gDNA mit 9 μ l Probenpuffer versetzt, 3 min bei 95 °C denaturiert und anschließend 1 min auf Eis gestellt. Die Probe wurde mit 9 μ l Reaktionspuffer und 1 μ l Enzym versetzt bei 30 °C amplifiziert und zur Inaktivierung der Polymerase für 10 min bei 65 °C inkubiert. Um mögliche Laborkontaminationen zu vermeiden erfolgte die Amplifizierung bei 30 °C für maximal 1,5 h. Zusätzlich wurden alle Schritte mit Filterpipettenspitzen und autoklavierten Reaktionsgefäßen unter sterilen Bedingungen durchgeführt. Zusätzlich wurde eine Negativkontrolle mit sterilem Wasser als Template mitgeführt, um unspezifische Amplifizierungen auszuschließen. Die WGA-Proben wurden anschließend in 0,7%igen Agarosegelen überprüft (vgl. 3.8.2) über Phenol/Chloroform (vgl. 3.8.5) extrahiert und bis zur weiteren Verwendung bei -20°C gelagert.

3.6 Selektive Amplifizierung der genomischen DNA

Für die selektive Amplifizierung der gDNA (Selective Whole Genome Amplification, SWGA) wurden Oligonukleotide generiert, die spezifisch die DNA des Parasiten amplifizieren (vgl. 3.11.4). Durchgeführt wurde die SWGA-Reaktion mit der phi29 DNA Polymerase (New England BioLabs, Frankfurt am Main) in einem 25 µl Ansatz. Der Ablauf und die Zusammensetzung der Reaktion sind in Tab. 3.1 zusammengefasst. Verwendete Primersets sind in Tab. 2.4 aufgeführt.

Tab. 3.1: Zusammensetzung und Temperaturprofil der SWGA-Reaktion.

SWGA-Ansatz		SWGA-Programm	
1 µl	gDNA	35 °C	10 min
1,75 µl	Primerset (50 µM)	34 °C	10 min
2,5 µl	10x phi29 Puffer	33 °C	10 min
0,25 µl	100x BSA	32 °C	10 min
2,5 µl	dNTP-Mix (je 10 mM)	31 °C	10 min
1,5 µl	phi29 DNA Polymerase (10 U/µl)	30 °C	16 h
15,5 µl	HPLC-H ₂ O	65 °C	15 min
		4 °C	∞

Die SWGA-Produkte wurden nach der Amplifizierung jeweils mit 25 µl HPLC-H₂O verdünnt, um die Viskosität der Lösung zu verringern. Anschließend wurden Konzentration und Qualität photometrisch und mittels Gelelektrophorese bestimmt (vgl. 3.8.2). Die Proben wurden über Phenol/Chloroform extrahiert (vgl. 3.8.5) und bis zur weiteren Verwendung -20 °C gelagert.

3.6.1 Komplementärstrangsynthese

SWGA-Produkte enthalten im Vergleich zu WGA-Produkten einen deutlich erhöhten Anteil einzelsträngiger DNA (ssDNA). Vor dem weiteren Gebrauch der Proben wurden deshalb komplementäre DNA-Stränge synthetisiert. Dafür wurde die DNA nach der Phenol/Chloroform-Extraktion und in 28 µl HPLC-H₂O gelöst und mit der LongAmp Taq Polymerase (New England BioLabs, Frankfurt am Main) und dem Enzym aus dem AccuPrime Taq DNA Polymerase System (Thermo Fisher Scientific, St. Leon-Rot) wie in Tab. 3.2 beschrieben, behandelt.

Tab. 3.2: Zusammensetzung und Temperaturprofil der Komplementärstrangsynthese.

Ansatz		Programm	
28 µl	gDNA	95 °C	2 min
10 µl	LongAmp Puffer	35 °C	30 sek
2,5 µl	Random Primer	65 °C	5 min
5 µl	dNTP-Mix (je 10 mM)	65 °C	10 min
2,5 µl	LongAmp Taq Polymerase	10 °C	∞
2 µl	Accuprime Taq		

5 x

Die Proben wurden anschließend über Phenol/Chloroform extrahiert (vgl. 3.8.5) und mit 1 x T4 DNA Ligase Buffer, 50 µM dNTPs und 250 U T4 DNA Ligase (New England BioLabs, Frankfurt am Main) in einem Volumen von 75 µl bei 25 °C inkubiert. Nach 30 min wurde der Ansatz mit 3,75 U T4 DNA Polymerase (New England BioLabs, Frankfurt am Main) versetzt und für 5 min

bei 37 °C inkubiert. Anschließend wurde die Reaktion zur Inaktivierung der Enzyme 10 min bei 75 °C erhitzt, über Ethanol präzipitiert (vgl. 3.8.6) und bis zur weiteren Verwendung bei 4 °C gelagert.

3.7 PCR-basierte Methoden

Eine Amplifizierung spezifischer DNA-Abschnitte erfolgte durch die Polymerasekettenreaktion (*Polymerase Chain Reaction*, PCR). Die dafür verwendeten Primer sind in Tab. 2.1 aufgeführt. Für die PCR wurden Enzyme und zugehörige Puffersysteme von New England BioLabs (Frankfurt am Main), Promega (Mannheim) und Thermo Fisher Scientific (St. Leon-Rot) verwendet. Ein Standard-PCR-Ansatz wurde abhängig von der verwendeten Polymerase in einem Endvolumen von 20 µl oder 25 µl durchgeführt (Tab. 3.3).

Tab. 3.3: Zusammensetzung und Temperaturprofil der Standard-PCR.

Polymerase		PCR-Ansatz		PCR-Programm	
AccuPrime Taq Polymerase System (Thermo Fisher Scientific, St. Leon-Rot)	DNA	2,5 µl	Puffer (I bis 200 bp, II von 0,2-4kb)	94 °C	2 min
		0,5 µl	forward Primer (10 µM)	40 x	94 °C 30 s
		0,5 µl	reverse Primer (10 µM)		X °C 30 s
		0,5 µl	AccuPrime Taq		66 °C 1 min/kb
		1 µl	DNA-Template		66 °C 5 min
		20 µl	HPLC-H ₂ O		12 °C ∞
Phusion High-Fidelity Polymerase (New England BioLabs, Frankfurt am Main)	DNA	5 µl	HF Puffer	98 °C	1 min
		0,25 µl	MgCl ₂	98 °C	10 s
		1 µl	forward Primer (10 µM)	X °C	30 s
		1 µl	reverse Primer (10 µM)	72 °C	30 s/kb
		0,25 µl	Phusion Taq	72 °C	5 min
		1 µl	DNA-Template	12 °C	∞
Pfu DNA Polymerase (Promega, Mannheim)		2,5 µl	10 x Puffer	95 °C	2 min
		1 µl	dNTP-Mix (je 10 mM)	95 °C	30 s
		0,5 µl	forward Primer (10 µM)	X °C	30 s
		0,5 µl	reverse Primer (10 µM)	68 °C	2 min/kb
		0,5 µl	Pfu DNA Polymerase	68 °C	10 min
		1 µl	DNA-Template	12 °C	∞
		20 µl	HPLC-H ₂ O	35 x	

Die Annealingtemperatur (X °C) richtet sich nach der Schmelztemperatur und ist abhängig von der Länge und Sequenz der Primer.

Die Anforderungen der entsprechenden Primer wurden jeweils aus der Literatur entnommen. Bei selbst entworfenen Primern wurde die Annealingtemperatur an die Schmelztemperatur angepasst. Die Elongationszeit wurde nach der erwarteten Größe des zu amplifizierenden DNA-Fragments verändert. In der Regel wurde in der PCR 1 µl der isolierten gDNA als Template verwendet. SWGA- und WGA-Produkte wurden entsprechend verdünnt. Zur Kontrolle von Kontaminationen mit Fremd-DNA wurde bei jeder durchgeführten PCR eine Negativkontrolle ohne DNA-Template mitgeführt. Ein Nachweis der Parasiten erfolgte mit einer zusätzlichen Positivkontrolle, die aufgereinigte gDNA des Parasiten *P. falciparum* enthielt. Nach der PCR wurden die entstandenen Amplifikate mittels Agarosegelelektrophorese überprüft (vgl. 3.8.2), aufgereinigt (vgl. 3.8.3), gegebenenfalls kloniert (vgl. 2.10) und sequenziert (vgl. 2.11.1).

3.7.1 Nested-PCR

Zur Erhöhung der Sensivität wurde bei den meisten Genen, wenn gDNA als Template verwendet wurde, eine Nested-PCR durchgeführt. Die Nested-PCR zeichnet sich dadurch aus, dass das Produkt einer Amplifizierung als Template bei einer zweiten PCR dient. Das Primerpaar der zweiten PCR bindet dabei an Sequenzbereiche innerhalb des ersten PCR-Produkts, so dass der gewünschte Sequenzbereich mit sehr hoher Spezifität generiert wird. Der verwendete Reaktionsansatz und die PCR-Bedingungen waren für beide PCRs weitgehend identisch und unterschieden sich nicht von der konventionellen PCR (vgl. 2.8). Bei der zweiten PCR wurde jedoch lediglich 0,5 µl des ersten PCR-Produkts als Template verwendet. Für ein Gesamtvolumen von 20 µl bzw. 25 µl wurde entsprechend mehr Wasser hinzugefügt. Die ursprünglichen Negativ- und Positivkontrollen wurden bei der zweiten PCR mitgeführt.

3.7.2 Stepdown-PCR

Eine Stepdown-PCR wurde angewandt, wenn eine Anpassung der Annealingtemperatur in der konventionellen oder in der Nested-PCR nicht zur Synthese des gewünschten Fragments führte. Dabei wurde die Annealingtemperatur so programmiert, dass sie im Verlauf der PCR allmählich auf eine ausreichend niedrige Temperatur verringert wurde. In der Regel wurde die Temperatur dafür, zunächst ausgehend von der Schmelztemperatur der Primer, zu Beginn des PCR-Programms alle zwei Zyklen schrittweise um 2 °C reduziert. Nach zehn Zyklen wurden 25 weitere Zyklen mit der bis dahin geringsten Annealingtemperatur durchgeführt.

3.7.3 Banden-Reamplifizierung

Bei schwachen Bandenintensitäten im Agarosegel aber auch bei leichten Kontaminationen mit unspezifischen PCR-Produkten wurde die Menge der PCR-Fragmente durch eine Banden-Reamplifizierung erhöht. Dafür wurde die PCR-Fragmente im Agarosegel aufgetrennt und die entsprechende Bande im Gel durch mehrfache Einstiche mit einer Kanüle aufgenommen. Die DNA wurde anschließend in den vorbereiteten PCR-Reaktionsansatz überführt. Die Proben wurden unter sonst identischen Bedingungen wie die erste, ursprüngliche PCR amplifiziert. Lediglich die Zyklenzahl während der Reamplifizierung wurde auf maximal 20 reduziert.

3.7.4 Kolonie-PCR

Mithilfe der Kolonie-PCR wurden positive rekombinante Klone ohne vorherige DNA-Präparation identifiziert. Dafür wurden die zu analysierenden *E. coli*-Zellen als Bakterienkolonie von der entsprechenden Agarplatte gepickt und direkt in den vorbereiteten PCR-Ansatz überführt. Dieser setzte sich unter Verwendung der *Taq*-Polymerase von Molzym (Bremen) wie in Tab. 3.4 beschrieben zusammen. Die Amplifizierung erfolgte jeweils mit einem vektor-spezifischen und einem Insert-spezifischen Primer. Nach der PCR wurden die Amplifikate mittels

Agarosegelelektrophorese analysiert (vgl. 3.8.2). Positive Einzelkolonien wurden anschließend in Übernachtskulturen angereichert (vgl. 2.10.2).

Tab. 3.4: Zusammensetzung und Temperaturprofil der Kolonie-PCR.

Polymerase	PCR-Ansatz		PCR-Programm	
MolTaq (Molzym, Bremen)	1 µl	10 x Puffer	94 °C	3 min
	0,2 µl	dNTP-Mix (je 10 mM)	94 °C	30 s
	0,2 µl	MgCl ₂	52 °C	30 s
	0,25 µl	forward Primer (10 µM)	72 °C	1 min/kb
	0,25 µl	reverse Primer (10 µM)	72 °C	10 min
	0,15 µl	Taq DNA Polymerase	12 °C	∞
	7,95 µl	HPLC-H ₂ O		

3.7.5 Quantitative Real-Time-PCR

Die quantitative Real-Time PCR (qPCR) ist eine Methode zur Quantifizierung von Nukleinsäuren und wird in der Regel dazu genutzt die Expressionen von Transkripten spezifischer Zielgene nachzuweisen. Im Rahmen dieser Arbeit wurde die qPCR verwendet, um das Mengenverhältnis der Parasiten-DNA zur Wirts-DNA zu bestimmen. Dabei fand der Maxima SYBR Green qPCR Master Mix von Fermentas (St. Leon-Rot) Anwendung, der bereits Puffer, DNA-Polymerase, dNTPs und den SYBR Green- Fluoreszenzfarbstoff beinhaltet. Die Zusammensetzung der Reaktionen erfolgten wie in Tab. 3.5 beschrieben jeweils in einem Gesamtvolumen von 10 µl. Zur Kontrolle der Pipettiergenauigkeit wurde stets mit einem Mastermix gearbeitet, der alle PCR-Reagenzien außer dem Template enthält. In der Regel wurde jede Probe in Triplikaten angesetzt. Zusätzlich wurden mindestens zwei Negativkontrollen ohne DNA-Template mitgeführt.

Tab. 3.5: Zusammensetzung und Temperaturprofil der quantitativen Real-Time PCR.

Polymerase			qPCR-Ansatz		qPCR-Programm	
Maxima SYBR Green qPCR Master Mix (Fermentas, St. Leon-Rot)			5 µl	2 x PCR Master Mix	95 °C	10 min
			0,25 µl	forward Primer (10 µM)	95 °C	15 s
			0,25 µl	reverse Primer (10 µM)	58 °C	15 s
			1 µl	DNA-Template	72 °C	45 s
			3,5 µl	HPLC-H ₂ O	95 °C	15 s
					58 °C	1 min
					95 °C	30 s

Zur Durchführung der qPCR wurde das 7500 Real-Time PCR System (Thermo Fisher Scientific, St. Leon-Rot) verwendet. Die Detektion der Fluoreszenz erfolgte von der systemeigenen Software während der Elongationsphasen. Am Ende der Reaktion wurde eine Schmelzkurvenanalyse durchgeführt, um unspezifische Amplifikate zu identifizieren. Der zur Quantifizierung nötige C_T-Wert (Cycle Threshold) wurde automatisch von der 7500 Software bestimmt. Dieser beschreibt die Zykluszahl einer Reaktion, bei der das erstandene Fluoreszenzsignal über einen vom System festgelegten Schwellenwert in der exponentiellen Phase der PCR steigt. Je niedriger der C_T-Wert, desto mehr Kopien der Zielsequenz lagen in der ursprünglichen Probe vor. Proben mit C_T-Werten über 38 wurden von der weiteren Analyse ausgeschlossen, da in diesen Bereichen keine zuverlässige Quantifizierung mehr möglich ist. Aus den technischen Replikaten wurde für die Quantifizierung

jeweils der Mittelwert der C_T -Werte gebildet. Dabei wurde eine Differenz von 0,3 als maximale Schwankung akzeptiert. Im Rahmen dieser Arbeit wurden zwei Quantifizierungsstrategien angewendet. Bei den WGA-Proben wurden mit Hilfe von Standards jeweils die absoluten Mengen des *sf3b1*-Gens aus Wasserfledermaus und *Polychromobilus* im ursprünglichen Template bestimmt. Bei SWGA-Proben erfolgte die Quantifizierung dagegen relativ, indem die C_T -Werte von jeweils drei Genen aus Wasserfledermaus und *Polychromobilus* gegenübergestellt wurden.

3.7.5.1 Absolute Quantifizierung mittels Standardkurve

Voraussetzung für die absolute Quantifizierung von Parasiten-DNA aus den infizierten Fledermausproben ist das Mitführen eines Standards. In der Regel werden dafür rekombinante Plasmide hergestellt, welche die gleichen Sequenzen wie die des zu untersuchenden Gens aufweisen. Im Rahmen dieser Arbeit erfolgte die absolute Quantifizierung über das *sf3b1*-Gen. Dafür wurde die entsprechenden Gene der Fledermaus und des Parasiten jeweils amplifiziert (vgl. 2.7. und 3.7) und in den Vektor pGEM-T (Promega, Mannheim) kloniert (vgl. 3.9). Ausgehend von der DNA-Konzentration der anschließend aufgereinigten Plasmide sowie der Sequenzlänge des Plasmids und des Integrats wurde die Kopienzahl des Gens pro μl wie folgt bestimmt:

$$\text{Kopienzahl pro } \mu\text{l} = \frac{\text{DNA-Konzentration [ng/}\mu\text{l}]}{\text{Sequenzlänge [bp]} \cdot 660} \cdot 6,022 \cdot 10^{14}$$

Mit der errechneten Kopienzahl pro μl konnte anschließend eine dekadische Verdünnungsreihe von 10^8 bis 10^3 angesetzt werden. Die Verdünnung erfolgte in HPLC- H_2O . Für jede Verdünnungsstufe wurden bei der qPCR Doppelbestimmung durchgeführt. Die sich ergebenden C_T -Werte dieser Ansätze wurden zur Erstellung einer Standardkurve genutzt. Dafür wurden die C_T -Werte gegen den dekadischen Logarithmus der Verdünnungsstufen aufgetragen und eine lineare Regressionsgerade bestimmt. Als Maß für die Pipettiergenauigkeit wurde der Korrelationskoeffizient R^2 der Regressionsgeraden bestimmt. Eine weitere Qualitätskontrolle erfolgte durch die Berechnung der Effizienz E der qPCR, die sich aus der Steigung m der Regressionsgeraden ergibt:

$$E = 10^{-1/m} - 1$$

Bei der Auswertung der qPCR sollten der Wert des Korrelationskoeffizienten größer als 0,99 sein und eine Steigung zwischen -3,2 und -3,6 annehmen, was einer Effizienz von etwa 95-105 % entspricht. Bei abweichenden Werten wurde die qPCR wiederholt und gegebenenfalls optimiert.

Die lineare Regression der Standardkurve diente anschließend als Basis für die Quantifizierung der jeweiligen Kopienzahl des *sf3b1*-Gens der Fledermaus und des *sf3b1*-Gens des Parasiten in Proben aus drei unabhängigen WGA-Reaktionen. Die statistische Auswertung der Ergebnisse erfolgte in Microsoft Excel. Zur graphischen Darstellung wurden Mittelwerte und Standardabweichung berechnet. Die Ermittlung der statistischen Signifikanz erfolgte durch einen zweiseitigen, ungepaarten t-Test unter Annahme gleicher Varianzen. Eine Irrtumswahrscheinlichkeit von kleiner als 5 % ($p < 0,05$) bzw. 1 % ($p < 0,01$) wurde dabei als statistisch signifikant angesehen.

3.7.5.2 Relative Quantifizierung

Bei der Quantifizierung des DNA-Gehalts von *Polychromophilus* und *Myotis* in SWGA-Proben wurde auf die Herstellung von Standardreihen verzichtet. Stattdessen erfolgte die Quantifizierung über die Δ -C_T-Methode. Im Vergleich zur absoluten Quantifizierung wird dabei nicht die Ausgangsmenge eines Gens bestimmt, sondern die Amplifizierung des zu untersuchenden Gens wird auf ein Referenzgen oder auf eine Kontrollprobe bezogen. In dieser Arbeit wurden die dafür die C_T-Werte von jeweils drei Genen aus Wasserfledermaus (*sf3b1*, *prlr*, *pdcd6*) und Parasit (*sf3b1*, *adh* und einem konservierten Gen, unbekannter Funktion) gegenübergestellt (vgl. Tab. 2.3). Die Amplifizierungen erfolgten jeweils in Duplikaten aus jeweils drei unabhängigen SWGA-Reaktionen. Als Kalibrator wurden zusätzlich WGA-Proben mitgeführt. Die Differenz der C_T-Werte zwischen SWGA-Probe und WGA-Probe (Δ -C_T-Wert) gibt Aufschluss über eine Änderung des relativen DNA-Gehalts des Parasiten bzw. der Fledermaus. Berechnen lässt sich diese Änderung wie folgt:

$$\text{Relative Änderung des DNA-Gehalts} = 2^{\Delta C_T}$$

Ausgegangen wird dabei von einer 100%igen Effizienz, bei dem sich der DNA-Gehalt mit jedem Replikationszyklus verdoppelt. Mittelwerte und Standardabweichungen der relativen Änderungen wurden in Microsoft Excel bestimmt und graphisch gegeneinander aufgetragen. Für statistische Analysen wurde der zweiseitige t-Test für abhängige Stichproben verwendet. Festgelegt wurde ein Signifikanzniveau von 5 % ($p < 0,05$) bzw. 1% ($p < 0,01$).

3.8 Analyse und Aufreinigung von Nukleinsäuren

3.8.1 Konzentrationsbestimmung von Nukleinsäuren

Die Konzentration von DNA-Lösungen wurde über die optische Dichte (OD) mit Hilfe des NanoDrop Spektralphotometers (Peglab, Erlangen) bestimmt. Für die Messung wurden bis zu 2 μ l der unverdünnten Nukleinsäurelösung auf die Messeinheit pipettiert und die Extinktion bei 260 nm und 280 nm gemessen. Liegt bei 260 nm eine optische Dichte von 1,0 vor, entspricht dies einer Konzentration von 50 ng/ μ l dsDNA beziehungsweise 33 ng/ μ l ssDNA. Das Verhältnis der OD von 260 nm zu 280 nm (OD₂₆₀/OD₂₈₀) ist ein Indikator für die Qualität der DNA und sollte bei einer reinen Probe zwischen 1,8 und 2,0 liegen.

Zusätzlich erfolgte eine Konzentrationsbestimmung mit dem Qubit 2.0 Fluorometer (Invitrogen, Karlsruhe). Bei erwarteten DNA-Konzentrationen von über 100 ng/ μ l wurde das Qubit dsDNA BR Assay Kit verwendet. Alternativ wurde das Qubit dsDNA HS Assay Kit genutzt. Die Durchführung erfolgte jeweils nach Angaben des Herstellers.

3.8.2 Elektrophoretische Auftrennung von Nukleinsäuren

Die Auftrennung von Nukleinsäuren erfolgte in nativen Agarosegelen. Abhängig von der Größe der aufzutrennenden Fragmente wurde dafür ein 0,5%iges bis 2%iges Agarosegel in 1x TAE-

Puffer oder 1x TBE-Puffer hergestellt. In der Regel wurden 5 µl PCR-Produkt mit 5x Loading Dye, und 2 µl WGA- und SWGA-Proben mit 6x Loading Dye (Fermentas, St- Leon-Rot) versetzt. Die Elektrophorese wurde bei einer Spannung von 60 V bis 130 V in 1x TAE-Puffer bzw. 1x TBE-Puffer mit einem geeigneten Größenstandard durchgeführt. Nach erfolgter Auftrennung wurde das Agarosegel für 10 min in einer Ethidiumbromidlösung inkubiert und in einem Wasserbad entfärbt. Anschließend wurden die Nukleinsäuren mit der Fusion Fx7 Geldokumentation (Peqlab, Erlangen) detektiert und konnten gegebenenfalls für eine weitere Aufreinigung aus dem Gel geschnitten werden (vgl. 3.8.3).

3.8.3 Extraktion von DNA aus Agarosegelen

PCR-Produkte, die im Weiteren in Plasmid-Vektoren kloniert werden sollten, wurden über das Crystal Gel-Extraktionskit (Biolab Products, Bebensee) aufgereinigt. Dafür wurden die Amplifikate zunächst im Agarosegel aufgetrennt (vgl. 3.8.2) und die entsprechenden Banden mit einem sterilen Skalpell unter UV-Licht ausgeschnitten. Die nachfolgenden Schritte erfolgten nach Angabe des Herstellers. Die DNA wurde mit 50 µl HPLC-Wasser eluiert und anschließend mit der Vakuumzentrifuge auf ca. 10 µl eingeeengt.

Eine Größenselektion nach einer WGA bzw. SWGA erfolgte dagegen mittels Elektroelution. Dafür wurden Bereiche entsprechender Größe aus dem Gel geschnitten und in Dialyseschläuche überführt. Die Elution der Nukleinsäuren erfolgte für 1 h bei 70 V in TAE-Puffer. Anschließend wurde die Polarität der Spannung für 5 min umgekehrt, die DNA aus dem Dialyseschlauch entnommen und über Isopropanol präzipitiert (vgl. 3.8.6).

3.8.4 Enzymatische Aufreinigung der PCR-Produkte

Um vor einer Sequenzierung überschüssige Oligonukleotide und dNTPs zu entfernen, wurden PCR-Produkte enzymatisch mit der ExoSAP-Methode aufgereinigt. Dafür wurden 45 U Exonuclease I (New England Biolabs, Frankfurt am Main) und 9 U Shrimp Alkaline Phosphatase (Roche, Mannheim) in einem Volumen von 50 µl gemischt. 3 µl dieses Ansatzes wurden zu 12,5 µl einer PCR-Reaktion gegeben. Die enzymatische Aufreinigung erfolgte für 25 min bei 37 °C. Nach der Hitzeinaktivierung der Enzyme bei 72 °C für 15 min konnte der Ansatz direkt für die Sequenzierung (vgl. 3.11.1) weiterverwendet werden.

3.8.5 Phenol/Chloroform-Extraktion

Die Phenol/Chloroform-Extraktion wurde angewendet um Nukleinsäurelösungen von Kontaminationen, wie Proteinen zu trennen. Dafür wurden die DNA im Verhältnis 1:1 mit dem Roti-Phenol/Chloroform/Isoamylalkohol-Gemisch (Carl Roth, Karlsruhe) versetzt, für 5 min invertiert und 10 min bei 4 °C und 20800 x g zentrifugiert. Die obere wässrige Phase wurde in ein neues Reaktionsgefäß pipettiert, mit einem Volumen Chloroform gemischt und erneut wie oben beschrieben zentrifugiert. Nach erfolgter Phasentrennung wurde die DNA in ein neues Reaktionsgefäß überführt. Um eine höhere Ausbeute zu erzielen, wurde die phenolische Phase

teilweise mit HPLC-H₂O reextrahiert. Bei einer stark ausgeprägten Interphase wurde die Extraktion über Chloroform wiederholt. Zur Entfernung verbliebender Chloroformreste wurde die DNA mit Isopropanol oder Ethanol präzipitiert (vgl. 3.8.6).

3.8.6 Ethanol- und Isopropanol-Präzipitation

Um Salze zu entfernen und Lösungen zu konzentrieren, wurde DNA mit dem 0,1-fachen Volumen einer 3 M Natriumacetatlösung (pH 5,2) und mit dem 2,5-fachen Volumen 99,8%igen (v/v) Ethanol versetzt. Nach einer Inkubationszeit von mindestens 1 h bei -80 °C oder über Nacht bei -20 °C wurde die DNA durch Zentrifugation (4 °C, 20800 x g, 60 min) pelletiert. Das Pellet wurde zweimal mit dem 5-fachen Volumen einer 70%igen (v/v) Ethanollösung gewaschen und an der Luft oder mit der Vakuumzentrifuge getrocknet. Abhängig von der weiteren Verwendung wurde das Präzipitat in einem geeigneten Volumen TE-Puffer, Tris/HCl-Puffer oder HPLC-H₂O gelöst.

Bei der Fällung großer Volumina wurde die DNA nach dem Versetzen mit Natriumacetatlösung mit dem 0,7-fachen Volumen Isopropanol vermengt. Der Fällungsansatz wurde bis zu 30 min bei Raumtemperatur inkubiert und 30 min bei 4 °C und 20800 x g zentrifugiert. Anschließend wurde die DNA wie oben beschrieben getrocknet und in einem geeigneten Puffer gelöst.

3.9 Klonierungsmethoden

Klonierung beschreibt in der Molekularbiologie alle Methoden, die zur Isolation und identischen Vervielfältigung von DNA-Fragmenten in Wirtsorganismen genutzt werden. Dies schließt die Ligation des entsprechenden DNA-Fragments in einen geeigneten Vektor, die Transformation in kompetente Bakterienzellen, die Selektion positiver Klone und die anschließende Präparation der Plasmid-DNA mit ein. In dieser Arbeit wurde die Klonierung einerseits zur Herstellung von Standardplasmiden für die qPCR verwendet. Aber auch einige PCR-Produkte, die im Weiteren sequenziert werden sollten, wurden zunächst kloniert, um einen Verlust von Sequenzinformationen vermeiden.

3.9.1 Ligation

Die Ligation von PCR-Produkten, die mit der Taq-Polymerase amplifiziert wurden mit dem pGEM-T Vektor System der Firma Promega (Mannheim) nach Herstellerangaben durchgeführt. In der Regel wurden jedoch 3 µl der aus dem Agarosegel aufgereinigten PCR-Produkte mit 1x Rapid Ligations-Puffer, 50 ng pGEM-T Vektor und 3 U T4-DNA-Ligase in einem 10 µl Ansatz für 1 h bei Raumtemperatur oder über Nacht bei 4 °C inkubiert.

PCR-Produkte, die dagegen keine A-Überhänge aufwiesen und beispielsweise durch die Pfu-Polymerase amplifiziert wurden, wurden in den pJET1.2/blunt Vektor ligiert. Die Ligation erfolgte nach Protokoll des Herstellers mit dem CloneJET PCR Cloning Kit (Thermo Fisher Scientific, St. Leon-Rot) im molaren Verhältnis von 1:1, 3:1 oder 7:1 für 1 h bei Raumtemperatur oder über

Nacht bei 4 °C. Die rekombinanten Plasmide wurden anschließend bei -20 °C gelagert oder direkt in kompetente Bakterienzellen transformiert (vgl. 3.9.2).

3.9.2 Transformation und Selektion

Die Transformation der rekombinanten Plasmide in *E. coli* erfolgte mittels Hitzeschock. Dafür wurden 25 µl der chemisch-kompetenten Zellen des Klonierungsstammes JM109 (Promega, Mannheim) dazu auf Eis aufgetaut, mit 4 µL des Ligationsansatzes versetzt und für mindestens 20 min auf Eis inkubiert. Die Zellen wurden für 50 s einem Hitzeschock bei 42 °C unterzogen, für 2 min auf Eis überführt und vorsichtig in 200 µl bis 375 µl LB-Medium resuspendiert. Anschließend wurde der Ansatz zur Regeneration der Kultur für 1,5 h bei 37 °C und 125 rpm im Brutschrank geschüttelt und auf ein bis zwei LB-Amp-Platten bzw. IPTG/X-Gal-Platten ausplattiert und erneut bei 37 °C über Nacht inkubiert. Die gewachsenen Bakterienkolonien konnten auf den entsprechenden Agarplatten mehrere Wochen bei 4 °C gelagert werden.

Einzelkolonien wurden mittels Kolonie-PCR (vgl. 3.7.4) überprüft und bei erfolgreicher Aufnahme des richtigen Plasmids mit einem sterilen Zahnstocher von der Agarplatte in 5 ml LB-Medium überführt. Für die Selektion wurde das Medium mit 10 µg/ml Ampicillin versetzt. Die Kultur wurde über Nacht bei 37 °C und 125 rpm im Brutschrank inkubiert und anschließend zur Plasmidpräparation (vgl. 3.9.3) verwendet werden.

3.9.3 Plasmidpräparation

Zur Isolierung der Plasmid-DNA aus *E. coli* wurden die Bakterien der Übernachtskultur zunächst vollständig sedimentiert (Raumtemperatur, 5.000x g, 10 min) und nach Anleitung des peqGOLD Plasmid Miniprep Kit I (Peqlab, Erlangen) lysiert. Die isolierten Plasmide wurden zweimal in 50 µl Wasser eluiert und mit der Vakuumzentrifuge auf etwa 10 µl bis 20 µl eingengt. Die Konzentration der Plasmid-DNA wurde photometrisch bestimmt (vgl. 3.8.1) und bis zur weiteren Verwendung bei -20°C gelagert.

3.10 Sequenzierung

3.10.1 Sanger-Sequenzierung

Sequenzierungen von Plasmiden und PCR-Produkten wurden durch die Firma GATC Biotech (Konstanz) mit dem ABI 3730xl DNA-Analyzer-System durchgeführt. Dafür wurden bis zu 300 ng Plasmid-DNA bzw. 100 ng PCR-Produkt und 50 pmol Primer in einem 10 µl Ansatz vorbereitet. Für die Sequenzierung wurden entweder vektor- oder sequenzspezifische Oligonukleotide verwendet (vgl. Tab. 2.1).

3.10.2 Paired-End-Sequenzierung

Für die Paired-End-Sequenzierungen einer infizierten und einer nicht-infizierten Fledermaus wurden jeweils zehn WGA-Reaktionen mit den Proben Mg43 bzw. WF32 durchgeführt (vgl. 2.6). Die WGA-Proben wurden vereinigt, über Phenol/Chloroform extrahiert, mit Ethanol gewaschen und in nukleasefreiem Tris/HCl-Puffer resuspendiert (vgl. 3.8.5 und 3.8.6). Die Erstellung der Paired-End-Bibliotheken mit anschließender Sequenzierung wurde durch die Firma GATC Biotech (Konstanz) durchgeführt. Für jede Probe wurden dafür 1,5 µg WGA-DNA zur Verfügung gestellt. Eine Übersicht über die verwendeten Sequenzierplattformen, die hergestellten Fragmentgrößen und sequenzierten Readlängen ist in Tab. 3.6 dargestellt.

Tab. 3.6: Eigenschaften der Paired-End-Sequenzierung.

Probe	Sequenzierplattform	Fragmentgröße	Readlänge	Einrichtung
Mg43	HiSeq 2500 (Illumina, San Diego, US)	750 bp	2 x 126 bp	GATC Biotech (Konstanz)
Mg43	MiSeq (Illumina, San Diego, US)	750 bp	2 x 300 bp	GATC Biotech (Konstanz)
WF32	MiSeq (Illumina, San Diego, US)	750 bp	2 x 300 bp	GATC Biotech (Konstanz)

3.10.3 Mate-Pair-Sequenzierung

Zur Konstruktion der Mate-Pair-Bibliothek wurden jeweils sechs SWGA-Reaktionen mit den Primersets 6 und 7 (Tab. 2.4) und den Proben Mg43 sowie WF43 durchgeführt. Wie auch bei der Vorbereitung zur Paired-End-Sequenzierung wurden die SWGA-Proben vereinigt und aufgereinigt. Nach der Komplementärstrangsynthese und einer weiteren Extraktion über Phenol/Chloroform wurde die SWGA-DNA in Tris/HCl-Puffer resuspendiert. 6,5 µg der genomischen DNA wurden anschließend für die Sequenzierung genutzt. Erstellung der Mate-Pair-Bibliotheken mit den mittleren Fragmentgrößen von 3 kbp und 8 kbp und die Sequenzierung dieser erfolgte durch die Firma StarSeq (Mainz) (Tab. 3.7).

Tab. 3.7: Eigenschaften der Mate-Pair-Sequenzierung.

Probe	Sequenzierplattform	Fragmentgröße	Readlänge	Einrichtung
Mg43/WF43	NextSeq 500 (Illumina, San Diego, US)	3.000 bp	2 x 150 bp	StarSeq (Mainz)
Mg43/WF43	NextSeq 500 (Illumina, San Diego, US)	8.000 bp	2 x 150 bp	StarSeq (Mainz)

3.11 Allgemeine bioinformatische Methoden

3.11.1 DNA-Sequenzanalyse

Die Analyse und Editierung von einzelnen Genen, Genfragmenten, Plasmiden und anderen DNA-Sequenzen erfolgte mithilfe der Software SnapGene v4.1 der Firma GSL Biotech LLC (Chicago, US). Die durch forward und reverse Primer entstandenen Sequenzdaten aus der Sanger-Sequenzierung wurden vor der Editierung visuell kontrolliert. Dafür wurden die entsprechenden

Chromatogramme mit dem Programm FinchTV v1.4.0 Geospiza Inc. (Seattle, US) analysiert. Bereiche schlechter Qualität und gegebenenfalls Vektoranteile wurden aus den Sequenzen entfernt.

Die Analyse der Referenzgenome und die Bearbeitung der in dieser Arbeit sequenzierten NGS Daten erfolgte, wenn nicht anders angegeben, mit dem Programmpaket CLC Genomics Workbench v7.5.1 (QIAGEN Bioinformatics, Aarhus, DK).

3.11.2 BLAST

Ein Abgleich von sequenzierten PCR-Produkten und erstellten Contigs über den BLAST-Algorithmus wurde in der Regel durchgeführt, um Sequenzen zu identifizieren, Sequenzvariationen zu ermitteln, Kontaminationen zu finden oder genomische Sequenzen zu vergleichen. In Abhängigkeit von der konkreten Fragestellung wurden dabei die Varianten blastn, blastx, oder blastp verwendet. Die Suchen erfolgten entsprechend gegen die Datenbanken des NCBI (nt bzw. nr) oder als Vergleich gegen annotierte Proteine und Proteome (Tab. 3.8). Homologievergleiche einzelner Sequenzen wurden in der Regel mit dem Web Tool des NCBI durchgeführt. Aufwendigere Suchen von mehreren Sequenzen z. B. auch gegen in dieser Arbeit erstellte Contigs erfolgten durch die in der CLC Genomics Workbench v7.5.1 (QIAGEN Bioinformatics, Aarhus, DK) implementierten BLAST-Algorithmen. Beim Vergleich von ganzen Assemblys gegen große Nukleotid- oder Proteindatenbanken wurde BLAST lokal oder auf dem High Performance Computing Cluster des regionalen Rechenzentrums der Universität Hamburg eingesetzt. Anwendung fanden die Programme ncbi-blast v2.3.0 sowie GHOSTX v1.3.7. Wenn nicht anders angegeben, wurden für die Homologiesuchen Standardeinstellungen verwendet und jeweils ein Schwellenwert des Erwartungswertes (e-value) von 10^{-5} angenommen.

Tab. 3.8: Übersicht der lokalen BLAST-Datenbanken.

Daten	Art	Sequenzen	Stand	Verwendung
UniProt „Apikomplexa“*	Protein	48.582.547	01/2018	Referenzproteine zur Sortierung
UniProt „Bakterien“	Protein	41.983.916	01/2018	Referenzproteine zur Sortierung
Chiroptera	Protein	405.305	11/2016	Annotation, erste Identifizierung
nr	Protein	103.788.118	11/2017	Sequenzidentifizierung
nt	DNA	41.640.128	08/2017	Sequenzidentifizierung
UniProt „Pilze“	Protein	5.610.606	01/2018	Referenzproteine zur Sortierung
<i>P. falciparum</i>	Protein	5548	04/2017	Annotation, erste Identifizierung
UniProt „Säugetiere“	Protein	885.137	01/2018	Referenzproteine zur Sortierung
WF32Rohdatenassembly	DNA	24.489.380	08/2017	Nicht infizierte Fledermaus zur Sortierung

*zusätzlich annotierte Proteine aus den Genomsequenzierungs-Projekten *B. bigemina*, *B. microti*, *C. muris*, *C. parvum*, *H. hammondi*, *T. annulata*, *T. gondii*, *T. parva*, *N. caninum*, *P. berghei*, *P. chabaudi*, *P. cynomolgi*, *P. falciparum*, *P. galinaceum*, *P. relictum*, *P. reichenowi*, *P. vivax*, und *P. yoelii* bezogen von EuPathDB [Release 29] sowie annotierte Proteine aus *H. tartakovskyi* (Bensch et al., 2016).

3.11.3 Primerdesign

Die verwendeten Primer wurden anhand von öffentlich zugänglichen Sequenzen oder auf Grundlage der in dieser Arbeit erstellten Contigs manuell in SnapGene v4.1 (Chicago, US) oder mithilfe der Tools Primer3 v4.1.0 (Untergasser et al., 2012) bzw. genefisher2 (Giegerich et al., 1996) generiert. Wenn möglich sollten die Primer mindestens 20 Nukleotide lang sein, einen GC-Gehalt von mindestens 40 % aufweisen und eine Schmelztemperatur zwischen 50 °C und 60 °C haben.

Zusätzlich wurden beim Entwurf auf eine geringe Sekundärstrukturbildung und 3'-Selbst-Komplementarität geachtet. Die Ausbildung von Heterodimeren und Schmelztemperaturunterschiede zwischen forward und reverse Primer von mehr als 5 °C wurden vermieden. Die Kontrolle dieser Kriterien erfolgte mit OligoCalc (Kibbe, 2007) und OligoAnalyzer (Integrated DNA Technologies, Illinois, US).

3.11.4 Herstellung der SWGA-Primer

Zur Identifizierung von Primern, die bevorzugt an das Genom von *P. murinus* binden, wurde ein veröffentlichtes Perl-Skript verwendet (Clarke et al., 2017). Ein Set aus 1.189 Sequenzen parasitären Ursprungs, bestehend aus 1.315.758 Nukleotiden parasitären Ursprungs diente dabei als Zielgenom, während 451.916 Contigs (1.329.408.124 Nukleotide) als Hintergrundgenom dienten. Das Hintergrundgenom basierte dabei auf den Sequenzen der infizierten Fledermaus (Mg43), die einen BLAST-Treffer in den Sequenzen der nicht infizierten Fledermaus (WF32) hatten und damit einen Ursprung im Wirtsgenom oder anderen kontaminierenden Organismen aus der ursprünglichen Probe haben (vgl. Abschnitt 4.4). Insgesamt wurden dabei 1.191 Motive mit einer Länge zwischen 8 bp und 12 bp identifiziert, die, relativ gesehen, in den Zielsequenzen häufig vorkamen (mindestens 30 Bindestellen; alle 45.000 bp) und seltener in den Hintergrundsequenzen (maximal 2500 Bindestellen; alle 500.000 bp). Die Liste der Primer wurde anschließend nach mehreren Kriterien gefiltert. Primer mit Schmelztemperaturen unter 25 °C und über 35 °C wurden ausgeschlossen. Um Sekundärstrukturen zu vermeiden, wurden zudem die Primer aus dem Datensatz entfernt, die mehr als drei komplementären Nukleotide an den 3'- und 5'-Enden aufwiesen. Für die Herstellung der Primersets wurden dann auch nur die Primer berücksichtigt, welche in den Sequenzen von *P. murinus* mindestens 15-mal häufiger binden, als im Hintergrundgenom. Pro Set waren zwei bis vier Primer erlaubt. Dabei wurden Sets ausgeschlossen, die Primerpaare mit mehr als drei komplementären Nukleotiden enthielten, um die Ausbildung von Heterodimeren zu verhindern. Als weiteres Kriterium für die Qualität der einzelnen Sets wurde die gleichmäßige Verteilung der Primer-Bindestellen in den Sequenzen von *P. murinus* berücksichtigt.

3.12 Allgemeine Methoden mit NGS-Daten

3.12.1 Bearbeitung von NGS-Daten

Bei der Bearbeitung und Manipulation von NGS-Daten sind unterschiedliche Skripte und Tools zur Anwendung gekommen. Verwendet wurden dabei einerseits Perl-Skripte, die in verwendeten öffentlich verfügbaren Programmpaketen enthalten waren, wie z. B. MAKER2 v2.31.9 und PAGIT v1.64. Andererseits erfolgten Konvertierungen in passende Datenformate beispielsweise mit unterschiedlichsten Funktionen von der CLC Genomics Workbench v7.5.1 (Qiagen) oder von Galaxy v18.05 (Afgan et al., 2016). Zusätzlich wurden auch individuell angepasste Ruby-Skripte verwendet, die von Dr. Janus Borner (Universität Hamburg) verfasst wurden. Diese kamen vor

allem zum Einsatz, um eine Vielzahl von Reads nach bestimmten Kriterien zu sortieren (z. B. GC-Gehalt, BLAST-Analysen).

3.12.2 Aufbereitung der Rohdaten

Die erhaltenen NGS-Rohdaten wurden zunächst mittels FastQC v0.11.5 auf ihre Qualität hin überprüft. Basierend auf den Ergebnissen der Qualitätskontrolle erfolgte die Aufbereitung von Reads mit Bereichen unzureichender Qualität mit der Funktion „Trim Sequences“ des Programmpaket CLC Genomics Workbench v7.5.1. Hierbei wurden üblicherweise die letzten 5 Basen am 3'-Ende der Reads sowie eventuell vorhandenen Adaptersequenzen entfernt. Zusätzlich wurde ein Qualitätsgrenzwert von 0,05 eingestellt und Bereiche von Reads ausgeschlossen, die mehr als zwei nicht eindeutige Basen enthielten. Im Anschluss an dieses Trimming wurden lediglich Sequenzen beibehalten, welche mindestens eine Länge von 50 bp aufwiesen und bei denen sowohl der forward als auch der reverse Read die Qualitätsfilter passiert haben. Im Anschluss wurden die Qualitätsmerkmale der Reads erneut mittels FastQC v0.11.5 überprüft.

Reads nach Mate-Pair-Sequenzierung wurden vor dem Qualitätstrimming in der CLC Genomics Workbench v7.5.1 zusätzlich mit dem Programm NxTrim v0.4.2 prozessiert. Dabei wurden die Standardfunktionen verwendet, um einerseits die Mate-Pair-Sequenzen von den ebenfalls im Datensatz enthaltenen Paired-End-Sequenzen und Einzel-Reads zu trennen und die Reads richtig zu orientieren. Alle erhaltenen Fraktionen wurden nach dem Qualitätstrimming als separate Sequenzier-Bibliotheken beibehalten und für die weiteren Analysen verwendet.

3.12.3 *De novo* Assembly

Im Gegensatz zum Referenz-basierten Assembly oder Mapping werden Reads im *de novo* Assembly ohne Zuhilfenahme von Referenzen zu längeren DNA-Sequenzen, den sogenannten Contigs, zusammengesetzt. Im Allgemeinen erfolgte ein *de novo* Assembly, im Folgenden nur noch als Assembly bezeichnet, mittels CLC Genomics Workbench v7.5.1., das auf der verteilten Repräsentation eines de Bruijn-Graphen beruht. In der Regel wurde unter Anpassung verschiedener Parameter wurden mehrere Contig-Sätze erstellt und empirisch das Assembly mit dem höchsten Nukleotid-Gehalt, der größten mittleren Sequenzlänge und dem höchsten N50-Wert ausgewählt. Wenn nicht anders angegeben, erfolgte ein Assembly mit den Standardeinstellungen unter der Funktion „Map reads back to contigs“. Die resultierenden Contigs mussten eine Mindestlänge von 500 bp besitzen. Auch „word“ und „bubble size“ wurden automatisch durch die Software festgelegt. Für das anschließende Mapping mussten 95 % eines Reads („Length Fraction“) zu 95 % mit der Contigsequenz („similarity fraction“) übereinstimmen.

3.12.4 Mapping

Ein Mapping wurde in der Regel mit der „Map reads to Contigs“-Funktion der CLC Genomics Workbench 7.5.1 durchgeführt. Die Parameter dafür wurden in Abhängigkeit der evolutionären Distanz der Referenzsequenz gewählt. Stammten Ursprungsorganismus und Referenzsequenz aus

derselben Gattung wurden stringenter Parameter eingesetzt als bei lediglich entfernt verwandten Taxa. Teilweise wurden die Parameter dafür empirisch bestimmt. Für Contigs, die aus den sequenzierten Reads dieser Arbeit generiert wurden, mussten jedoch immer mindestens 95 % eines Reads zu 95 % mit der Referenzsequenz übereinstimmen, um in ein entsprechendes Mapping aufgenommen zu werden.

3.12.5 Mapping als Sortierprozess

Abhängig von der jeweiligen Fragestellung wurde die in CLC Workbench 7.5.1 integrierte Funktion „Map Reads to Contigs“ auch als Teil des Sortierprozesses verwendet. Zum Einsatz kam dies vor allem bei der Identifizierung und der Zuordnung der Reads nach dem Ursprungsgenom, wie das des Parasiten oder des Wirts. Da zu Beginn dieser Arbeit weder das Wirtsgenom noch das Genom des Parasiten zur Verfügung stand, erfolgte die Reduktion des Ausgangsdatensatzes auf die zu suchenden Sequenzen auf zwei unterschiedliche Weisen. Abhängig war dies von dem verwendeten Referenzgenom. Wurden beispielsweise *Myotis*-Genome als Referenz verwendet (GCA_000147115.1, GCA_000412655.1, GCA_000327345.1), wurden Reads, die nicht gegen diese positioniert werden konnten, zunächst als Reads mit putativ parasitärem Ursprung interpretiert. Im umgekehrten Fall wurden Reads als putative *P. murinus* Reads identifiziert, sobald sie gegen Referenzgenome aus dem Phylum der Apikomplexa positioniert werden konnten.

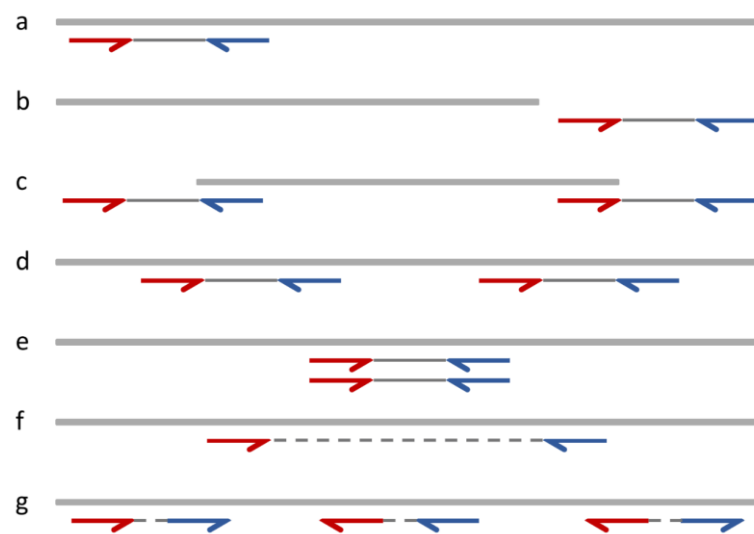


Abb. 3.2: Schematische Darstellung von Positionierungsmöglichkeiten beim Mapping.

Schematische Darstellung von Positionierungsmöglichkeiten beim Mapping. Beim Mapping können sequenzierte Readpaare (rot und blau) in unterschiedlicher Weise an die Referenzsequenz (grau) positioniert werden. Die Paare können dabei unter Einhaltung der Orientierung und des Readabstands genau an eine (a), an keine (b), oder an mehrere Stellen der Referenzsequenz zugeordnet werden (d). Möglich ist auch die Zuordnung von mehreren Paaren als Duplikate an gleicher Stelle (e), die Positionierung mit unerwarteten Abstand (f) oder mit Richtungswechsel (g). Besonders häufig bei nicht nah verwandten Referenzsequenzen ist auch die Positionierung nur eines Reads eines Readpaares (c).

In beiden Fällen wurden Readpaare dabei als Ganzes betrachtet. Reads eines Readpaares, bei dem nur ein Read jeweils gegen die Referenz positioniert werden konnte (Abb. 3.2 c), wurden also so behandelt, als wenn beide Reads in das Mapping integriert worden wären. Extrahierte Reads, die

nur als einzelne Reads (singles) an die Referenz kartiert wurden, wurden dafür mit dem dazugehörigen Read des Readpaares nach dem Mapping vervollständigt. Beim Mapping wurde daher die Funktion „collect un-mapped reads“ genutzt. Im umgekehrten Fall wurden dagegen nur die Reads für die weitere Analyse verwendet, bei denen jeweils beide Reads eines Readpaares nicht gegen eine entsprechende Referenzsequenz positioniert werden konnte.

3.13 Beurteilung des Anteils an Kontaminationen im Assembly

Zur Identifizierung des Ursprungs einzelner Contigs wurden Homologie-Suchen über den BLAST-Algorithmus der Programme ncbi-blast v2.3.0 oder GHOSTX v1.3.7 durchgeführt (vgl. 3.11.2). Eine solche Herangehensweise berücksichtigt allerdings nicht den Umfang aller Kontaminationen einer Sequenzierung. Zur Visualisierung der Kontaminationen der Rohdaten-Assemblys wurde daher ein Blobplot mit BlobTools v1.0.1 erstellt. Dabei wurden alle Contigs gegen die nicht redundante Nukleotid-Datenbank des NCBI abgeglichen und in unterschiedliche Taxa zugeordnet, dabei wurde für den Erwartungswert ein Schwellenwert von 10^{-25} angenommen. Zudem werden durch BlobTools v1.0.1 die Eigenschaften der Contigs, wie GC-Gehalt und Coverage bei der Darstellung berücksichtigt. Für die entsprechenden Assemblys wurden daher sowohl die Contigs im fasta-Format als auch die dazugehörigen Mappings im Format SAM aus der CLC Workbench 7.5.1 exportiert.

3.14 Auftrennung der wirts- und parasitenspezifischen Reads

Die Auftrennung der parasitenspezifischen Sequenzen aus dem NGS-Datensatz der infizierten Wasserfledermaus erfolgte über zwei Herangehensweisen. Einerseits erfolgte dies relativ unspezifisch anhand von Referenzgenomen nah verwandter Parasiten oder Wirtsfledermäusen, oder etwa dem GC-Gehalt der Sequenzen (vgl. 3.14.1), andererseits aber auch durch die Sequenzierung der Wirtsfledermaus und einer stringenten Aufreinigung weiterer Verunreinigungen (vgl. 3.14.2).

3.14.1 Readextraktion über verwandte Referenzgenome und GC-Gehalt

Bei der Extraktion der Pair-End-Reads mit putativ parasitärem Ursprung wurde großen Wert daraufgelegt, leicht reproduzierbare Methoden zu verwenden, die darauf ausgelegt sind, in kurzer Zeit biologisch relevante Informationen zu generieren. Grundlage dafür waren drei unterschiedliche Strategien (Abb. 3.3). Einerseits wurden die überrepräsentierten Sequenzen des Wirtes durch Positionierung der Reads (Mapping) an Referenzgenome reduziert, andererseits erfolgte, ebenfalls über eine Mapping-Strategie, eine Anreicherung von Reads parasitären Ursprungs. Da zu dem Zeitpunkt dieser Arbeit weder das Genom der Wasserfledermaus noch das Genom eines Parasiten der Gattung *Polychromophilus* vorlag, wurden als Referenzgenome jeweils verwandte Organismen verwendet. Bei beiden Methoden stand lediglich die Identifizierung der entsprechenden Reads im Vordergrund, der eigentliche Kartierungsprozess beim Mapping war weniger von Bedeutung, daher wurden mehrere Referenzen zugleich verwendet. Im sogenannten

Myotis-Filter erfolgte ein Mapping gegen drei Genome der Gattung *Myotis*, während zur Identifizierung der putativen Reads parasitären Ursprungs, im Folgenden als Apikomplexa-Filter bezeichnet, ein Mapping mit 24 Referenzgenomen des Phylums Apikomplexa durchgeführt wurde (vgl. 3.12.5). In einem weiteren Ansatz wurden die Reads aufgrund ihres GC-Gehalts vorsortiert (GC-Filter).

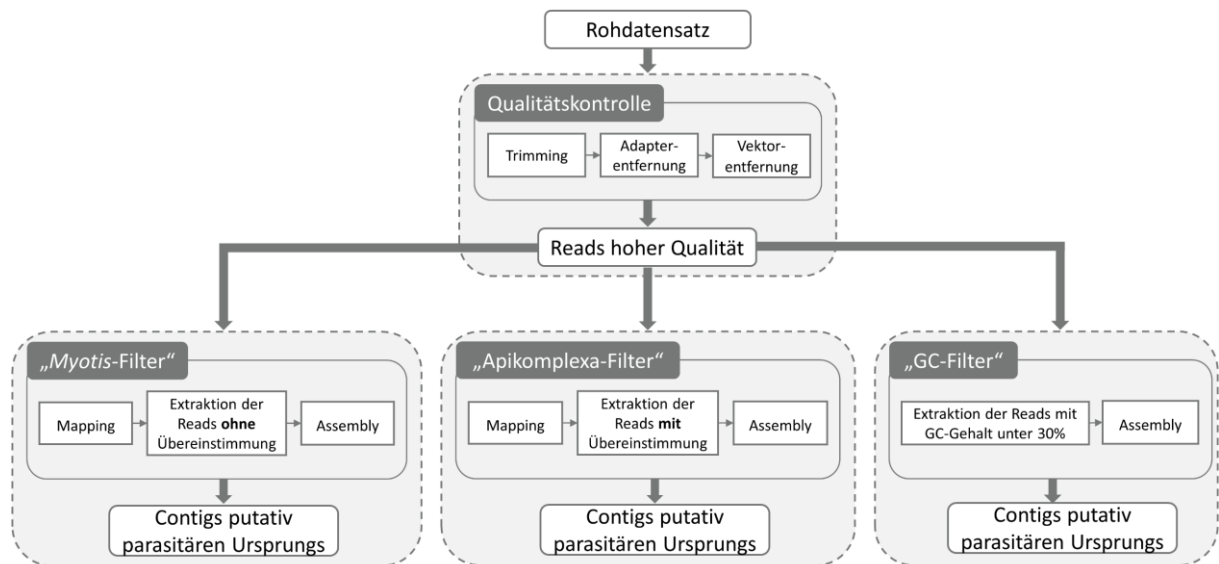


Abb. 3.3: Übersichtsschema zur Identifikation der Paired-End-Reads von *P. murinus*.

Sequenziert wurde eine Wasserfledermaus, die mit dem Erreger *P. murinus* infiziert war. Die Extraktion der Sequenzen parasitären Ursprungs erfolgte über das Mapping gegen drei verwandte Fledermäuse der Gattung *Myotis* (*Myotis*-Filter), gegen 24 Arten des Phylums Apikomplexa (Apikomplexa-Filter) oder anhand des GC-Gehalts der Reads (GC-Filter).

Im Vorfeld wurden unterschiedliche Parameter im Aufreinigungsprozess gegeneinander ausgetestet. Entscheidend für die Auswahl war dabei die Größe und der durchschnittliche GC-Gehalt des anschließenden Assemblys. Aber auch die Abundanz der bislang von *P. murinus* sequenzierten Gene (vgl. Tab. 7.4) spielten nach einer vorläufigen Annotation (vgl. 3.16) spielten bei der Bewertung optimaler Einstellungen eine Rolle. Als vergleichsweise gute Parameter erwiesen sich für den Sortierprozess (vgl. 3.12.5) im *Myotis*-Filter eine 90%igen Homologie über 90% der Sequenzlänge (0,9/0,9), während im Apikomplexa-Filter eine 80%ige Homologie über 80 % der Readlänge (0,8/0,8) bereits ausreichend war. Bei der Aufteilung der Reads nach ihrem GC-Gehalt wurden Readpaare weiterverwendet, bei dem beide Reads einen GC-Gehalt von unter 30 % aufwiesen. Genutzt wurde hierfür ein individuell angepasstes Ruby-Skript, das von Dr. Janus Borner (Universität Hamburg) verfasst wurde.

3.14.2 Readextraktion basierend auf genomischen Sequenzen der Wirtsfledermaus

Das Sortieren der Reads aus der Sequenzierung der infizierten Fledermaus erfolgte in drei Phasen (Abb. 3.4). Zunächst wurden die Reads gegen eine Datenbank, die Sequenzen der nicht infizierten Wasserfledermaus enthielt abgeglichen (vgl. 3.11.2). Verwendet wurden dabei einerseits die Contigs des Rohdaten-Assemblys und andererseits einzelne Reads, die nicht zu Contigs zusammengefügt

werden konnten. Die Reads die bei einem e-value von maximal 10^{-5} aufwiesen (90.556.992) wurden zum Assembly der infizierten Wasserfledermaus genutzt und für die Herstellung der SWGA-Primer verwendet (vgl. 3.11.4) von dem restlichen Datensatz getrennt. Alle Reads, die keinen signifikanten Treffer erzielten, wurden in einem Assembly bestehend aus 382.549 Contigs zusammengesetzt (N50: 637 bp). Entstandene Contigs waren zwischen 300 bp und 82.340 bp lang und wurden mit dem blastx Algorithmus von GHOSTX v1.3.7 gegen annotierte Proteine unterschiedlicher Taxa abgeglichen (vgl. Tab. 3.8). Contigs die gar keinen Treffer erzielten und solche, die homologe Sequenzen zu Proteinen aus dem Phylum der Apikomplexa aufwiesen, wurden beibehalten. Insgesamt konnten auf diese Weise 110.167 Contigs mit putativ parasitärem Ursprung identifiziert werden. Die Reads, aus denen die entsprechenden Contigs zusammengesetzt wurden anschließend extrahiert. Einzelne Reads wurden mit dem dazugehörigen Read des Readpaares über ein Ruby-Skript (zur Verfügung gestellt von Dr. Janus Borner (Universität Hamburg)) vervollständigt. Eine erste Analyse der identifizierten Reads in der CLC Workbench 7.5.1 zeigte eine deutlich bimodale Verteilung (nicht dargestellt). Daher wurden lediglich solche extrahiert, die einen mittleren GC-Gehalt von unter 30 % aufwiesen. Die entsprechenden Reads wurden anschließend in einem Assembly bestehend aus 9.140 Contigs zusammengesetzt (N50: 870 bp). Auch die Contigs aus dem anschließenden Assembly wurden anschließend noch einmal nach dem GC-Gehalt gefiltert ($< 25\%$). Insgesamt wurde der Datensatz damit auf 1.189 Contigs mit insgesamt 1,29 Mbp (bestehend aus 1,3 % der ursprünglich sequenzierten Reads) verkleinert. Die finalen Contigs wurden ebenfalls zur Herstellung der SWGA-Primer verwendet (vgl. 3.11.4).

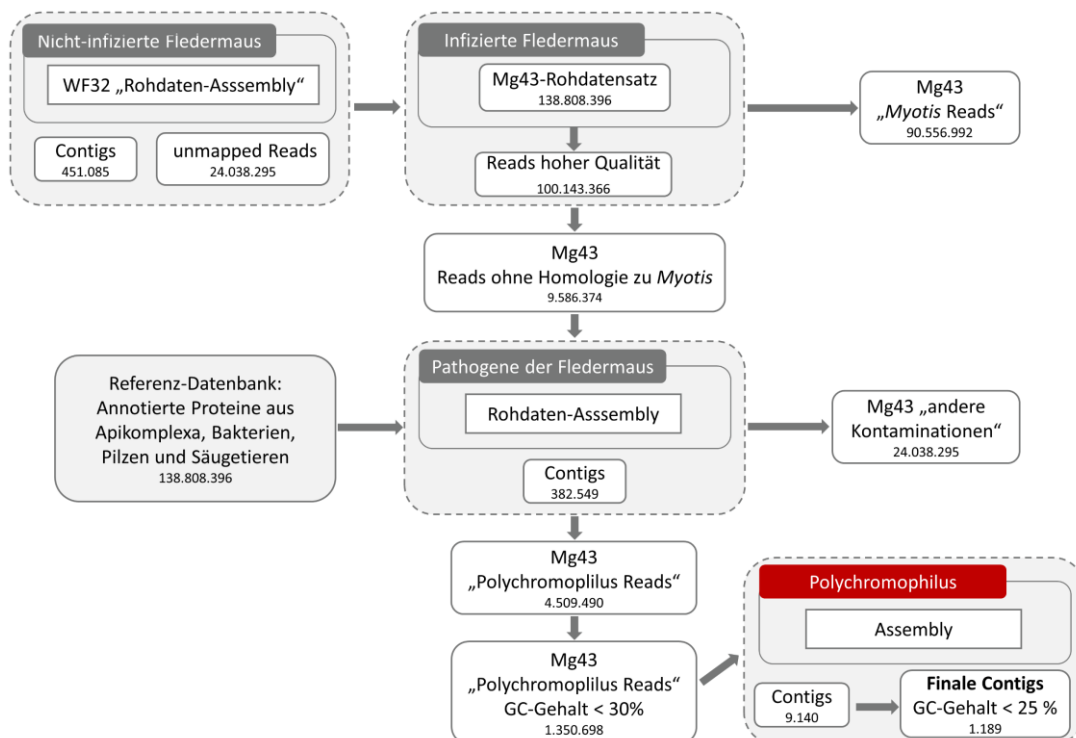


Abb. 3.4: Herstellung von Contigs parasitären Ursprungs auf hohem Zuverlässigkeitsniveau.

Schematische Darstellung der methodischen Vorgehensweise zur Sortierung von Reads des Parasiten *Polychromophilus murinus* aus den Sequenzen einer infizierten Wasserfledermaus (Mg43). Die Reads wurden mit dem blastn-Algorithmus gegen die Sequenzen einer nicht infizierten Wasserfledermaus (WF32) und mit dem blastx-Algorithmus gegen Proteine-Sequenzen der RefSeq-Datenbank abgeglichen. Anschließend erfolgte eine Auftrennung der Sequenzen nach GC-Gehalt.

3.15 Referenzbasiertes Assembly

3.15.1 Herstellung der genomischen Sequenzen von *P. murinus*

Für die Zusammensetzung des Genoms von *P. murinus* wurde zunächst ein Assembly mit den Reads der Mate-Pair-Sequenzierung und den gefilterten Reads aus der Paired-End-Sequenzierung (GC-Gehalt unter 30 %, vgl. 3.14) erstellt. Angewendet wurde dabei die in CLC Workbench v7.5.1 integrierte Funktion „*De novo* Assembly“ mit einer „word size“ von 64 bp und einer minimalen Contiglänge von 500 bp. Die erstellten Contigs (N50: 1.632; zwischen 500 bp und 31.978 bp) wurden anschließend anhand des Referenzgenoms von *Plasmodium falciparum* [3D7; PlasmoDB Release 32] orientiert und geordnet. Die Contigs wurden dafür in der PAGIT v1.64 Pipeline mit dem Tool ABACAS an die Sequenz des Referenzgenoms positioniert. Da eine gewisse Divergenz zwischen den Sequenzen erwartet wurde, erfolgte ein Alignment basierend auf den Translationen beider Genome in Aminosäuresequenzen mit dem Skript PROmer aus dem MUMmer Paket v3.2; abgesehen davon, wurden bei der Positionierung Standardeinstellungen verwendet. 1.553 Contigs von insgesamt 13.292 Contigs wurden dabei in größeren Scaffold platziert. Um entstandene Lücken zu schließen, wurden an diesen und an jedem Contig-Ende lokale Assemblys hergestellt. Verwendet wurde dafür das Tool IMAGE v2.4.1, das ebenfalls Teil der PAGIT Pipeline ist. Insgesamt wurden lokale Assemblys in Gruppen von jeweils drei Iterationen mit den k-mer Größen 75, 63, 51 und 39 durchgeführt (Insgesamt 12 Iterationen). Dabei wurden 131 Lücken (von 26.309) geschlossen und 3.428 Contigs verlängert. Die Contigs wurden anschließend in Scaffolds zusammengeführt mit Lücken, die 300 Nukleotide überspannen. Die am Referenzgenom orientierten Sequenzen wurden anschließend weiter vervollständigt. Dafür wurde eine erste Annotation Companion v1.0.1 durchgeführt (vgl. 3.16). Genfragmente, die eindeutig als unterschiedliche Fragmente desselben Gens identifiziert werden, wurden in der CLC Workbench v7.5.1 manuell zusammengeführt. Fehlende Gene, die im Referenzgenom *P. falciparum* in direkter Nachbarschaft zu den bereits geordneten Contigs von *P. murinus* positioniert vorlagen, und noch ungeordnet aus der PAGIT Pipeline hervorgegangen sind, wurden manuell in die Contigs integriert. Die Qualität der bearbeiteten Contigs wurde anschließend über ein Mapping validiert. Regionen, die nicht über gepaarte Reads verbunden wurden, wurden aus den größeren Contigs wieder entfernt und fehlerhaft zusammengesetzte Contigs mit REAPR v1.0.18 wieder aufgebrochen. Neu entstandene Lücken, zwischen den neu zusammengestellten größeren Contigs wurden anschließend ein weiteres Mal mit IMAGE v2.4.1 wie oben beschrieben vervollständigt und manuell, nach einem weiteren Mapping, editiert.

3.15.2 Herstellung der genomischen Sequenzen von *M. daubentonii*

Für die Rekonstruktion des Genoms von *Myotis daubentonii* wurden die Reads aus der Paired-End-Sequenzierung der nicht infizierten Wasserfledermaus (WF32) verwendet. Dabei wurde mit der CLC Workbench v7.5.1 zunächst ein *de novo* Assembly zusammengesetzt (vgl. 4.3.1). Verwendet wurden dabei die Standardeinstellungen mit einer „word size“ von 50 bp und einer minimalen Contiglänge von 1000 bp. Reads, aus der Paired-End-Sequenzierung, die nicht gegen die erstellten

Contigs positioniert werden konnten (24.038.295 Reads), wurden mit der Funktion „collect unmapped reads“ gespeichert und mit der mit der „Map Reads to Reference“ Funktion an die Contigs des Rohdaten-Assemblys der infizierten Fledermaus kartiert (vgl. 3.12.4). Contigs, die mithilfe von BlobTools v1.0.1 als Kontaminationen identifiziert wurden, wurden zuvor aus dem Assembly entfernt (vgl. 3.13). Aus dem hergestellten Mapping wurden anschließend Konsensus-Sequenzen mit einer minimalen Größe von 500 bp generiert. Contigs, die Regionen mit einer geringen Abdeckung (weniger als 2 Reads) aufwiesen, wurden gespalten. Reads, die nicht mit im Mapping integriert worden sind, wurden anschließend noch einmal gegen das Genom von *M. davidii* (GCA_000412655.1) kartiert. Aufgrund der evolutionären Distanz zwischen *M. daubentonii* und *M. davidii* musste mindestens 90 % eines Reads zu 90 % mit der Referenzsequenz übereinstimmen, um in das Mapping aufgenommen zu werden. Auch hier wurden wie oben Konsensus-Sequenzen generiert.

Alle hergestellten Sequenzen wurden mit BlobTools v1.0.1 analysiert. Contigs, die eindeutig einen anderen Ursprung als die Wasserfledermaus hatten, wurden aus dem Assembly entfernt (vgl. 3.13). Der Datensatz wurde anschließend mit RaGOO v1.0.18 (Alonge et al., 2019) an dem Genomsequenzen der Referenz-Fledermaus *M. davidii* (GCA_000412655.1) orientiert. Contigs die nicht platziert werden konnten wurden dabei einzeln gespeichert. Lücken in den Sequenzen wurden mit Gapfiller 1.10 dreimal in 3 Iterationen weitgehend vervollständigt. Dabei wurden die K-mere 75, 65 und 55 verwendet. Andere Parameter (-n 50, -d 75) wurden während der Iterationen nicht verändert. Das resultierende finale Genom wurde zur weiteren Analyse genutzt.

3.16 Annotation

Im Anschluss an das Assembly wurden die Sequenzen annotiert. Einfache Zuordnungen von kodierenden Bereichen, die innerhalb dieser Arbeit als Qualitätskontrolle während unterschiedlicher Teilschritte genutzt wurden, erfolgten durch Homologievergleiche. Dabei wurden Contigs mit der „Find Open Reading Frames“-Funktion der CLC Workbench v7.5.1 auf offene Leserahmen ohne Startkodon untersucht, die identifizierten Bereiche in die jeweiligen Aminosäuresequenzen übersetzt und gegen die nichtredundante Proteindatenbank des NCBI abgeglichen (Abschnitt 3.11.2). Zur Identifizierung repetitiver Elemente wurde RepeatMasker v4.0.7 verwendet, wobei als Datenbank eukaryotischer repetitiver Elemente Repbase Update hinterlegt wurde (Bao et al., 2015). Erste Annotationen von Sequenzen mit Ursprung im Parasiten *P. murinus* erfolgten unter Standardeinstellung mit Companion v1.0.1, wobei *P. falciparum* [3D7; PlasmoDB Release 32] als Referenzorganismus ausgewählt wurde. Nichtcodierende RNAs wurden mit tRNAscan-SE v2.0, RNAmmer v1.2 und der Rfam Datenbank identifiziert. Die Annotation der kodierenden Gene erfolgte dagegen mit der Pipeline MAKER2 v2.31.9.

3.16.1 Annotation des Genoms von *P. murinus*

Zur Annotation des Genoms von *P. murinus* wurden zunächst die Contigs des Rohdaten-Assemblys mit den Reads der Mate-Pair-Sequenzierung in erster Annäherung mit Companion v1.0.1 annotiert.

Als Referenzdatensatz diente das Genom von *Plasmodium falciparum* [3D7; PlasmoDB Release 32]. Die dabei vorhergesagten Proteinsequenzen (2.101 Proteine, bestehend aus durchschnittlich 574 Aminosäuren) wurden bei der Annotation des finalen Assemblys in der Pipeline MAKER2 v2.31.9 als Eingangsdatei verwendet. Zusätzlich wurden in der initialen Analyse Gene aus dem in Augustus v3.2.2 implementierten Genmodell für *P. falciparum* abgeleitet. Die ersten Ergebnisse der Annotation wurden genutzt, um ein Hidden-Markow-Modell (HMM) für das Gen-Identifizierungs-Programm SNAP [2013] zu erstellen. Das Genmodell wurde anschließend in einem zweiten Durchlauf von MAKER2 v2.31.9 verwendet. Als zusätzliche Referenzen wurden Aminosäuresequenzen aus den Referenzorganismen *P. falciparum* [3D7; PlasmoDB Release 32], *P. gallinaceum* [8A; PlasmoDB Release 32], *P. relictum* [SGS-1 like; PlasmoDB Release 32], und *Heamoproteus tartakovskiy* (<http://mbio-serv2.mbioekol.lu.se/Malavi/Download>; Zugriff im Februar 2017) hinzugezogen. Basierend auf den Ergebnissen der zweiten MAKER2-Analyse wurde anschließend ein neues HMM für die Genvorhersage durch SNAP erstellt und in einer weiteren Re-Annotation mit Maker2 verwendet. Um artspezifische Exon-Intron-Grenzen optimal voraussagen zu können, wurden die Parameter zur Genvorhersage mit dem Programm Augustus optimiert. Dafür wurde ein Training-Set aus 100 vorhergesagten Genen des dritten MAKER2-Durchlaufs erstellt und spezie-spezifische Genstrukturen daran in einem Generalisierten HMM abgeleitet. Das erstellte Genmodell wurde anschließend in einer weiteren Re-Annotation mit der MAKER2-Pipeline verwendet.

3.16.2 Annotation der mitochondrialen und apikoplastiden DNA von *P. murinus*

Eine erste Annotation der extrachromosomalen DNA von *P. murinus* erfolgte mit der Pipeline Companion v1.0.1, die auf die Annotation eukaryotischer Parasiten (Apikomplexa und Kinetoplastida) ausgelegt ist. Als Referenzsequenzen wurde dabei die extrachromosomalen Sequenzen von *P. falciparum* [3D7; PlasmoDB Release 32] verwendet. Der Abgleich erfolgte unter Standardeinstellungen. Nichtkodierende RNAs wurden in erster Näherung ebenfalls unter Standardeinstellung mit tRNAscanSE v2.0, Aragorn v1.2.38, und RNAmmer 1.2 identifiziert.

Die Identifizierung der fragmentierten rDNA der mitochondrialen erfolgte nach (Feagin et al., 2012). Dafür wurde die mitochondriale Sequenz mit blastn gegen die entsprechend annotierten Fragmente sequenzierter Plasmodien abgeglichen. Alle auf diese Weise annotierten Gene wurden über BLAST-Suchen verifiziert und gegebenenfalls manuell korrigiert.

3.16.3 Annotation des Genoms von *M. daubentonii*

Die Annotation der Wasserfledermaus erfolgte analog zu der des Parasiten *P. murinus* in der MAKER2 v2.31.9 Pipeline. Zur Reduktion des Rechenaufwands wurden repetitive Sequenzen des Genoms mit der Standardkonfiguration für eukaryotische Genome des Repeatmaskers v4.0.7 im Vorfeld maskiert. In einem ersten Durchlauf mit MAKER2 v2.31.9 dienten die Proteinsequenzen von *Myotis davidii* [NCBI Annotation Release 101] als Referenz. Für die erste *ab initio* Annotation

mit dem implementierten Tool Augustus v3.2.2 wurde das Genmodell des Menschen herangezogen. Insgesamt wurden drei konsekutive MAKER2-Durchläufe vorgenommen, wobei die Ergebnisse des vorausgegangenen Durchlaufs jeweils die automatische Annotation des darauffolgenden unterstützten. Die genutzten Genmodelle für die automatische Genvorhersage mit SNAP v2013 und Augustus v3.2.2 wurden analog zur Annotation des Genoms von *P. murinus* zwischen den einzelnen MAKER2-Durchläufen optimiert.

3.16.4 Annotation der mitochondrialen DNA *M. daubentonii*

Zur Annotation der mitochondrialen DNA der Wasserfledermaus wurde der Mito WebServer (Bernt et al., 2013) verwendet. Regulatorische Elemente (konservierte Sequenzierblöcke, OriL) innerhalb der Sequenz wurden durch eine BLAST-Suche gegen entsprechende Bereiche von anderen annotierten mtDNA-Sequenzen anderer Wirbeltiere identifiziert. Zusätzlich erfolgte mithilfe des TRF v4.09 eine Analyse auf repetitive Regionen (Benson, 1999).

3.17 Evaluierung der Annotation

Zur Beurteilung der Qualität der Gen-Annotationen wurden die kodierende Sequenzbereiche auf Vollständigkeit universell orthologer Gene untersucht. Für die Annotation von *P. murinus* wurde die Gendatenbank für Protisten herangezogen (Datenbank vom 18.11.2016, Anzahl orthologer Gene 215). Die Analyse erfolgte im Modus für Transkripte. Zum Vergleich wurden zusätzlich kodierende Sequenzbereiche von *Plasmodium falciparum* [3D7; PlasmoDB Release 32], *P. gallinaceum* [8A; PlasmoDB Release 44], *P. relictum* [SGS1-like; PlasmoDB Release 44] und *H. tartakovskyi* (Bensch et al., 2016), unter gleichen Bedingungen analysiert.

Die Annotation der Wasserfledermaus wurde dagegen anhand des BUSCO-Datensatzes für Laurasiatheria bewertet (Datenbank odb9 vom 23.02.2017, Anzahl orthologer Gene 6.253). Auch hier erfolgte eine Analyse anhand der Transkripte. Analog wurden auch die Annotationen von *M. brandtii* (GCA_000147115.1), *M. davidii* (GCA_000412655.1), *M. lucifugus* (GCA_000327345.1) und *M. myotis* (GCA_0004026985.1) untersucht.

3.18 Funktionelle Annotation

Proteinfunktionen wurden in einem ersten Schritt durch Homologieinformationen zugewiesen. Dafür wurden die Genprodukte von *P. murinus* gegen die Proteinsequenzen der verwandten Haemosporiden *H. tartakovskyi* (Bensch et al., 2016), *P. falciparum* [3D7; PlasmoDB Release 32] und *P. relictum* [SGS-1 like; PlasmoDB Release 32], abgeglichen (blastp, e-value cutoff 10^{-10} ; vgl. 3.11.2). Zudem wurden die Sequenzen auf Signalsequenzen untersucht und funktionell eingeteilt.

3.18.1 Vorhersage von Domänen und Signalsequenzen

Zur Vorhersage der subzellulären Lokalisation der abgeleiteten Proteine wurden transmembraner Domänen mit TMHMM v2.0 (Sonnhammer et al., 1998) identifiziert und Signalpeptide mit SignalP v4.1 (Petersen et al., 2011) vorhergesagt. Zudem wurde eine weitere Analyse mit Phobius v1.01 (Käll et al., 2004) durchgeführt, um N-terminale Aminosäuresequenzen, die sowohl als Signalsequenz als auch als transmembrane Domäne vorhergesagt wurden, weiter einzugrenzen. Signalsequenzen, welche eine Lokalisierung in das Mitochondrium oder den Apikoplasten initiieren, wurden mit jeweils mit MitoProt II 1.101 (Claros et al., 1996) oder PATS [30/04/2018] (Zuegge et al., 2001) identifiziert. Für die Vorhersage von PEXEL oder dem ER-Retentionssignal wurden die Aminosäuresequenzen in der CLC Workbench 7.5.1 nach den Sequenzmotiven RxLxE/Q/D, RxLxx, R/KxL/IxE/Q/D oder R/KxL/IxxE/Q/D bzw. xDEL durchsucht, wobei x jede beliebige Aminosäure darstellt. Das ER-Retentionssignal musste sich dabei am Ende der jeweiligen Sequenz befinden, während das PEXEL N-Terminal innerhalb der ersten 100 Aminosäuren positioniert sein musste in Proteinen, die gleichzeitig ein Signalpeptid oder mindestens eine Transmembrandomäne aufwiesen. Funktionelle Domänen wurden darüber hinaus mit InterProScan v.5.18-57.0 (Jones et al., 2014) vorhergesagt. Dabei wurden die Analysen TIGRFAM, ProDom, SMART, ProSiteProfiles, ProSitePatterns, SUPERFAMILY, PRINTS, PANTHER, Gene3d, PIRSF, Pfam, Coils aktiviert.

3.18.2 Gene Ontology (GO) Annotation

Die weitere funktionelle Charakterisierung der Proteine erfolgte über die Zuordnung zu Gene Ontology Kategorien (The Gene Ontology Consortium., 2019). Dabei werden Genprodukte hierarchisch in unterschiedliche Aspekte biologischer Prozesse, molekularer Funktionen und/oder zellulärer Komponenten eingeordnet. Die Zuordnung zu GO-Kategorien erfolgte durch eine Analyse mit InterProScan v.5.18-57.0 (Jones et al., 2014). Zudem konnten die entsprechenden GO-Kategorien auch durch die homologen Proteine des Parasiten *P. falciparum* auf dem Gene Ontology Server zugeordnet werden, welcher online unter <http://geneontology.org/> abgerufen werden kann (v10.5281/zenodo.2529950, zuletzt verwendet 10.03.2020). Für die weitere Analyse wurde mithilfe von PANTHER v14.0 (Mi et al., 2019) überprüft ob einzelne GO-Kategorien in ausgewählten Datensätzen unter- bzw. überrepräsentiert sind. Bei dem Genom des Parasiten wurden daher zunächst homologe Proteine identifiziert, die (1) in Piroplasmiden und Haemosporiden, (2) in Haemosporiden oder (3) lediglich in der Familie Plasmodiidae konserviert sind. Ausgehend von den Gen-IDs von der orthologen Gene von *P. falciparum*, wurden statistische Überrepräsentationstests für die kompletten Kategorien biologische Prozesse, molekulare Funktionen und zellulärer Komponenten durchgeführt.

3.19 Identifizierung orthologer Proteine

Die Zuordnung zu orthologen Proteinen erfolgte mit OrthoMCL v5 (Li, 2003). Dafür wurden zunächst 127.395 phylogenie-basierte orthologe Proteinsequenzen der Proteome von 7

Piroplasmiden und 16 Plasmodienarten (vgl. Tab. 7.5) bezogen. Die Proteine sind dabei bereits aufgrund Ihrer OrthoMCL-Annotation in distinkte orthologen Gruppen (OG) eingeteilt. Mithilfe eines individuell angepasste Ruby-Skripts (zur Verfügung gestellt von Dr. Janus Borner, Universität Hamburg) wurden die Proteinsequenzen innerhalb dieser OG in einzelnen Dateien zusammengefasst, sodass pro OG jeweils eine fasta-Datei erstellt wurde, die die entsprechenden Proteine aller untersuchten Arten enthielt. Zudem wurden in diesem Schritt bereits alle OG ausgeschlossen, die Multikopie-Gene enthalten.

Eine Zuweisung der Proteinsequenzen von *H. tartakovskyi* und *P. murinus* zu den OG erfolgte durch den WebServer von OrthoMCL v5 (<https://orthomcl.org/>; letzter Zugriff 8. November 2018). Anschließend wurden diese in die erstellten fasta-Dateien integriert. Auch dafür wurde ein von Dr. Janus Borner (Universität Hamburg) erstellte Ruby-Skripts verwendet. Insgesamt wurden auf diese Weise 2.149 OG erstellt. Die Zuordnungen wurden einerseits für die Anreicherungs-Analysen der GO-Kategorien verwendet (vgl. 3.18.2) oder bei der Herstellung phylogenetischer Analysen weiterverwendet (vgl. 3.20).

3.20 Phylogenetische Analyse der Sequenzdaten

Für die phylogenetische Einordnung der erzeugten Sequenzdaten wurden neben den eigenen Sequenzen je nach Fragestellung öffentlich zugängliche Sequenzen der NCBI- bzw. EuPath-Datenbank genutzt. Die Erstellung der multiplen Protein- und Nukleinsäurealignments erfolgte einzeln für jede orthologe Sequenz mit dem L-INS-i-Algorithmus des Programms MAFFT v7.3.13. Lagen für die entsprechenden Proteinalignments keine Proteinsequenzen vor, wurden diese zuvor mithilfe des ExPASy Translate Tools in Proteinsequenzen translatiert. Schlecht konservierte Bereiche oder andere qualitativ mangelhafte Alignmentabschnitte wurden durch Gblocks v0.91b aus den Alignments entfernt. Aus den einzelnen Alignments wurde anschließend ein konkateniertes Alignment hergestellt. Verwendet wurde dabei ein Ruby-Skript, das von Dr. Janus Borner (Universität Hamburg) zur Verfügung gestellt wurde. Die erhaltenen Alignments wurden bei Bedarf in GeneDoc v2.7 oder Seaview v4.2.3 manuell korrigiert. Um ein geeignetes Substitutionsmodell bei der phylogenetischen Analyse zu verwenden, wurden Proteinalignments durch ProtTest v3.4.2 und Nukleinsäurealignments durch jModeltest v2.1.10 analysiert. Die Auswahl eines geeigneten Modells erfolgte anhand des akaiken Informationskriteriums (AIC). Eine Übersicht über die erstellten Alignments sowie die im Weiteren verwendeten Substitutionsmodelle für dieser Datensätze finden sich in Tab. 3.9.

Tab. 3.9: Übersicht der Datensätze für die phylogenetischen Analysen.

Name	Datentyp	Außengruppe (Anzahl)	Taxa	Gene/ Proteine	Positionen	fehlende Daten	Modell
Genomanalyse	aa	Piroplasmida (7)	26	315	108.862	5,3 %	WAG+G-F
Myotis mtDNA	nuc	Murina (2)	31	13	11.421	2,59 %	GTR+G
Api mtDNA (1)	nuc	Leucocytozoon (4)	41	-	6.714	<0,001 %	GTR+I+G
Api mtDNA (2)	aa	Leucocytozoon (4)	43	3	1.134	<0,001 %	LG+I+G+F
Api pDNA	aa	Babesia (3)	24	24	7.590	0,25 %	JTT+G+F

*aa, Aminosäuren; Api, Apikomplexa; nuc, Nukleinsäuren.

Die phylogenetischen Analysen der unterschiedlichen Datensätze erfolgten in der Regel über mehrere Methoden. Maximum-Parsimony-Rekonstruktionen (MP) erfolgten, abhängig vom Alignment, entweder mit Protpars oder mit Dnapars. Verwendet wurden dabei die implementierten Algorithmen aus dem PHYLIP v3.696 Paket unter Standardeinstellungen. Maximum Likelihood-Analysen wurden mit Garli v2.0.1 unter Berücksichtigung von Außengruppen und dem ermittelten Substitutionsmodell berechnet. Die Bewertung der statistischen Signifikanz der ermittelten Topologien aus MP- und ML-Analysen erfolgte durch Bootstrap-Analysen. Die Bootstrapwerte wurden auf einen Konsensusbaum kartiert. Bei den ML-Analysen erfolgte dies mit dem im Paket DendroPy v4.3.0 implementierten Algorithmus SumTrees. Phylogenetische Rekonstruktionen, basierend auf der Bayes'schen-Interferenz (BI), wurden mit MrBayes v3.2.6 durchgeführt. Auch hier wurden die ermittelten Substitutionsmodelle berücksichtigt. Die Erstellung eines 50% majority rule Konsensus Baums mit Posterior-Probability-Werten (pp) als statistische Unterstützung erfolgte in MrBayes v3.2.6 über den Befehl „sumt“. Eine Übersicht über alle verwendeten Berechnungsparameter der einzelnen phylogenetischen Analysen ist in Tab. 3.10 dargestellt. Die grafische Darstellung der Topologien erfolgte über FigTree v1.4.3 oder Treeview v1.6.6.

Tab. 3.10: Verwendete Parameter der phylogenetischen Analysen.

Name	Software	Parameter
Genomanalyse	Garli 2.01	datatype=aminoacid ratematrix=wag ratehetmodel=gamma numratecats=4 statefrequencies=empirical invariantsites=estimate searchpreps=10 bootstrappreps=10
Myotis mtDNA	Garli 2.01	datatype=nucleotide ratematrix=6rate ratehetmodel=gamma numratecats=4 statefrequencies=empirical invariantsites=none searchpreps=10 bootstrappreps=100
	MrBayes v3.2.6	lset nucmodel=4by4 nst=6 rates=gamma; mcmc nruns=2 nchains=4 ngen=10000000 printfreq=1000 samplefreq=1000 savebrlens=yes; sumt Burnin=5000
	PHYLIP 3.696	Seqboot: 1000 Bootstrap replicates
Api mtDNA (1)	Garli 2.01	datatype=nucleotide ratematrix=6rate ratehetmodel=gamma numratecats=4 statefrequencies=estimate invariantsites=estimate searchpreps=10 bootstrappreps=100
	MrBayes v3.2.6	lset nucmodel=4by4 nst=6 rates=invgamma; mcmc nruns=2 nchains=4 ngen=10000000 printfreq=1000 samplefreq=1000 savebrlens=yes; sumt Burnin=5000
	PHYLIP 3.696	Seqboot: 1000 Bootstrap replicates
Api mtDNA (2)	Garli 2.01	Streefname=LGmodel.mod datatype=aminoacid ratematrix=fixed ratehetmodel=gamma numratecats=4 statefrequencies=estimate invariantsites=estimate searchpreps=10 bootstrappreps=100
	MrBayes v3.2.6	prset aamodelpr=fixed(lg) statefreqpr=fixed(empirical); lset rates=invgamma; mcmc nruns=2 nchains=4 ngen=10000000 printfreq=1000 samplefreq=1000 savebrlens=yes; sumt Burnin=5000
	PHYLIP 3.696	Seqboot: 1000 Bootstrap replicates
Api plDNA	Garli 2.01	datatype=aminoacid ratematrix=jones ratehetmodel=gamma numratecats=4 statefrequencies=empirical invariantsites=none searchpreps=10 bootstrappreps=100
	MrBayes v3.2.6	prset aamodelpr=fixed(Jones) statefreqpr=fixed(empirical); lset rates=gamma; mcmc nruns=2 nchains=4 ngen=10000000 printfreq=1000 samplefreq=1000 savebrlens=yes; sumt Burnin=5000
	PHYLIP 3.696	Seqboot: 1000 Bootstrap replicates

4 Ergebnisse

Ein großer Teil dieser Arbeit bestand in der Etablierung geeigneter Methoden zur Herstellung von genomischen Sequenzen des Parasiten *P. murinus* in hoher Qualität. In Anlehnung an der dafür entwickelten experimentellen Strategie gliedert sich die Ergebnisse dieser Arbeit in mehrere Abschnitte (Abb. 4.1). In den ersten Abschnitten werden die Ergebnisse der Entwicklung von Methoden zur Anreicherung der DNA des Parasiten (Abschnitt 4.1) und die initiale Sequenzierung dieser (Abschnitt 4.2) beschrieben. Gemeinsam mit den genomischen Sequenzen einer nicht infizierten Wirtsfledermaus (Abschnitt 4.3) bilden diese die Grundlage zur Genomresequenzierung des Parasiten (Abschnitt 4.4). Die genomischen Sequenzen von *P. murinus* stellen wiederum den Ausgangspunkt für die weitere molekulare Charakterisierung des Erregers dar (Abschnitt 4.5).

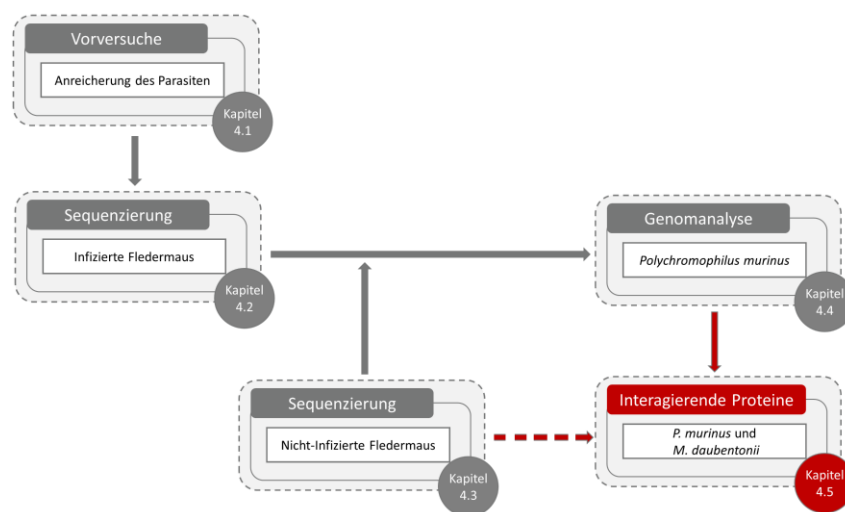


Abb. 4.1: Experimentelle Strategie zur Sequenzierung des Genoms von *P. murinus*.

Schematische Darstellung der methodischen Vorgehensweise dieser Arbeit. Teilergebnisse der einzelnen Bereiche finden sich in den entsprechend aufgeführten Kapiteln. Pfeile markieren Teilergebnisse, die wesentlich zu weiteren Ergebnissen beigetragen haben.

4.1 Vorexperimente zur Anreicherung des Parasiten

4.1.1 Prävalenz und Parasitenintensität von *P. murinus*

Die Isolierung des intrazellulären Parasiten *Polychromophilus murinus* erfolgte aus dem Wirtsorganismus der Wasserfledermaus *Myotis daubentonii*. Zur Identifizierung infizierter Wasserfledermäuse wurden daher Blutproben dieser auf genetisches Material des Parasiten hin untersucht. Unter den 21 untersuchten Wasserfledermäusen sind 17 Individuen (80,9 %), wie in Abschnitt 3.7.1, positiv auf *P. murinus* getestet worden (Abb. 4.2, Übersicht im Anhang Tab. 7.5). Lediglich vier männliche Tiere waren nicht infiziert.

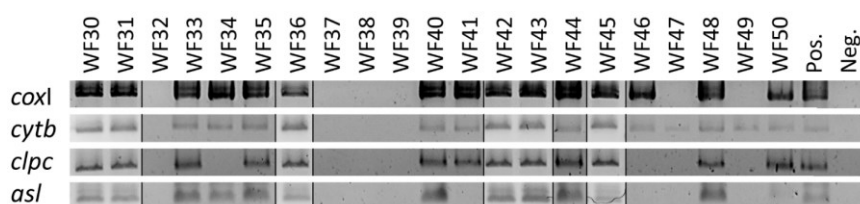


Abb. 4.2: Nachweis der parasitären DNA in den Wasserfledermäusen.

Gelelektrophoretische Auftrennung der nach Martinsen et al. (2008) amplifizierten Genabschnitte. Abkürzungen: asl, Adenylosuccinate Lyase; clpc, Caseinolytic Protease; coxI, Cytochrom Oxidase I; cytb, Cytochrom b. Pos, interne Positivkontrolle; Neg., PCR ohne Template.

Tab. 4.1: Parasitämie der infizierten Fledermäuse.

Fledermaus	Parasitämie*
WF30	0,003 %
WF31	0,003 %
WF35	<0,001 %
WF36	0,006 %
WF42	0,020 %
WF43	0,016 %
WF45	0,010 %
WF48	<0,001 %

* Anteil parasitenhaltiger Erythrozyten

Die mikroskopische Untersuchung der Blutausstriche zeigte dagegen eine weitaus geringere Prävalenz. Infizierte Erythrozyten konnten lediglich in 8 Individuen (38,1 %) nachgewiesen werden. Die Parasitämie dieser Proben (vgl. 3.2) schwankte dabei von einer minimalen Parasitämie an der Nachweisgrenze mit dem Mikroskop (<0,001 %) bis zu 0,02 % (Tab. 4.1). Die Blutausstriche enthielten hauptsächlich reife Gametozyten (Abb. 4.3). Andere Parasiten wie Helminthen aber auch andere Apikomplexa wie Trypanosomen oder Kokzidien konnten in den mit *P. murinus* infizierten Blutausstrichen nicht entdeckt werden.

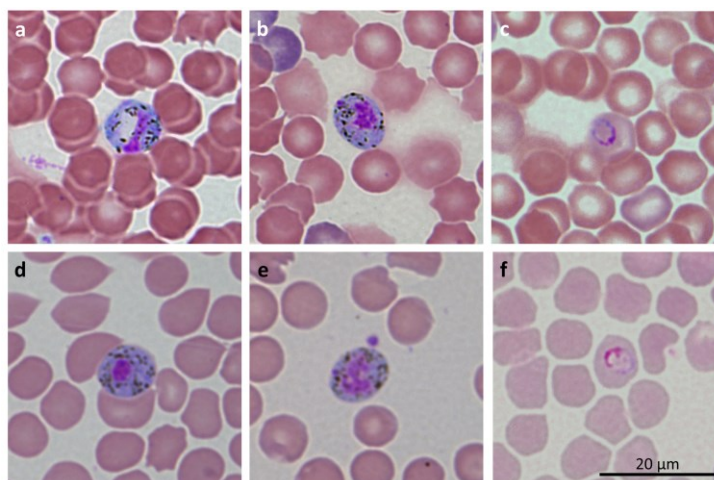


Abb. 4.3: Repräsentative lichtmikroskopische Aufnahmen infizierter Erythrozyten.

Makro- (a, d), Mikrogametozyten (b, e) und Ringstadien aus den Wasserfledermäusen WF43 (a, b, c) und WF42 (d, e, f) bei unterschiedlicher Kontrasteinstellung am Mikroskop. Längenmaßstab: 20 µm.

Um für die weiteren Analysen Proben mit Mehrfachinfektionen auszuschließen, wurden die amplifizierten *cytb*-Fragmente aus den infizierten Wasserfledermäusen, die eine Parasitämie von 0,003 % und mehr enthielten, sequenziert (vgl. 3.10.1). Keine der Sequenzen enthielt mehrdeutige Sequenzen. Alle Sequenzen wiesen die höchste Übereinstimmung zu der von Megali et al. (2011) als Haplotyp 7 beschriebenen *cytb*-Sequenz von *Polychromophilus murinus* auf (Genbank Accession HM055589). Die Sequenzen der Proben WF30, WF42 sowie WF43 unterschieden sich von dieser gar nicht, während die *cytb*-Sequenzen der Proben WF36 sowie WF45 jeweils ein und WF31 drei Nukleotidaustausche enthielt. Die Proben dieser sechs infizierten Fledermäuse wurden für die

weiteren Analysen berücksichtigt. Zusätzlich wurden die entsprechenden Proben nach der *in vitro* Exflagellation hinzugezogen. Diese werden im weiteren Verlauf dieser Arbeit mit Mg gekennzeichnet.

Neben den aus dem Blut der infizierten Wasserfledermäuse gewonnenen Proben sollten für die weiteren Analysen auch Parasiten aus den Fledermausfliegen isoliert werden. Zwölf der gefangenen Wasserfledermäuse (57,1%) waren mit Fledermausfliegen infiziert (Übersicht im Anhang Tab. 7.5). Insgesamt konnten 28 Fledermausfliegen der Art *Nycteribida kolenatii* eingesammelt werden. Ein bis zwei Fledermausfliegen konnten dabei im Schnitt auf einer infizierten Fledermaus gefunden werden. Lediglich die Fledermäuse WF32, WF37 und WF38 wiesen vier oder fünf Fledermausfliegen auf. Oozystenähnliche Strukturen in den Mitteldärmen der Fledermausfliegen konnten nicht festgestellt werden. Lediglich fünf Fledermausfliegen (17,8 %) wurden wie in Abschnitt 3.7.1 positiv auf *Polychromophilus murinus* getestet (nicht dargestellt). Für den weiteren Verlauf wurde nur die genomische DNA der Mitteldärme dieser fünf Fledermausfliegen verwendet (Nyc32, Nyc36, Nyc38, Nyc40, Nyc45) (vgl. 3.3.1).

4.1.2 Qualitätsbeurteilung der WGA-DNA

Aufgrund der geringen DNA-Menge nach der Extraktion, welche teilweise photometrisch nicht nachweisbar war, wurde ausgewählten Proben eine WGA durchgeführt (vgl. 3.5, Abb. 4.4 a,). Sowohl die DNA vor als auch nach WGA zeigte vorwiegend eine Fragmentgröße von deutlich über 10 kbp. Zusätzlich konnte in den Proben eine relativ gleichmäßig verteilte DNA-Spur nachgewiesen werden, die einerseits als Fragmentierung der DNA interpretiert werden könnte, andererseits jedoch aufgrund der Amplifizierung methodisch bedingt zu erwarten war.

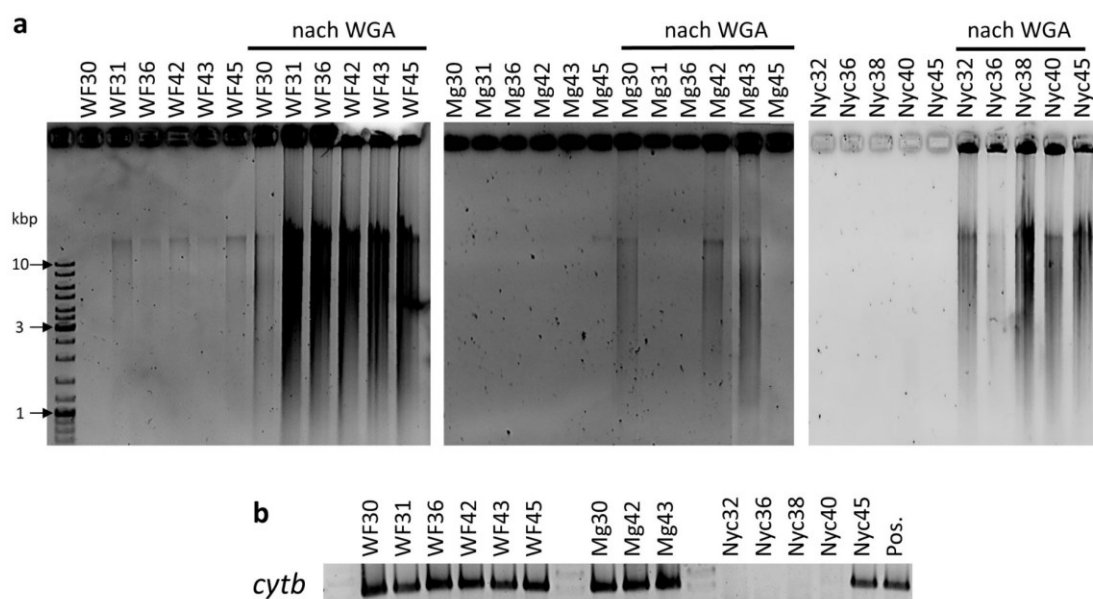


Abb. 4.4: Whole-Genome-Amplifikation (WGA) des Untersuchungsmaterials.

0,7%iges Agarosegel von gDNA vor (5 µl) und nach der WGA (2 µl) (a) sowie Nachweis der parasitären DNA in den WGA-Proben (b). Abkürzungen: *cytb*, Cytochrom b; Mg, DNA des mikrogametenreichen Blut-Überstands der Wasserfledermaus; Nyc, DNA aus dem Mitteldärmen der Fledermausfliegen; Pos., interne Positivkontrolle; WF, DNA aus den aufgereinigten Erythrozyten der Wasserfledermaus; WGA, Whole Genome Amplifikation.

Die Spezifität der WGA-Reaktionen wurde durch Kontroll-PCRs und Re-Sequenzierungen validiert (vgl. 3.7 und 3.10.1). In keiner der Proben unterschieden sich die Sequenzen der amplifizierten Genabschnitte *asl*, *clpc*, *coxI*, und *cytb* nach der WGA von den Sequenzen dieser vor der WGA (nicht dargestellt). Allerdings war nicht mit allen WGA-Proben eine Amplifizierung dieser Genabschnitte problemlos möglich. Die Proben Nyc32, Nyc36, Nyc38 und Nyc40 enthielten deutlich weniger Parasiten-DNA und wurden daher von den weiteren Untersuchungen ausgeschlossen (Abb. 4.4 b).

4.1.3 Quantifizierung des Anteils der Parasiten-DNA nach WGA

Für die Genomsequenzierung des Parasiten war es wichtig eine Probe auszuwählen, die einen möglichst hohen Anteil an Parasiten-DNA im Verhältnis zur Wirts-DNA aufweist. Daher wurde die WGA-DNA wie in Abschnitt 3.7.5 beschrieben untersucht und die Anzahl der Genomkopien von Wirt und Parasit in den unterschiedlichen Proben bestimmt (Abb. 4.6). Der Einfachheit halber wird im Folgenden lediglich von der *Polychromophilus*- bzw. *Myotis*-Kopienzahl gesprochen. Die ermittelten Kopienzahlen unterschieden sich in den einzelnen Proben teilweise recht erheblich. Eine Korrelation der Genomkopienzahl der Parasiten zur Parasitämie der einzelnen Fledermäuse konnte jedoch nicht festgestellt werden (vgl. Tab. 4.1).

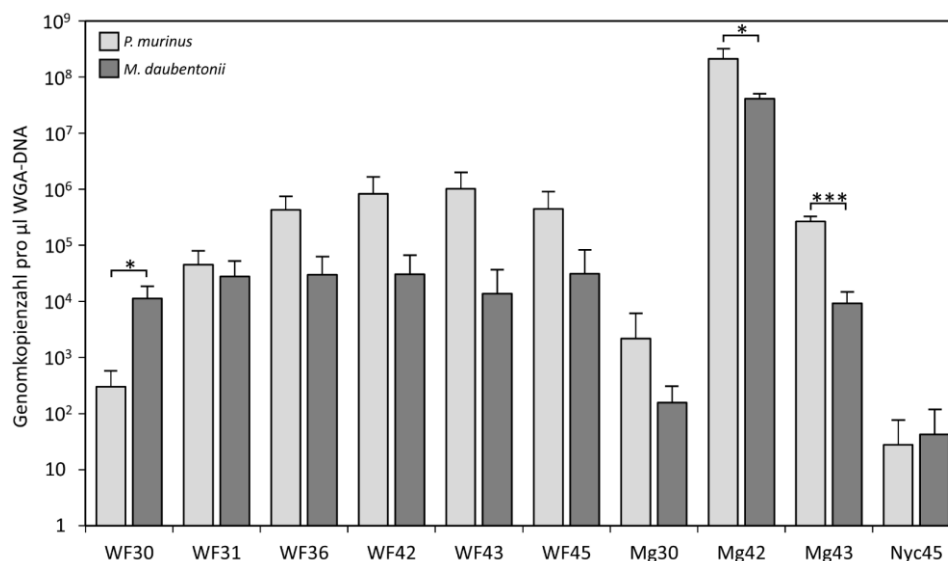


Abb. 4.6: Analyse der DNA-Zusammensetzung des Untersuchungsmaterials.

Gezeigt ist die sf3b1-Kopienzahl, interpretiert als Genomkopienzahl, von *Polychromophilus murinus* und *Myotis daubentonii* in der WGA-DNA. Die Signifikanzen sind durch Sternchen repräsentiert, wobei ein Sternchen (*) einem $p < 0,05$ entspricht und drei Sternchen (***) einem $p < 0,01$ entsprechen. Abkürzungen: Mg, DNA des mikrogametenreichen Blut-Überstands der Wasserfledermaus; Nyc, DNA aus den Mitteldärmen der Fledermausfliegen; WF, DNA aus den aufgereinigten Erythrozyten der Wasserfledermaus; WGA, Whole Genome Amplification.

Mit Ausnahme der Probe WF30, zeigten alle DNA-Proben, die ursprünglich aus den Erythrozyten der Wasserfledermäuse aufgereinigt wurden, eine deutlich höhere Kopienzahl des Parasiten als die des Wirts. Gleichzeitig konnte gezeigt werden, dass durch die *in vitro* Exflagellation der Anteil der DNA parasitären Ursprungs deutlich gesteigert werden konnte. Tatsächlich erwiesen sich die Proben nach *in vitro* Exflagellation im Hinblick auf ihre Zusammensetzung aus DNA parasitären

Ursprungs und wirtsspezifischer DNA auch am günstigsten. Die höchste *Polychromophilus*-Kopienzahl konnte in der Probe Mg42 nachgewiesen werden. Zudem wies die Probe eine signifikant ($p < 0,05$) niedrigere *Myotis*-Kopienzahl auf. Noch deutlich größer war der Unterschied in der Probe Mg43. Hier konnte, auch wenn die Kopienzahlen insgesamt geringer waren als in der Probe Mg42, eine etwa 300fach höhere Kopienzahl von *Polychromophilus* nachgewiesen werden als von *Myotis* ($p < 0,01$). Da für die Genom-Sequenzierung von *P. murinus* nicht die absolute Genomkopienzahl des Parasiten, sondern vielmehr die relative DNA-Menge im Vergleich zur wirtsspezifischen DNA von Bedeutung ist, sollte die Probe Mg43 sequenziert werden.

4.2 Sequenzierung der WGA-DNA mittels Illumina

Die Generierung der Sequenzbibliothek und die anschließende Sequenzierung der WGA-DNA erfolgte wie in Abschnitt 3.10.2 beschrieben durch GATC Biotech (Konstanz). Insgesamt wurden etwa 40,3 Gb Rohdaten bestehend aus etwa 8 Millionen und 130 Millionen Paired-End Reads auf zwei unterschiedlichen Sequenzierplattformen aus derselben Probe generiert (Tab. 4.2). Nach der Aufarbeitung von Read-Abschnitten schlechter Qualität (vgl. 3.12.2) konnten etwa 100 Millionen qualitativ hochwertig Sequenzen für die weitere Analyse genutzt werden (Tab. 4.2).

Tab. 4.2: Zusammenfassende Statistik über Roh- und gefilterte Sequenzierungsdaten.

Parameter	HiSeq 2500 Reads	Miseq Reads	Gesamt
Rohdaten			
Mittlere Readlänge	126	301	290
Anzahl Reads	8.353.922	130.454.474	138.808.396
Anzahl der Nukleotide	1.052.594.172	39.266.796.674	40.319.390.846
Bereinigte Daten*			
Mittlere Readlänge	115	286	272
Anzahl der Reads	7.814.710 (93,5 %)	92.328.656 (70,7 %)	100.143.366 (72,1 %)
Anzahl der Nukleotide	941.324.440	26.431.225.808	27.372.550.248

*Rohdaten nach Qualitätstrimming, Entfernung von Adaptern sowie Vektorsequenzen.

4.2.1 *De novo* Assembly und Identifizierung parasitenspezifischer Contigs

Für eine erste Analyse der erhaltenen Sequenzinformationen wurden die Rohdatensätze zunächst ohne weitere Filterschritte in einem *de novo* Assembly zusammengesetzt (vgl. 3.12.3). Das resultierende Assembly enthielt 400.977 Contigs, 1,0 Gbp Nukleotide und einen GC-Gehalt von 41,7 % (Tab. 4.3).

Tab. 4.3: Statistiken des Rohdaten-Assemblys.

Parameter	Assembly ≥ 1 kbp
Anzahl der Contigs	400.977
Anzahl der Nukleotide	1.005.239.925
Minimale Contiglänge	1000
Maximale Contiglänge	120.446
N50	2.849
Mittlerer GC-Gehalt	41,7 %
Anzahl der Lücken (Gaps)	20.682
Anzahl der Gap-Nukleotide	1.772.184

Die Genome der bislang sequenzierten Haemosporida sind, mit Ausnahmen von *P. vivax* und eng verwandten Parasiten, besonders AT-reich (vgl. Nikbakht et al., 2014 und Galen et al., 2018). Auch die Basenkomposition der bislang sequenzierten Loci genomischen Ursprungs von *Polychromophilus spp.* legen mit einem GC-Gehalt von 28,1 % (Übersicht im Anhang Tab. 7.4) eine ähnliche Genom-Zusammensetzung nahe. Der durchschnittliche GC-Gehalt des Rohdaten-Assemblys von 41,7 % und auch die Verteilung des GC-Gehaltes über alle Sequenzen hinweg (Abb. 4.7) lässt deshalb auf einen sehr hohen Anteil von Sequenzen der Wasserfledermaus schließen. Tatsächlich erscheint auf die Häufigkeitsverteilung des GC-Gehalts dem ersten Blick auch unimodal. Erst bei einer genaueren Betrachtung der Sequenzen mit einem GC-Gehalt zwischen 10 % und 30 % wird die bimodale Verteilung erkennbar. Ein geringer Anteil der Sequenzen weist dabei einen mittleren GC-Gehalt von etwa 20 % auf (Abb. 4.7 b).

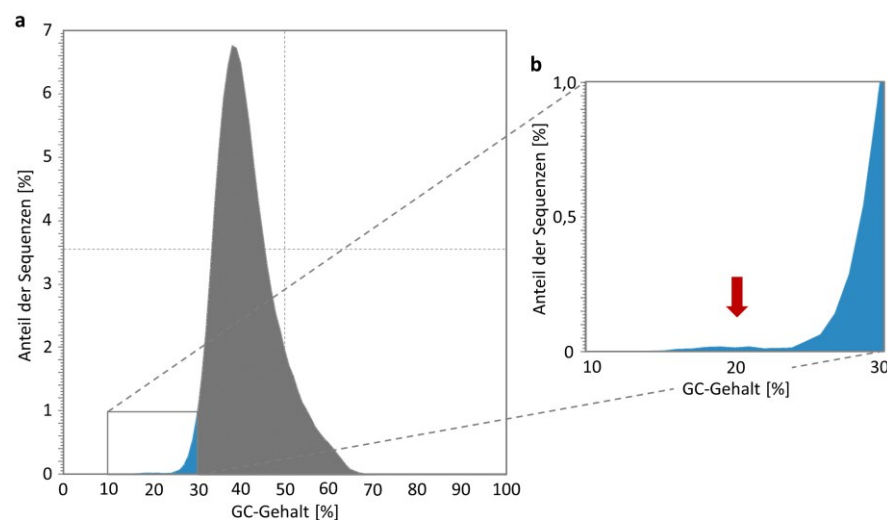


Abb. 4.7: Verteilung des GC-Gehalts im Rohdaten-Assembly.

Dargestellt ist der relative Anteil aller Contigs des Rohdaten-Assemblys (≥ 1 kbp) in Abhängigkeit zum GC-Gehalt (a). Der Anteil der Contigs potenziell parasitären Ursprungs mit einem GC-Gehalt zwischen 10 % und 30 % ist vergrößert dargestellt und durch einen roten Pfeil markiert (b).

Eine Detailanalyse (vgl. 3.13) des Assemblys deutet, in Hinblick auf das Genom von *P. murinus*, ebenfalls auf einen sehr hohen Grad an Kontamination (> 99 %, Abb. 4.8). Neben den Sequenzen der Wasserfledermaus wurde eine Reihe weiterer Contigs identifiziert, die für eine Zusammensetzung des Genoms von *P. murinus* hinderlich sind. Überrepräsentiert waren dabei vor allem Sequenzen aus Vertretern der Bacteroidetes, Proteobacteria, Chlorophyta, Arthropoda und Nematoda. Allerdings konnten die Contigs auch noch 23 weiteren Ordnungen zugeordnet werden. Nur ein sehr geringer Anteil der Sequenzen konnte den Apikomplexa zugeordnet werden. Die Contigs waren dabei wesentlich kürzer als Contigs anderer Organismen und wiesen zusätzlich nur eine sehr geringe Abdeckung im einstelligen Bereich auf. Wie durch die Verteilung des GC-Gehalts bereits vermutet, bilden Contigs, die potenziell Sequenzen des Parasiten darstellen, eine relativ homogene Gruppe mit einem GC-Gehalt von unter 30 % (eingekreist in Abb. 4.8 a).

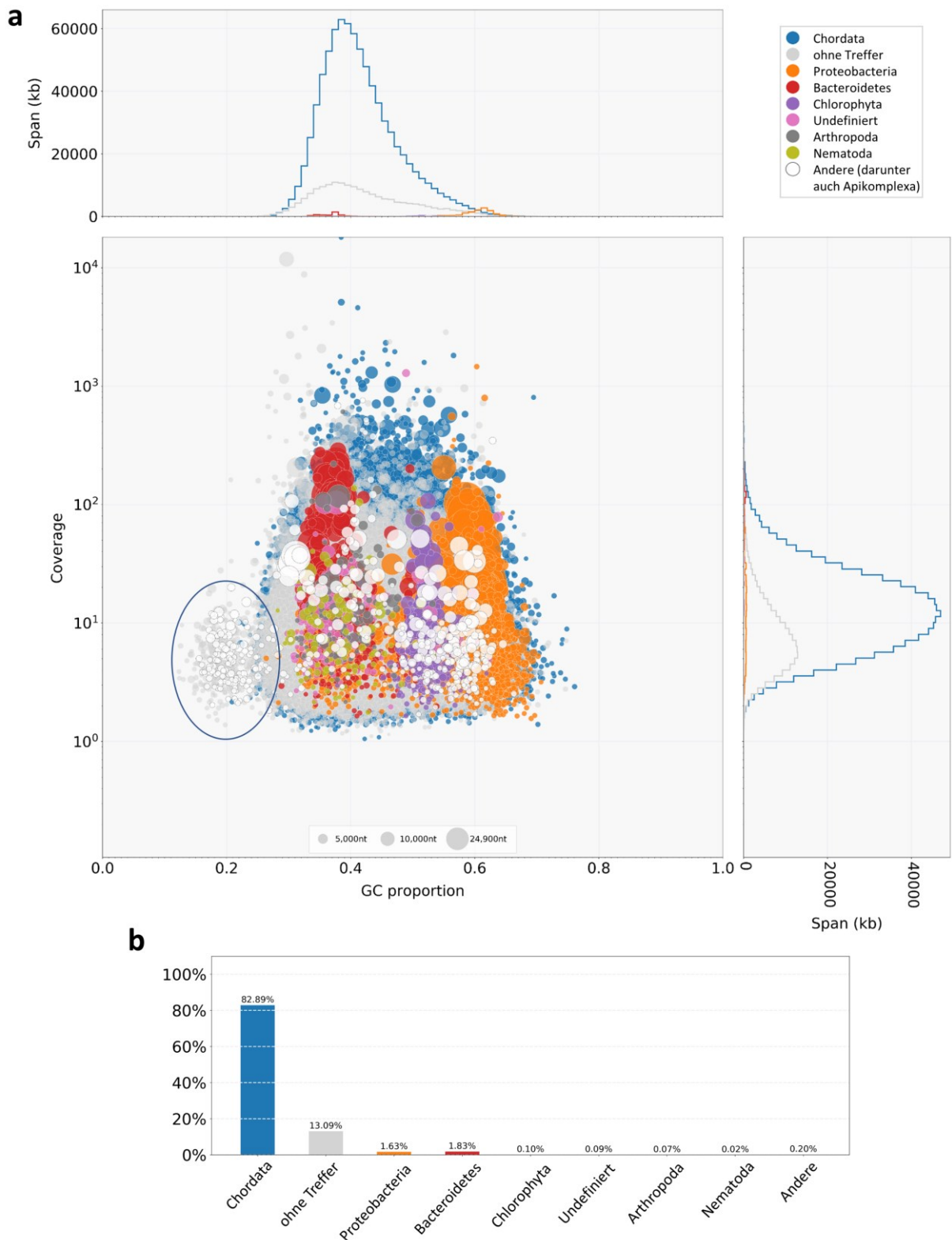


Abb. 4.8: Grad der „Kontamination“ im Assembly der Paired-End-Sequenzierung.

Gezeigt ist der GC-Gehalt der einzelnen Contigs in Abhängigkeit zu ihrer durchschnittlichen Abdeckung (Coverage) (a). Die Sequenzen sind dabei als Kreise dargestellt, deren Durchmesser proportional zur Sequenzlänge skaliert sind. Die taxonomische Zugehörigkeit der Sequenzen (auf der Rangebene "Ordnung"), basierend auf einer blastn-Suche, ist farblich kodiert. Die Histogramme zu Abdeckung und GC-Gehalt sind jeweils an Hand der kumulativen Länge der Sequenzen (Span (kb)) innerhalb eines Taxons gewichtet. Eingekreist: für diese Arbeit interessante Sequenzen. Zusätzlich dargestellt ist der Anteil aller Reads, eingeteilt nach den identifizierten taxonomischen Zugehörigkeiten (b). Die wenigen Contigs mit der Zuordnung Apikomplexa fallen dabei in die Kategorie „Andere“.

Insgesamt wiesen nur 5.608 Contigs des Rohdaten-Assemblys einen GC-Gehalt von $< 30\%$ auf. Der Großteil davon (91 %, 5.094 Contigs) resultierte immer noch aus dem Genom der Wirtsfledermaus (Abb. 4.9). Mit 6 % ist der Anteil der Sequenzen, die signifikante BLAST-Treffer innerhalb von Organismen aus dem Phylum der Apikomplexa hatten (e-value cutoff 10^{-10} , vgl. 3.11.2), dagegen gering. Von den 370 identifizierten Contigs wies die Mehrheit der Sequenzen die höchste Signifikanzidentität zu Organismen der Gattung *Plasmodium* auf (368 Contigs). Insgesamt konnten auf diese Weise 639,8 kbp identifiziert werden, die putativ aus dem Genom des Parasiten *P. murinus* stammten. Geht man bei *P. murinus* von einer ähnlichen Genomgröße, wie die der bislang sequenzierten Haemosporida aus (~ 23 Mbp), entspricht dies etwa 2,8 % des Genoms.

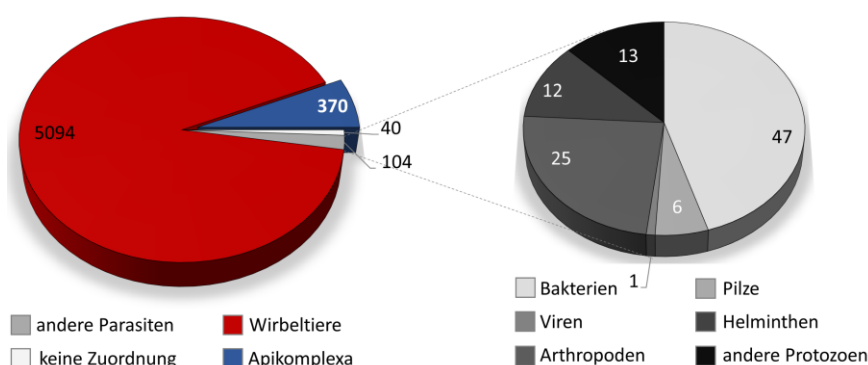


Abb. 4.9: Taxonspezifische Einordnung der Rohdatensequenzen (GC-Gehalt $< 30\%$).

Die Klassifizierung der Sequenzen erfolgte mit der nt NCBI-Protein-Datenbank durch BLAST. Das linke Diagramm zeigt die Einteilung in die Taxa andere Parasiten (grau), Wirbeltiere (rot) und Apikomplexa (blau). Der Anteil ohne Zuordnung (weiß) bezieht sich auf Gene ohne Treffer oder mit signifikanten Treffern in unterschiedlichen Taxa. Rechts sind alle weiterhin identifizierten Sequenzen parasitären Ursprungs weiter aufgeschlüsselt.

4.2.2 Extraktion der Sequenzen parasitären Ursprungs

Für die Extraktion der Sequenzen parasitären Ursprungs wurden, wie in Abschnitt 3.14.1 beschrieben, unterschiedliche Methoden gegenübergestellt. Eine Übersicht der Analysedaten ist in Tab. 4.4 dargestellt. Insgesamt konnten durch den *Myotis*-Filter 16,1 % der Reads identifiziert werden, die nicht dem Wirt zugeordnet werden konnten und damit einen putativ parasitären Ursprung aufweisen können. Die Anzahl der Reads ist damit deutlich größer als die durch den Apikomplexa-Filter (1,23 %) oder den GC-Filter (3,95 %) identifizierten Reads. Allerdings deuten Größe Assemblys (97,04 Mbp), die Anzahl der hier annotierten Gene (20.402) und der relativ hohe GC-Gehalt (46,3 %) auf einen großen Anteil an kontaminierenden Sequenzen.

Tab. 4.4: Vergleich der Readextraktionsmethoden.

Parameter	Myotis-Filter	Apikomplexa-Filter	GC-Filter
Extrahierte Reads	16.113.068 (16,1 %)	1.231.763 (1,2 %)	3.955.662 (3,6%)
Contigs	91.651	1.514	81.890
Assembly Größe (Mbp)	97,04	1,28	0,08
GC-Gehalt (%)	46,3	38,5	28,1
Annotierte Gene	20.402	610	2.561
Pseudogene	1.283	68	138
Nicht-kodierende Gene	417	17	21
GC-Gehalt kodierender Bereiche (%)	48,2	31,9	26,1

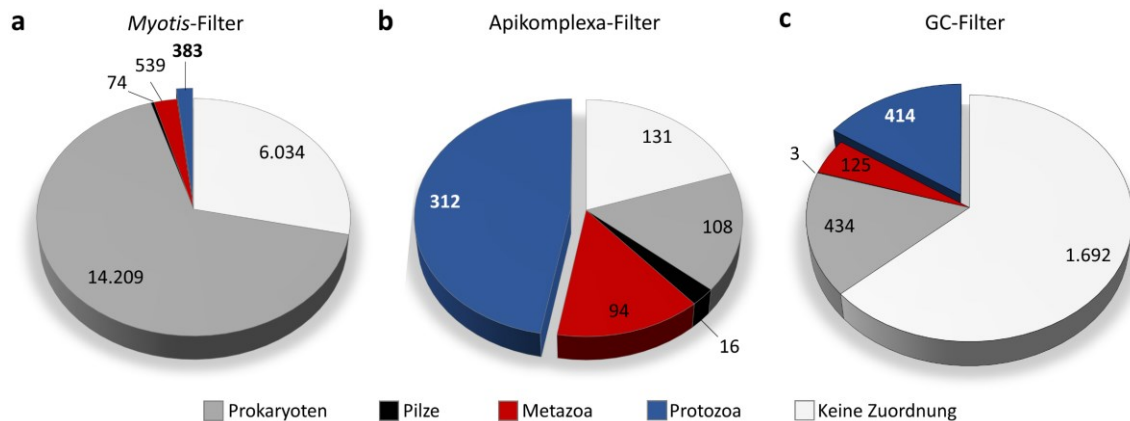


Abb. 4.10: Verteilung identifizierter Gene nach unterschiedlichen Readextraktionsmethoden.

Die Klassifizierung der annotierten Gene aus den Assemblys nach dem Myotis- (a), Apikomplexa- (b), und GC-Filter (c) erfolgte mit der nr NCBI-Protein-Datenbank durch BLAST. Die Diagramme zeigen jeweils die Einteilung in die Taxa Bakterien (grau), Pilze (schwarz), Säugetiere (rot) und Apikomplexa (blau). Der Anteil ohne Zuordnung (weiß) bezieht sich auf Gene ohne Treffer oder mit signifikanten Treffern in unterschiedlichen Taxa.

Tatsächlich konnte im Assembly nach dem *Myotis*-Filter mithilfe einer ausgedehnten Homologiesuche gegen die nr Datenbank des NCBI (vgl. 3.11.2). ein großer Anteil an Genen identifiziert werden, die ihren Ursprung vermutlich in anderen mit sequenzierten Organismen hatten (Abb. 4.10). Obwohl durch die Extraktionsmethode Reads herausgefiltert wurden, die von der Wasserfledermaus stammten, konnten mit 539 Genen mehr Gene in die Gruppe Metazoa als in die Gruppe Protozoa (383) geordnet werden. Die Mehrheit der homologen Sequenzen war hier aus der Klasse der Säugetiere und enthielt neben unterschiedlichen Modellorganismen (z. B. Mensch, Maus, Ratte) auch einen nicht unerheblichen Anteil von homologen Sequenzen aus der Gattung *Myotis* (nicht dargestellt). Gleichzeitig enthielt das Assembly jedoch auch einen besonders großen Anteil an Kontaminationen bakteriellen Ursprungs. Gemessen an der Gesamtzahl der annotierten Gene enthielt das Assembly des Apikomplexa-Filters den höchsten Anteil an Genen parasitären Ursprungs (47,2 %). Insgesamt konnten hier jedoch deutlich weniger Gene annotiert werden (Tab. 4.4). Mit 2.461 annotierten Genen wies das Assembly nach dem GC-Filter viermal mehr Gene auf, als das Assembly nach dem Apikomplexa-Filter (610 Gene), war jedoch mit einer wesentlich höheren Anzahl an Contigs auch deutlich fragmentierter. Auch dies spiegelt sich bei der Klassifizierung der Gene wider. Die Mehrheit konnte dabei gar nicht oder gleich mehreren unterschiedlichen Gruppen zugeordnet werden. Dennoch konnte hier mit 414 Genen eine relativ große Anzahl identifiziert werden, die signifikante Ähnlichkeiten zu Genen aus Protozoa hatten.

Auf diese Weise konnten insgesamt 661 unterschiedliche Gene aus *P. murinus* identifiziert werden. Die Schnittmenge aller drei Extraktionsmethoden betrug dabei 81 Gene. In den Assemblys nach dem *Myotis*-Filter und dem GC-Filter konnten mit 258 Genen parasitären Ursprungs die meisten gemeinsamen Gene identifiziert werden (39 %). Obwohl das Assembly nach dem Apikomplexa-Filter dagegen zahlenmäßig die wenigsten Gene des Parasiten aufwies, waren hier deutlich mehr Gene zu finden, die durch die anderen beiden Methoden nicht identifiziert werden konnten.

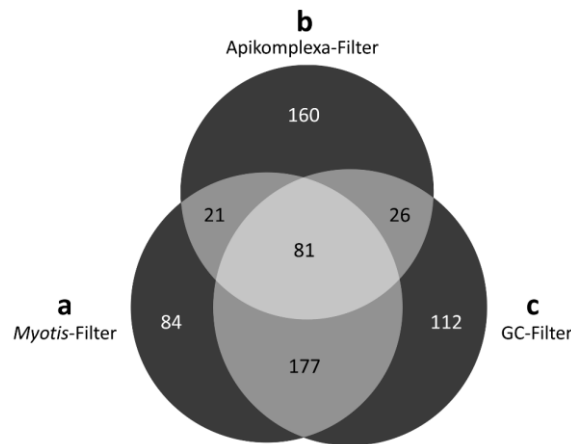


Abb. 4.11: Gemeinsame Gene nach *Myotis*- (a), Apikomplexa- (b), und GC-Filter (c).

Gemeinsame Gene wurden durch identische BLAST-Treffer charakterisiert. Fragmentierte Gene wurden nur einmal gezählt.

4.2.3 Phylogenetische Position von *P. murinus* als Schwesterngruppe zu *Plasmodium*

Unter den hier erstellten Proteinsequenzen von *P. murinus* konnten 315 Proteine identifiziert werden, die orthologe Sequenzen in 18 weiteren Haemospiriden und 7 Piroplasmiden aufwiesen und als Gen in Einzelkopie kodiert werden. Nach Entfernung schlecht konservierter Bereiche (vgl. 3.20) ergab sich daraus ein Proteinalignment mit 108.862 Aminosäurepositionen, das zur phylogenetischen Analyse mittels ML-Algorithmen verwendet wurde (vgl. 3.20). Die Sequenzen der Gattungen *Babesia* und *Theileria* dienen in der Topologie als Außengruppe.

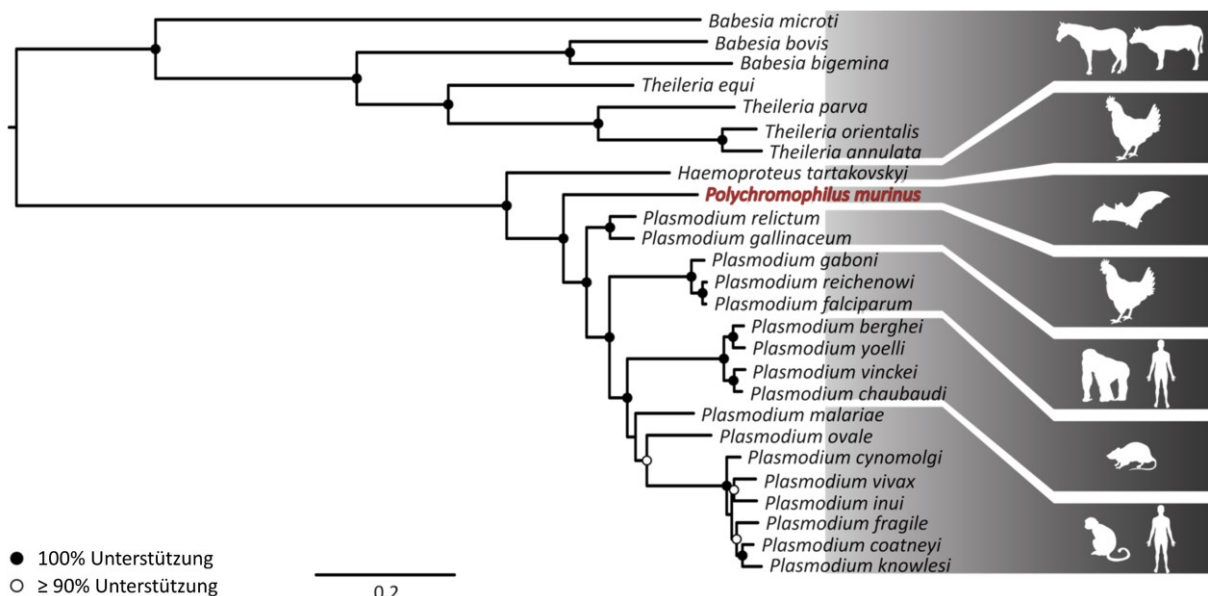


Abb. 4.12: Phylogenetische Beziehungen der Haemosporida basierend auf Genomdaten.

Die dargestellte Topologie basiert auf einer ML-Analyse. Die Sequenzen der Gattungen *Babesia* und *Theileria* dienen in der Topologie als Außengruppe. Die Verzweigungslängen sind proportional zur durchschnittlichen Anzahl der Substitutionen. Die Länge des Balkens entspricht 0,2 Substitutionen pro Position. Verzweigungen mit Unterstützungswerten größer als 90 % wurden durch einen weißen Punkt, Unterstützungswerten von 100% durch einen schwarzen Punkt markiert. Rot hervorgehoben: in dieser Arbeit sequenziertes Genom.

Die daraus resultierende Phylogenie ist weitgehend kongruent mit der vielfach akzeptierten Topologie der Apikomplexa (Abb. 4.12). Dabei zweigt *H. tartakovskyi*, stellvertretend für die Gattung *Haemoproteus*, basal von den Vertretern der Gattungen *Plasmodium* und *Polychromophilus* ab. Die Spezies der Gattung *Plasmodium* bilden ein sehr gut abgesichertes Monophylum, das mit einem maximalen Unterstützungswert eine Schwesterngruppe mit dem hier untersuchten Erreger bildet. Innerhalb der Gattung *Plasmodium* ist eine deutliche Gliederung zu erkennen, welche die Einteilung der Parasiten in die entsprechenden Untergattungen (*Haemamoeba*, *Laverania*, *Vinckeia*, *Plasmodium*) widerspiegelt. Mit Ausnahme der Verzweigungen innerhalb der Untergattung *Plasmodium* sind diese allesamt sehr gut unterstützt. Die wenig unterstützten Verzweigungen innerhalb der Untergattung *Plasmodium*, deuten auf eine nicht-konfliktfreie Auflösung der Topologie hin. Die hier implizierte monophyletische Gruppierung wurde, insbesondere in Hinblick auf die phylogenetische Stellung der Art *Plasmodium malariae*, auch bereits in einer anderen Arbeit in Frage gestellt (Böhme et al., 2018).

4.3 Sequenzierung des Wirtsgenoms

Zur Sequenzierung der Wasserfledermaus wurde eine der vier Proben ausgewählt, in denen keine Infektion mit *P. murinus* nachgewiesen wurde (WF32, vgl. Abb. 4.2). Die Sequenzierung erfolgte mit der MiSeq Sequenzieretechnologie (Illumina, San Diego, US) durch GATC Biotech (Konstanz) und produzierte 18,8 Gb bestehend aus etwa 63 Millionen Paired-End Reads mit einer maximalen Readlänge von 300 bp (vgl. 3.10.3). Nach Aufbereitung der Rohdaten (Abschnitt 3.13.2) wurden 55,5 Millionen qualitativ hochwertige Sequenzen für die weitere Analyse genutzt (Tab. 4.5).

Tab. 4.5: Statistik der Sequenzierungs- und Mappingdaten von *M. daubentonii*.

Parameter			
Sequenzierung	Rohdaten	Bereinigte Daten*	
Mittlere Readlänge	300	284	
Anzahl Reads	62.974.800	55.508.094 (88,14 %)	
Anzahl der Nukleotide	18.892.440.000	15.764.298.696	
Mapping	M. brandtii	M. davidii	M. luzifugus
Mapped Reads	42.695.135 (76,92 %)	42.425.836 (76,43 %)	42.515.467 (76,59 %)
1x Genome Coverage	46,50 %	63,86 %	95,77%
5x Genome Coverage	11,93 %	27,19 %	40,34%
10x Genome Coverage	2,58 %	7,03 %	5,83 %
1x CDS Coverage	92,57 %	93,23 %	99,45 %
5x CDS Coverage	67,62 %	66,17 %	72,78 %
10x CDS Coverage	0,95 %	1,17 %	0,90 %
Assembly	Gesamtes Assembly	Assembly ≥ 3 kbp	
Anzahl der Contigs	451.085	69.598	
Anzahl der Nukleotide	942.349.704	300.936.503	
Minimale Contiglänge	566	3.000	
Maximale Contiglänge	33.075	33.075	
N50	2.228	4.205	
Mittlerer GC-Gehalt	40,2 %	39,4 %	

*Rohdaten nach Qualitätstrimming, Entfernung von Adaptern sowie Vektorsequenzen.

Für eine erste Analyse der erhaltenen Sequenzinformationen wurden initiale Mappings gegen die veröffentlichten Genome der Gattung *Myotis* (*M. brandtii*, *M. davidii*, *M. luzifugus*) durchgeführt (vgl. 3.12.4). Etwa 76 % der Reads konnten dabei jeweils an die Referenzgenome positioniert werden

(Tab. 4.5). Die Abdeckung des Genoms variierte jedoch stark. Während etwa 46,5 % des Genoms von *M. brandtii* mit mindestens einem Read abgedeckt wurde (Coverage), konnte die höchste Coverage durch das Mapping gegen *M. lucifugus* erreicht werden. Hier wurde etwa 95,8 % des Genoms durch mindestens einen Read abgedeckt; 40,3 % durch mindestens 5 Reads und 5,8 % wies eine 10fache Coverage auf. Insgesamt konnten durch das Mapping gegen *M. lucifugus* 19.288 Gene (653.561.471 bp) von *M. daubentonii* identifiziert werden. Ein Großteil dieser (7.784 Gene; 40,4 %) waren jedoch hochgradig fragmentiert und wiesen mit insgesamt 19.881 (14.133.629 bp) eine hohe Anzahl an Lücken auf.

Die Ergebnisse der ersten Mappings versprechen bereits eine mehr als ausreichend hohe Qualität der Sequenzierung. Allerdings war das erste Assembly mit 0,9 Gbp nur etwa halb so groß wie erwartet, wenn von einer ähnlichen Genomgröße wie die der anderen *Myotis*-Arten ausgegangen wird (~2,2 Gbp) (Kasai et al., 2013; Seim et al., 2013; Bhak et al., 2017). Der GC-Gehalt von 40,2 % war mit dem der anderen *Myotis*-Arten dagegen gut vergleichbar (40-42 %). Tatsächlich konnten nur ~1 % der Contigs als Kontaminationen durch andere Arten identifiziert werden. Im Vergleich zum Genom der infizierten Fledermaus (vgl. 4.2.1) waren diese zudem deutlich weniger divers. Neben den Phyla Bacteroidetes, Proteobacteria, Arthropoda, Nematoda und Tenericutes konnten nur noch fünf weitere bakterielle Ordnungen den Contigs zugewiesen werden.

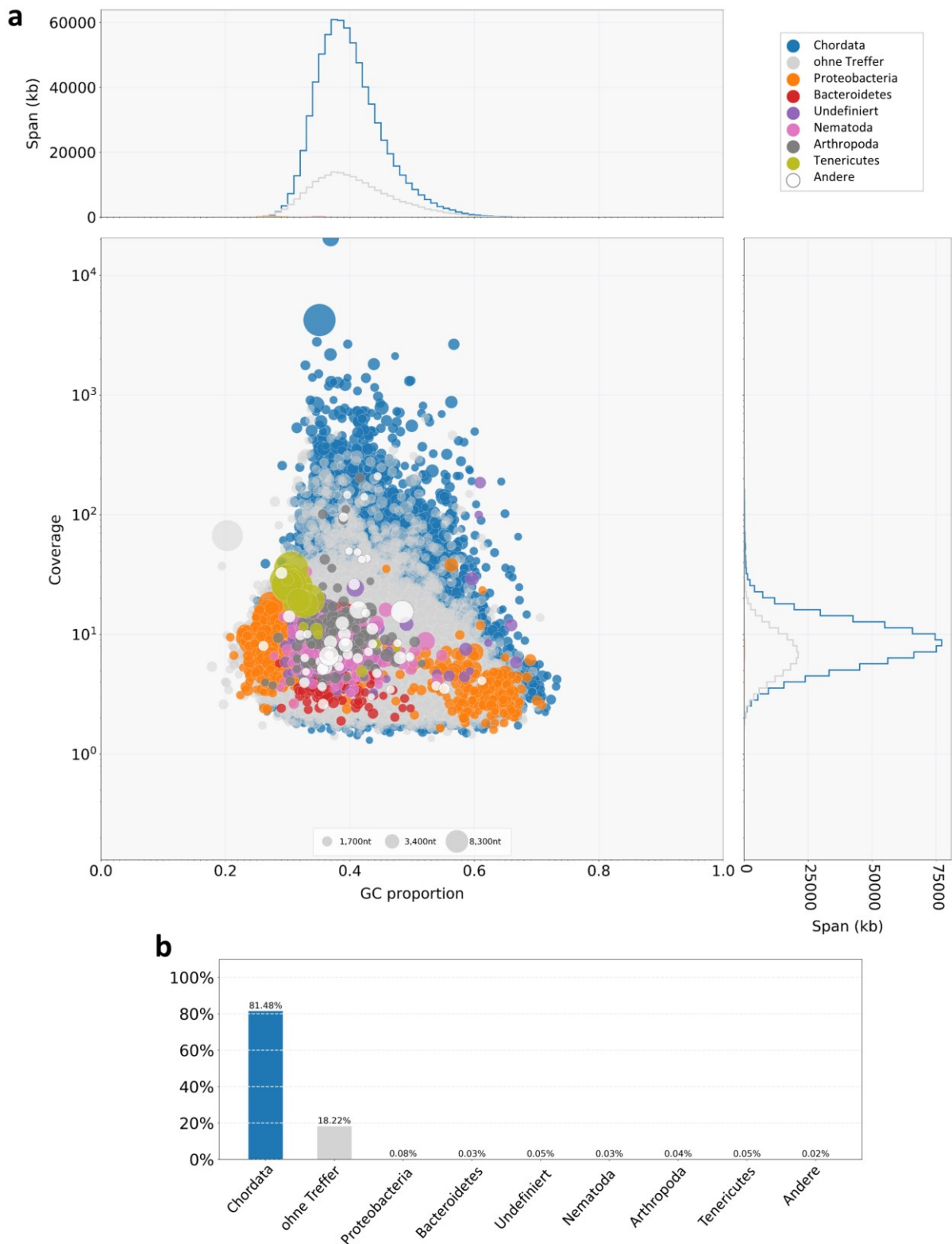


Abb. 4.13: Grad der Kontamination bei der Sequenzierung der nicht infizierten Fledermaus.

Gezeigt ist der GC-Gehalt der einzelnen Contigs in Abhängigkeit zu ihrer durchschnittlichen Abdeckung (Coverage) (a). Die Sequenzen sind dabei als Kreise dargestellt, deren Durchmesser proportional zur Sequenzlänge skaliert sind. Die taxonomische Zugehörigkeit der Sequenzen (auf der Rangebene "Ordnung"), basierend auf einer blastn-Suche, ist farblich kodiert. Die Histogramme zu Abdeckung und GC-Gehalt sind jeweils an Hand der kumulativen Länge der Sequenzen (Span (kb)) innerhalb eines Taxons gewichtet. Zusätzlich dargestellt ist der prozentuale Anteil der Reads eingeteilt nach den identifizierten taxonomischen Zugehörigkeiten (c).

4.3.1 Die mitochondriale DNA von *M. daubentonii*

Die mitochondriale DNA der Wasserfledermaus WF32 besteht aus 17.149 bp und weist einen GC-Gehalt von 35,6 %. Insgesamt konnten auf der mtDNA 13 proteinkodierende Gene, 22 tRNA-Gene und 2 rRNA-Gene (12S und 16S) annotiert werden (Abb. 4.15). Alle charakteristischen Merkmale, wie die Organisation und die Genanordnung, sind dabei identisch zu denen eines typischen mitochondrialen Genoms eines Wirbeltiers (vgl. Pereira, 2000). Der in Abb. 4.15 abgebildete forward Strang, setzt sich aus 34,2 % A, 30,5 % T, 22,6 % C und 12,7 % G zusammen und entspricht damit dem historisch bedingt als H (heavy) bezeichneten purinreichen Strang. Mit Ausnahme des Gens für die NADH Dehydrogenase Untereinheit 6 (*nad6*) sind alle proteinkodierenden Gene, sowie die beiden rRNA-Gene ausschließlich auf dem H-Strang kodiert.

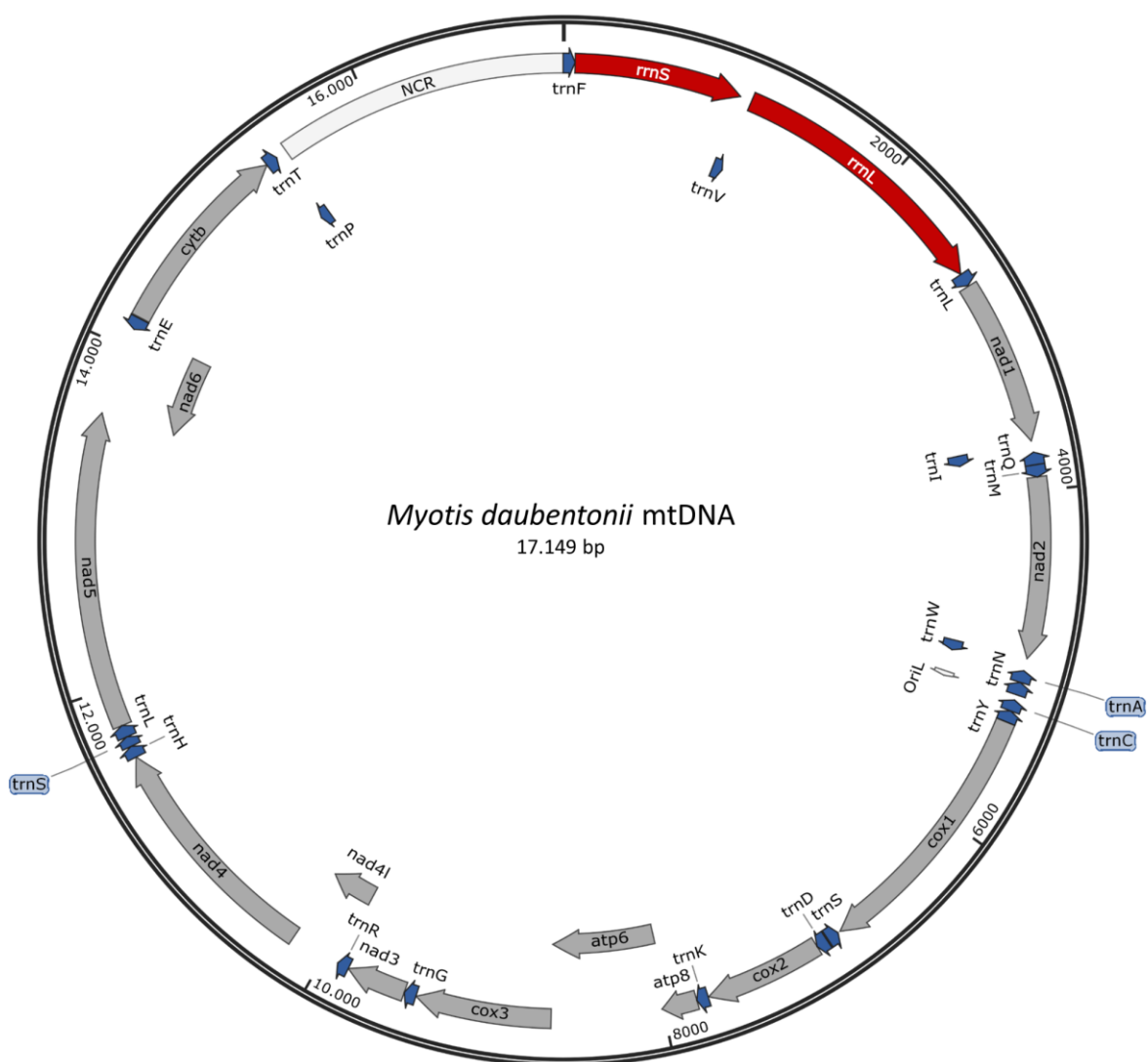


Abb. 4.15: Organisation der mtDNA von *M. daubentonii*.

Das zirkuläre Molekül besteht aus zwölf Genen für Proteine (graue Pfeile), zwei rRNAs (rote Pfeile) und 22 tRNAs (blaue Pfeile, Großbuchstabe steht jeweils für die entsprechende Aminosäure). Die Orientierung der Pfeile repräsentiert die Kodierung auf dem H- (Uhrzeigersinn) bzw. L-Strang (entgegen des Uhrzeigersinns) der DNA. Abkürzungen: *atp*, ATPase Untereinheit; *cox*, Cytochrom Oxidase; *cytb*, Cytochrom b; NCR, non coding region (Kontrollregion der mtDNA); *nad*, NADH Dehydrogenase Untereinheit; OriL, Origin of Replication (L-Strang); *rrn*, rRNA; *trn*, tRNA.

Die zwischen mitochondrialer und nukleärer DNA abweichende Kodierung wird insbesondere bei den Start- und Stopkodons der proteinkodierenden Gene deutlich. Während neun proteinkodierenden Gene mit dem typischen ATG beginnen (Übersicht im Anhang: Tab. 7.7), weisen vier der Strukturgene (*nad2*, *nad3*, *nad5* und *nad6*) ATT als Startkodon auf. Als Stopkodon konnte bei jeweils zwei Genen (*cox1*, *nad2*) das Kodon TAG und bei sieben Genen (*atp6*, *atp8*, *cox2*, *nad1*, *nad4l*, *nad5*, *nad6*) das Kodon TAA identifiziert werden. Das Kodon TGA kodiert anstelle eines Stopkodons dagegen für die Aminosäure Tryptophan und wird 99-mal innerhalb der Leseraster aller Gene verwendet. Das Gen *cytb* terminiert dagegen mit dem alternativen Basentriplett AGA. Bei den restlichen drei Strukturgenen konnten durch Sequenzvergleiche mit annotierten Genen anderer mtDNAs die unvollständigen Stopkodons T* (*cox3*, *nad4*) oder TA* (*nad3*) identifiziert werden.

Wie für andere mitochondriale Genome aus Wirbeltieren typisch, liegt die mtDNA der Wasserfledermaus in einer sehr kompakten Form vor. Die mtDNA weist keine Introns auf und die Gene sind teilweise nur durch ein oder zwei Nukleotide separiert. In einigen Fällen überlappen sich zudem die Leseraster benachbarter Gene (*atp8* und *atp6*, *nad4l* und *nad4* sowie *nad5* und *nad6*). Mit 90,19 % ist der kodierende Anteil der mitochondrialen DNA entsprechend hoch. Insgesamt weist die mtDNA von *M. daubentonii* zwei nichtkodierende Regionen auf. Der Replikationsursprung des L-Strangs (OriL) ist 33 bp groß und innerhalb der sogenannten WANCY-Region lokalisiert, die fünf tRNA-Gene enthält (tRNA^{Trp}, tRNA^{Ala}, tRNA^{Asn}, tRNA^{Cys} und tRNA^{Tyr}). Der Replikationsursprung des H-Strangs (OriH) konnte nicht genauer lokalisiert werden, befindet sich jedoch innerhalb der zweiten nichtkodierenden Region (NCR). Diese ist zwischen den tRNA-Genen tRNA^{Pro} und tRNA^{Phe} lokalisiert und erschreckt sich über 1.686 bp. Neben dem OriH enthält diese hauptregulatorische Elemente für die Transkription beider Stränge, sowie den D-Loop (displacement loop, Verdrängungsstrang). Die genaue Position dieser Sequenzabschnitte konnte jedoch nicht weiter lokalisiert werden. Innerhalb der NCR konnte jedoch ein Motiv aus 6 bp identifiziert werden (5'-GCATAC-3'), das 40-mal zwischen den konservierten Sequenzblöcken (conserved sequence block) CSBI (5'-TTAATGGTCACAGGACATAA-3') und CSBII (5'-AAACCCCCCTTACCCCCCG-3') vorlag. Zusätzlich konnten fünf Kopien einer 81 bp langen Sequenz identifiziert werden (5'-GTACATTAATTTATTTCCACATGAATATTAAGCATGTACATAYTTATRTTAAYATTACATAATACATAWYATGYRTRATC-3'). Ähnliche sich wiederholende, speziesspezifische Einheiten wurden auch bereits in anderen mtDNAs der Gattung *Myotis* beschrieben (vgl. Kim et al., 2015).

4.3.2 Phylogenetische Position von *M. daubentonii*

Die Rekonstruktion der internen Verwandtschaftsverhältnisse der Gattung *Myotis* basierte auf Grundlage der 13 proteinkodierenden Gene der mtDNA (vgl. 3.20). Zum Zeitpunkt der Erstellung dieser Arbeit lagen bereits 28 sequenzierte mitochondriale Genome der Gattung *Myotis* vor (Übersicht im Anhang: Tab. 7.8). Bei den Analysen dienten jeweils die homologen Sequenzen von *Murina leucogaster* und *Murina ussuriensis* als Außengruppe. Die Analysen führten, mit Ausnahme von einigen Knoten bei der MP-Analyse, zu sehr hohen Unterstützungswerten, wiesen jedoch

untereinander teilweise Konflikte in den unterstützten Topologien auf (vgl. Anhang: Abb. 7.2-7.3). Während sich die Topologien nach BI- und ML-Algorithmen in nur wenigen Punkten unterscheiden, weist das Ergebnis der MP-Analyse eine hohe Anzahl an Beziehungen auf, die dem widersprechen. Zur besseren Übersicht zeigt Abb. 4.16 den Konsensus-Baum nach BI-Analyse. Aufspaltungen, die durch alle Analysen unterstützt werden, sind hervorgehoben.

In sämtlichen abgeleiteten Topologien ist innerhalb der Gattung *Myotis* eine Gliederung in zwei Gruppen klar erkennbar. Die statistische Unterstützung für diese Aufspaltung war bei allen durchgeführten Analysen maximal. Abgesehen von *M. atacamensis* und *M. brandtii* teilt sich die Gattung damit in biogeographisch differenzierte Gruppen auf („Altwelt“- und „Neuwelt“-*Myotis*). Innerhalb der Gruppe der „Altwelt“-*Myotis*, zu denen auch die hier untersuchte Wasserfledermaus gehört, lassen sich die Verwandtschaftsverhältnisse zwischen den Arten relativ schlecht auflösen. In allen Topologien sind die paläarktischen Fledermäuse gegenüber den orientalischen *Myotis*-Arten paraphyletisch. Lediglich die Schwesternverhältnisse einzelner Arten finden sich gleichermaßen in den Topologien aller Analysen. Die Beziehungen dieser Gruppen zueinander können aber mit dem vorliegenden Datensatz nicht mit Sicherheit aufgelöst werden.

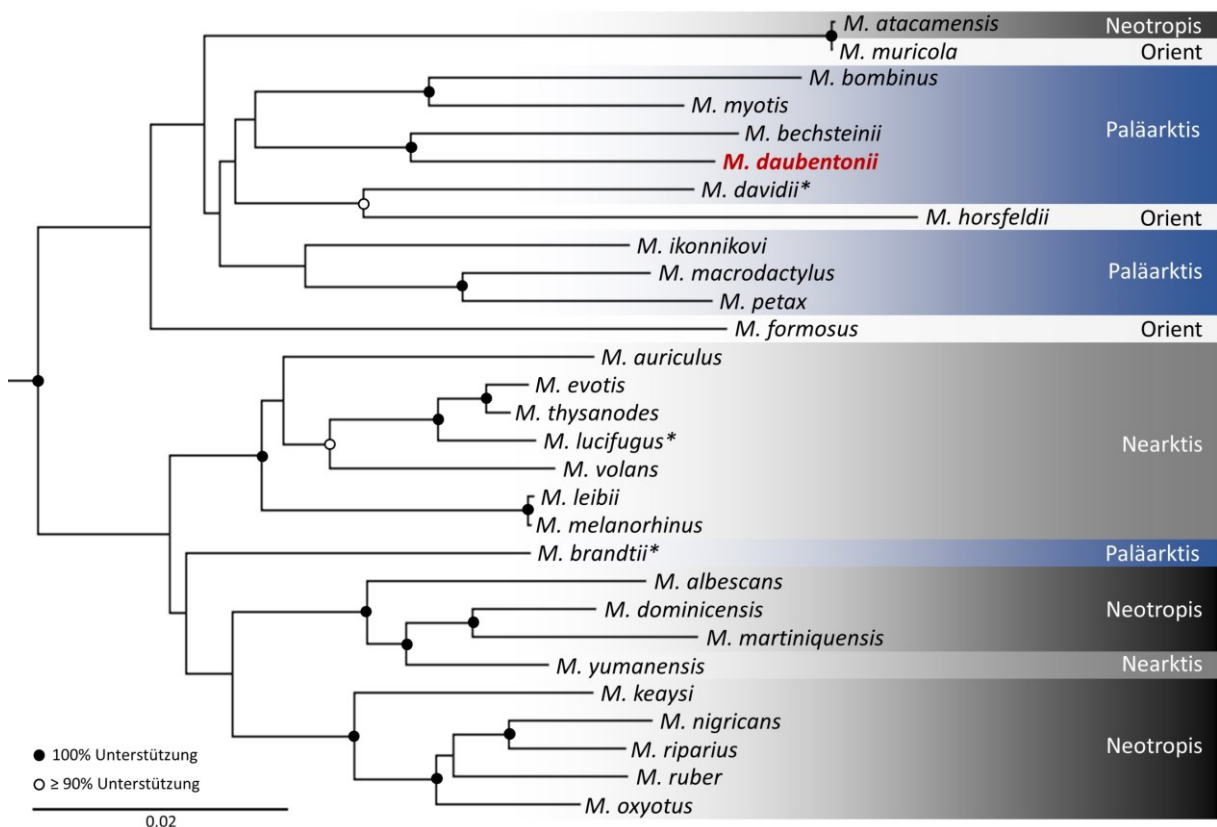


Abb. 4.16: Phylogenie von Fledermäusen der Gattung *Myotis*.

Die dargestellte Topologie basiert auf einer MrBayes-Analyse. Die Verzweigungslängen sind proportional zur durchschnittlichen Anzahl der Substitutionen. Die Länge des Balkens entspricht 0,02 Substitutionen pro Position. Verzweigungen mit Unterstützungswerten größer als 95 % aus ML-, MP- und den Mr-Bayes-Analysen wurden durch einen weißen Punkt, Unterstützungen von 100% durch einen schwarzen Punkt markiert. Die farbliche Hinterlegung repräsentiert die biogeographische Herkunft der Fledermäuse. Sternchen: Fledermäuse, deren Genom bislang sequenziert wurde; rot hervorgehoben: in dieser Arbeit sequenzierte mtDNA.

Auch wenn die Verwandtschaftsverhältnisse innerhalb der Gattung durch die durchgeführten Analysen nicht im Detail gelöst werden können, können aus den unterschiedlichen Topologien klar aufgegliederte Abspaltungsfolgen abgeleitet werden, die nicht mit der traditionellen Unterteilung der Gattung in Speziesgruppen übereinstimmt. Wie bereits durch die Blast-Analysen zu erwarten, bilden *M. daubentonii* und *M. bechsteinii* eine statistisch gut unterstützte monophyletische Gruppe. Während *M. bechsteinii* aufgrund äußerlicher Merkmale zu der Gruppe *Mystacinus* gezählt wird, wird *M. daubentonii* traditionell in die Untergattung *Leuconoe* gruppiert. Die Art *M. petax* dagegen, die lange auch zu der Art *M. daubentonii* gezählt wurde und von dieser morphologisch kaum zu unterscheiden ist (Matveev et al., 2005), scheint mit *M. daubentonii* nicht näher verwandt zu sein. Von den bislang sequenzierten Fledermäusen scheint *M. daubentonii* zusätzlich am ehesten mit der Fledermaus *M. davidii* verwandt zu sein. Traditionell wird *M. davidii*, wie auch *M. bechsteinii*, zu der Untergattung *Mystacinus* gezählt, und gehört durch alle phylogenetischen Rekonstruktionen gut unterstützt zu der Gruppe der „Altwelt“-*Myotis*. *M. lucifugus* (Untergattung *Leuconoe*) und *M. brandtii* (Untergattung *Mystacinus*) haben dagegen einen statistisch gut belegten Ursprung in der Gruppe der „Neuwelt“-*Myotis*, obwohl *M. brandtii* tatsächlich biogeographisch zu den paläarktischen Fledermäusen gehört.

4.3.3 Das Genom der Wirtsfledermaus

Aufgrund der näheren Verwandtschaft zwischen der Wasserfledermaus und der David-Langfußfledermaus (vgl. 4.3.2), wurde zur Orientierung der Contigs von *M. daubentonii* *M. davidii* als Referenz herangezogen (vgl. 3.15.2). Vor dem referenzbasierten Assembly wurde das Genom der Wasserfledermaus von 1,39 Mbp verteilt auf 946.963 Sequenzen repräsentiert (Tab.4.6). Durch die Positionierung am Referenzgenom konnte die Anzahl der Sequenzen auf 459.518 reduziert werden (vgl. 3.15.2). Dabei wurden 492.955 Contigs zu 5.510 Scaffolds mit einer Größe von bis zu 13,9 Mbp zusammengefasst und insgesamt 477.270 Lücken zwischen den Contigs neu eingeführt.

Tab. 4.6: Statistiken zur Herstellung des ersten Genom-Entwurfs von *M. daubentonii*.

Parameter	Rohdaten-Assembly	Referenz-basierte Contigs	Finales Genom
Anzahl der Sequenzen	946.963	459.518	459.518
Anzahl der Nukleotide	1.389.784.001	1.438.518.501	1.441.045.229
Minimale Sequenzlänge	500	500	500
Maximale Sequenzlänge	33.075	13.865.009	13.867.220
N50	1.712	1.396.960	1.399.523
Anzahl der Lücken	362	487.785	477.270
Mittlerer GC-Gehalt	40,4 %	40,4 %	40,4 %

Bislang wurde noch kein Vertreter der Gattung *Myotis* vollständig auf Chromosom-Ebene sequenziert worden. Ausgehend von einer ähnlichen Genomgröße, wie die der anderen *Myotis*-Arten (~2,2 Gbp), entsprechen die genomischen Sequenzen von *M. daubentonii* etwa 65 % des gesamten Genoms (Kasai et al., 2013; Seim et al., 2013; Bhak et al., 2017). Damit repräsentieren die hier erstellten genomischen Sequenzen im Vergleich zu den Sequenzen anderer Vertreter aus der Gattung *Myotis* einen deutlich geringeren Anteil der Gesamtsequenz (Tab. 4.7). Mit 459.518

Scaffolds ist das Genom der Wasserfledermaus zudem deutlich fragmentierter als die Genome von *M. davidii*, *M. brandtii* oder *M. lucifugus*. Lediglich das kürzlich sequenzierte Genom von *Myotis myotis* ist, obwohl es mit 2,3 Gbp einen deutlich größeren Anteil des gesamten Genoms repräsentiert, ebenfalls stark fragmentiert.

Dass ein großer Teil an Regionen nur unzureichend bzw. gar nicht durch die genomischen Sequenzen repräsentiert werden, spricht für eine geringe statistische Redundanz im ursprünglichen Datensatz. Aufgrund der geringen Anzahl von 55,5 Millionen qualitativ hochwertigen Reads war dies allerdings zu erwarten. Tatsächlich wurde jede Base mit einer 7,8fachen Coverage stärker abgedeckt, als durch die Rohdaten impliziert (6,2fach). Der Datensatz enthielt jedoch auch 1.121 Sequenzen (1,1 Mbp), die eine Coverage von 100-2.398 aufwiesen und teilweise unterschiedlich zusammengesetzte repetitive Bereiche enthielten (nicht dargestellt). Daher lag die Vermutung nahe, dass es sich zum Teil um Sequenzen handeln musste, die vermehrt im Genom positioniert vorliegen und daher nicht korrekt erweitert und zugeordnet werden konnten. Interessanterweise wiesen diese Sequenzen mit 44,1 % auch einen relativ hohen GC-Gehalt auf. Im Schnitt lag der GC-Gehalt des ersten Entwurfs des Genoms von *M. daubentonii* mit 40,4 % deutlich unter dem GC-Gehalt der Genome der anderen Vertreter der Gattung *Myotis* (42,5-43,2 %).

Tab. 4.7: Statistiken zum Entwurf des Genoms von *M. daubentonii* im Vergleich.

Parameter	<i>M. davidii</i>	<i>M. brandtii</i>	<i>M. lucifugus</i>	<i>M. myotis</i>	<i>M. daubentonii</i>
Gesamtgröße	2.059.799.708	2.107.242.811	2.034.575.300	2.294.472.406	1.441.045.229
Größe ohne Lücken	1.878.461.987	1.981.769.726	1.966.419.868	2.293.659.106	1.385.078.613
Anzahl der Scaffolds	101.769	169.750	11.654	1.045.197	459.518
Scaffold N50	3.454.484	3.225.832	4.293.315	21.824	1.399.523
Scaffold L50	163	181	106	27.249	225
Anzahl der Contigs	325.280	325.414	72.785	1.053.330	946.963
Contig N50	15.182	23.289	64.330	19.322	1.712
Contig L50	36.107	24.039	8.574	31.447	247.936
Abdeckung/Coverage	110x	120x	7x*	47,8x	7,8x
GC-Gehalt	42,7	42,6	42,5	43,2	40,4

*Sanger-Sequenzierung

Repetitive Sequenzen, die sich aus bekannten Transposons ableiten, repräsentieren etwa 29,2 % des Genoms (Abb. 4.17; vgl. Tab 7.9). Im Vergleich zu anderen Säugetieren (~40 %) weist die Wasserfledermaus deutlich weniger repetitiven Elemente auf, hat unter allen bislang sequenzierten Fledermäusen jedoch den höchsten Anteil (vgl. Anhang Tab. 7.10). Der Anteil der repetitiven DNA wird, wie auch bei anderen Fledermäusen, stark von LINEs (Long interspersed nuclear elements) dominiert (16,4 %). Die zweit größte Klasse bilden die LTR-Retrotransposons (long terminal repeats, LTR) (5,5 %). DNA-Transposons und SINEs (short interspersed nuclear elements) umfassen jeweils 3,9 % bzw. 3,3 % der genomischen Sequenzen.

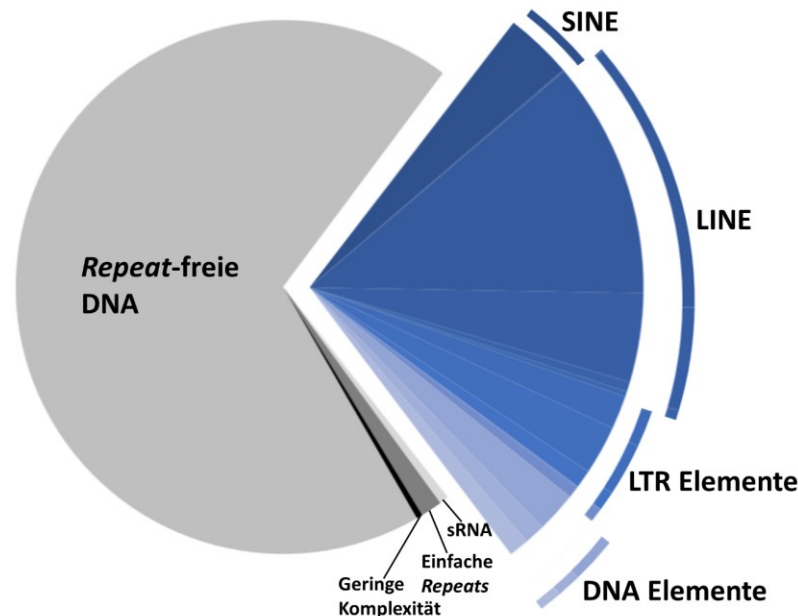


Abb. 4.17: Verteilung repetitiver DNA im Genom von *M. daubentonii*.

Dargestellt ist der Anteil der DNA an unterschiedlichen Klassen repetitiver Sequenzen, die durch eine RepeatMasker-Analyse ermittelt wurden. Abkürzungen: SINE, short interspersed nuclear elements; sRNA, small RNA repeats; LINE, long interspersed nuclear element; LTR, long terminal repeats.

4.4 Genomanalyse von *Polychromophilus murinus*

4.4.1 Selektive Anreicherung und Genomresequenzierung

Mithilfe der genomischen Sequenzen der nicht infizierten Fledermaus konnten mit den Reads der infizierten Fledermaus 1.189 Contigs (~1,3 Mbp) hergestellt werden, die aus dem Genom von *P. murinus* stammten (vgl. 3.14.2). Die so sortierten Contigs aus der infizierten Fledermaus wurden herangezogen, um Sequenzmotive zu identifizieren, die in dem Genom des Parasiten stärker vertreten sind als in der Wirtsfledermaus (vgl. 3.11.4). 17 Primer wurden auf diese Weise ausgewählt, in acht unterschiedliche Sets zusammengefasst und, wie in Abschnitt 3.7.5.2 beschrieben, gegenübergestellt (vgl. Tab. 2.4, Tab. 7.12). Basierend auf den Ergebnissen der Quantifizierung (nicht dargestellt), wurden die Proben Mg43 und WF43 zur Vorbereitung der Mate-Pair-Sequenzierung in unabhängigen Reaktionen zu gleichen Teilen mit den Primersets Set6 und Set7 amplifiziert (vgl. Abschnitt 3.6). Die Generierung der 3 kbp und 8 kbp Sequenzbibliotheken und die anschließende Sequenzierung der SWGA-DNA erfolgte durch StarSeq (Mainz). Insgesamt wurden etwa 9,5 Gb Rohdaten bestehend aus 30 Millionen (8 kbp) und 32 Millionen (3 kbp) Reads generiert (Tab. 4.8).

Im Vergleich zu dem Paired-End-Datensatz konnten durch die Mate-Pair-Sequenzierung deutlich mehr Reads parasitären Ursprungs generiert werden. Deutlich wird dies durch ein stringentes Mapping der Reads an die genomischen Sequenzen des Parasiten (vgl. 3.12.4). Während 23.857 Paired-End Reads an diese Contigs positioniert werden konnten, waren dies beim Mate-Pair-Datensatz 313.240 Reads. Abb. 4.19 visualisiert die durchschnittliche Abdeckung der Contigs

parasitären Ursprungs dargestellt als Konkater. Sowohl die WGA- als auch die SWGA-Probe erzeugten eine relativ ungleichmäßige Abdeckung mit einigen, allerdings unterschiedlichen, Bereichen, die stark überrepräsentiert waren.

Tab. 4.8: Zusammenfassende Statistik der Mate-pair-Sequenzierung.

Parameter	3 kbp	8 kbp
Rohdaten/bereinigte Daten*		
Mittlere Readlänge	151	151
Anzahl Reads	32.790.722 (32,574,889)	30.429.240 (30.322.216)
Anzahl der Nukleotide	4.951.399.022	4.594.815.240
Orientierung		
Paired-End	6.324.984 (28,8 %)	5.767.905 (28,3 %)
Mate-Pair	6.242.635 (28,4 %)	5.755.461 (28,2 %)
unbekannt	3.619.253 (16,5 %)	3.533.191 (17,3 %)
Single	5.796.652 (26,4 %)	5.326.748 (26,1 %)

*Rohdaten nach Qualitätstrimming, Entfernung von Adaptern sowie Vektorsequenzen.

Die Paired-End-Sequenzierung der WGA-Probe erzeugte eine im Durchschnitt 4,9fache Abdeckung der Referenzsequenz. 25 Bereiche in der Größe von 579-5.521 bp (56.970 bp) zeigen dagegen eine mindestens 10fache Abdeckung. Dabei handelt es sich vor allem um Regionen, die für Gene und Genfragmente von konservierten, lysinreichen Proteinen mit unbekannter Funktion kodieren (vgl. 3.16). Teilweise war die Abdeckung auch besonders hoch in Regionen, in denen Gene identifiziert werden konnten, die in den Genomen von *Plasmodium*-Spezies in mindestens zwei Genomkopien vorliegen (z. B. cation-transporting P-ATPase, Ubiquitin fusion degradation protein 1).

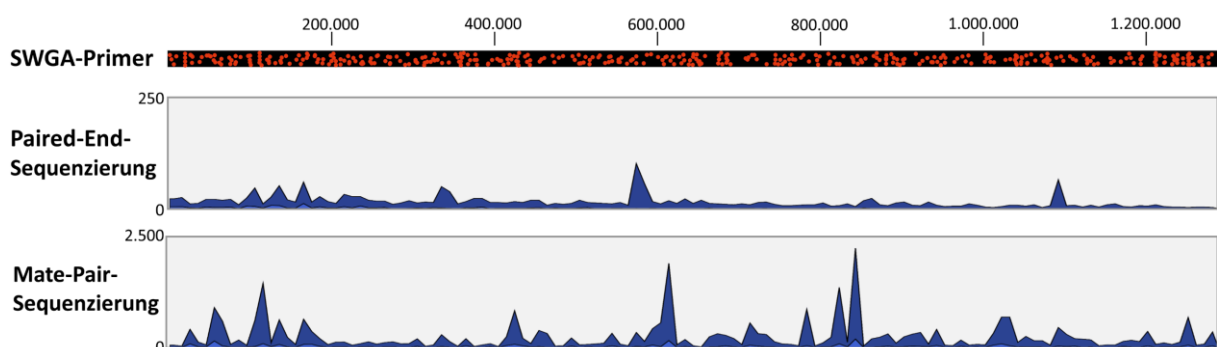


Abb. 4.19: Abdeckung der Sequenzen von *P. murinus* im Überblick.

Dargestellt sind die aus Probe WF43 identifizierten Sequenzen von *Polychromophilus murinus* (1,3 Mbp) als Konkater. Das oberste Felder zeigt, dargestellt durch rote Punkte, Positionen, an denen die SWGA-Primer binden. Darunter befinden sich die Abdeckung der Sequenz in 1 kbp-Fenstern nach Paired-End-Sequenzierung eines WGA-Produkts bzw. nach Mate-Pair-Sequenzierung eines SWGA-Produkts.

Auch nach der Mate-Pair-Sequenzierung konnten überrepräsentierte Regionen hauptsächlich als Gene oder Genfragmente von konservierten Proteinen mit unbekannter Funktion identifiziert werden. Allerdings wies dieser Datensatz mit 64 überrepräsentierten Regionen (529-3.380 bp; insgesamt 72.689 bp) eine deutlich ungleichmäßigere Verteilung der Reads auf. Eine Korrelation mit den Primer-Bindungsstellen der verwendeten SWGA-Primer konnte nicht nachgewiesen werden. Zudem konnten 23 Regionen identifiziert werden, die durch keine Reads repräsentiert

wurden. Diese Bereiche erstreckten sich über insgesamt 25.508 bp (444-4401 bp; 1,9 % der gesamten Sequenz). Die Mehrheit dieser Regionen enthielt einen nicht unerheblichen Anteil an repetitiver AT-reicher DNA. Insgesamt konnte durch die Mate-Pair-Sequenzierung im Durchschnitt eine 29,9fache Abdeckung der Referenzsequenz erreicht werden. Im Vergleich zur einfachen WGA, konnte die durchschnittliche Abdeckung der Sequenzen des Genoms von *P. murinus* durch die SWGA also um ein 6faches erhöht werden.

Dies konnte auch durch ein erstes Assembly bestätigt werden. Im Vergleich zum Rohdaten-Assembly nach der Paired-End-Sequenzierung (42,2 %; vgl. 4.2.2), weist das Rohdaten-Assembly mit den Reads der Mate-Pair-Sequenzierung mit nur 15 % einen deutlich geringeren GC-Gehalt auf (Tab. 4.9, Abb. 4.20). Mit 15,6 Mbp ist das Mate-Pair-Assembly auch deutlich kleiner als das vorige. Zudem deutet die geringe Gesamtzahl der erstellten Contigs (4.241) auf eine weniger starke Fragmentierung.

Tab. 4.9: Statistiken des Mate-Pair-Assemblys.

Parameter	Assembly
Anzahl der Contigs	4.241
Anzahl der Nukleotide	15.671.834
Minimale Contiglänge	956
Maximale Contiglänge	25.508
N50	6.101
Mittlerer GC-Gehalt	15,0

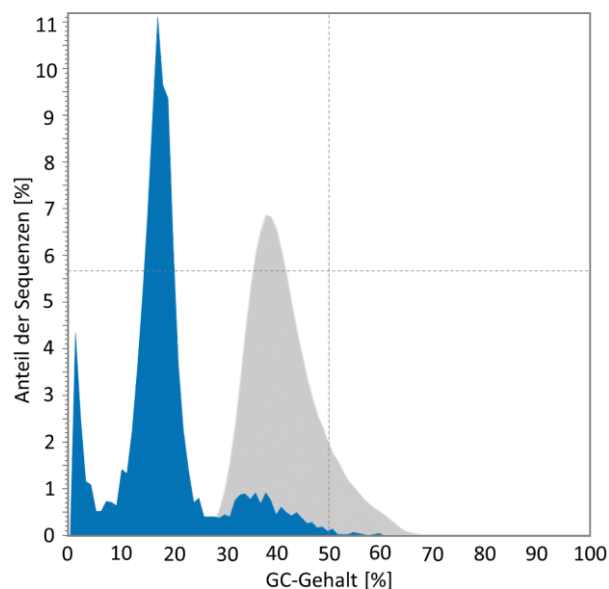


Abb. 4.20: Verteilung des GC-Gehalts im Mate-Pair-Assembly. Dargestellt ist der relative Anteil aller Contigs in Abhängigkeit zum GC-Gehalt. Als Vergleich ist in grau die Verteilung der Paired-End-Sequenzierung dargestellt.

Wie auch die Verteilung des GC-Gehalts zu erwarten (Abb. 4.20), resultieren die Mehrheit der Contigs aus dem Genom des Parasiten *P. murinus* (Abb. 4.21). In einer BLAST-Analyse (vgl. 3.11.2) wurden 2.681 Contigs (63,5 %) identifiziert, die signifikante Signifikanzidentitäten mit Sequenzen aus dem Phylum der Apikomplexa aufweisen. Contigs, die ihren Ursprung wahrscheinlich im Genom der Fledermaus hatten, sind dagegen mit 26,2 % (1.107 Contigs) vertreten. Contigs, die nicht zugeordnet werden konnten (3,2 %), oder die durch andere Kontaminationen entstanden sind und signifikante BLAST-Treffer zu Sequenzen von Arten aus den Taxa Bakterien, Pilze, Viren, Helminthen oder Arthropoden haben (7,1 %) waren dagegen die Minderheit. Im Vergleich zum Rohdaten-Assembly der Paired-End-Sequenzierung, bei dem auch die längsten Contigs ihren Ursprung in bakteriellen Genomen hatten (vgl. 4.2.1), haben die fünf längsten Sequenzen (>20 kbp) hier homologe Sequenzen innerhalb der Gruppe der Apikomplexa. Insgesamt konnte

auf diese Weise mit 10,1 Mbp ein deutlich größerer Anteil des Genoms von *P. murinus* identifiziert werden. Ausgehend von der durchschnittlichen Genomgröße von ~23 Mbp, entspricht dies etwa 43 % des Genoms.

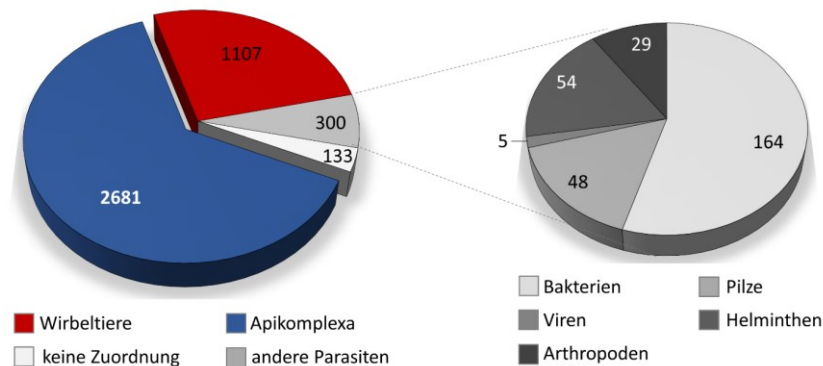


Abb. 4.21: Taxonspezifische Einordnung des Mate-Pair-Rohdatenassemblys.

Die Klassifizierung der Sequenzen erfolgte mit der nt NCBI-Protein-Datenbank durch BLAST. Das linke Diagramm zeigt die Einteilung der Contigs in die Taxa Wirbeltiere (rot) Apikomplexa (blau), und andere Parasiten. Der Anteil ohne Zuordnung (weiß) bezieht sich auf Gene ohne Treffer oder mit signifikanten Treffern in unterschiedlichen Taxa. Rechts sind alle identifizierten Sequenzen parasitären Ursprungs weiter aufgeschlüsselt.

4.4.2 Das mitochondriale Genom von *P. murinus* und darauf basierende phylogenetische Analysen

Das als mitochondrial identifizierte Contig im Mate-Pair-Assembly erstreckte sich über 5.995 bp. Von der Größe ist das mitochondriale Genom damit vergleichbar mit anderen sequenzierten mtDNAs aus dem Phylum der Haemosporida. Im Durchschnitt wurde jede Base durch 2.395 Reads abgedeckt. Auf der mtDNA konnten drei proteinkodierende Gene (*cox1*, *cox2* und *cytb*) sowie zwei hochgradig fragmentierte rRNA-Gene annotiert werden (Abb. 4.22). Gene für tRNA fehlten dagegen. *Cox1* und *cytb* sind, durch lediglich 24 bp getrennt, hintereinander angeordnet und werden von einem Strang der mtDNA kodiert, während *cox3* auf dem anderen Strang lokalisiert ist. Als Stopkodon konnte bei allen drei Strukturgenen das Kodon TAA identifiziert werden (Übersicht in Tab. 7.11). *Cytb* weist das typische Startkodon ATG auf, während bei den anderen proteinkodierenden Genen kein konventionelles Startkodon gefunden werden konnte. Die Startkodon-Zuordnungen erfolgten bei diesen Fällen durch den Vergleich mit öffentlich verfügbaren vollständigen mt-Genomen verschiedener Haemosporida-Arten. Insgesamt weist die mtDNA mit 31% einen recht niedrigen GC-Gehalt auf. Die Verschiebung der Basenkomposition in Richtung A und T ist auch auf die Kodonverwendung in den Proteinkodierenden Bereichen zurückzuführen. Bei den allen synonymen Kodons werden A und T an den entsprechenden Wobble-Positionen bevorzugt gegenüber C und G verwendet (Übersicht im Anhang: Tab. 7.11). Dadurch ist der GC-Gehalt der dritten Kodon-Position mit 8,8 % auch deutlich kleiner als an Position 1 oder 2 (31,9 %, bzw. 36,4 %). Interessanterweise ist der GC-Gehalt der Proteinkodierenden Gene tatsächlich sogar viel kleiner (25,7 %) als in den nichtkodierenden Regionen, bzw. jenen, die für rRNA-Fragmente kodieren (37,5 %).

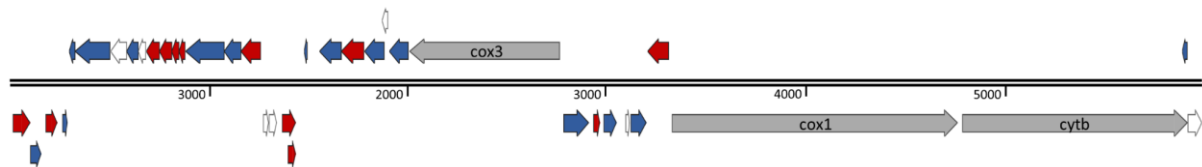


Abb. 4.22: Organisation der mtDNA von *P. murinus*.

Dargestellt ist das repetitive 6 kbp Elements der mitochondrialen DNA. Start- und Endpunkt orientiert sich dabei an der homologen Sequenz von *P. falciparum* (Genbank Accessionnr. M76611). Die Orientierung der Pfeile sowie die Darstellung ober- bzw. unterhalb der Skalierung repräsentieren die Richtung der Transkription. Grau: Proteinkodierende Gene; rot: Fragmentierte Gene der 16S-rRNA, blau: Fragmentierte Gene der 23S-rRNA, weiß: andere nicht zugeordnete rRNA-Gene.

Die Regionen, die für die 16S- und 23S-rRNA, kodieren sind über das gesamte mitochondriale Genom verteilt. Insgesamt konnten 12 16S- und 15 23S-rRNA-kodierende Fragmente annotiert werden. Die Gene lagen dabei unregelmäßig fragmentiert vor. Das bedeutet, es liegen einerseits unterschiedlich große Fragmente vor (15-195 bp) und andererseits sind diese scheinbar zufällig auf dem Genom verteilt, und zwar auf beiden DNA-Strängen. Die Mehrheit dieser Fragmente ist zwischen den Genen für *cytb* und *cox3* lokalisiert. Drei 16S- und zwei 12S-rRNA-kodierende Fragmente befinden sich dagegen zwischen den Genen für *cox3* und *cox1*. Sowohl die Position als auch die Sequenz dieser Fragmente ist, wenn man von den mtDNAs zweier *Nycteria*-Arten absieht (Karadjian et al., 2016), hoch konserviert.

Angeichts der Tatsache, dass sowohl die Gene als auch die nichtkodierenden Regionen stark konserviert sind, wurden phylogenetische Analysen einerseits basierend auf der vollständigen Sequenz der mtDNA und andererseits mit den Aminosäuresequenzen der drei Strukturgene durchgeführt (vgl. Anhang: Abb. 7.5-7.9, Tab. 7.13). In allen Topologien zeigte sich die Gattung *Plasmodium* paraphyletisch in Hinblick auf andere fledermausinfizierende Gattungen (Abb. 4.23). Dies wurde bei allen Analysen, mit Ausnahme der MP-Analyse basierend auf den Aminosäuresequenzen, maximal unterstützt. Dabei bildet die Gattung *Hepatocystis* ein Monophylum mit der Untergruppe *Plasmodium* (*Plasmodium*). Die phylogenetische Position der Gattung *Nycteria* konnte dagegen nicht mit Sicherheit bestimmt werden. Sowohl eine engere Verwandtschaft mit allen primateninfizierenden Plasmodienarten und *Hepatocystis* als auch eine monophyletische Gruppierung mit vogel- und reptilieninfizierenden Parasiten konnten gezeigt werden. *P. murinus* steht dagegen, gemeinsam mit hufentierinfizierenden Plasmodienarten, an der Basis des Clusters, das alle anderen *Plasmodium*-, *Hepatocystis*- und *Nycteria*-Spezies enthält. Diese Gruppierung wurde in allen Topologien, mit Ausnahme der MP-Analyse basierend auf den Aminosäuresequenzen, gut unterstützt und stimmt mit anderen phylogenetischen Studien, die teilweise nur auf einzelnen Genfragmenten basieren, überein (vgl. Boundenga et al., 2016; Templeton et al., 2016; Galen et al., 2018).

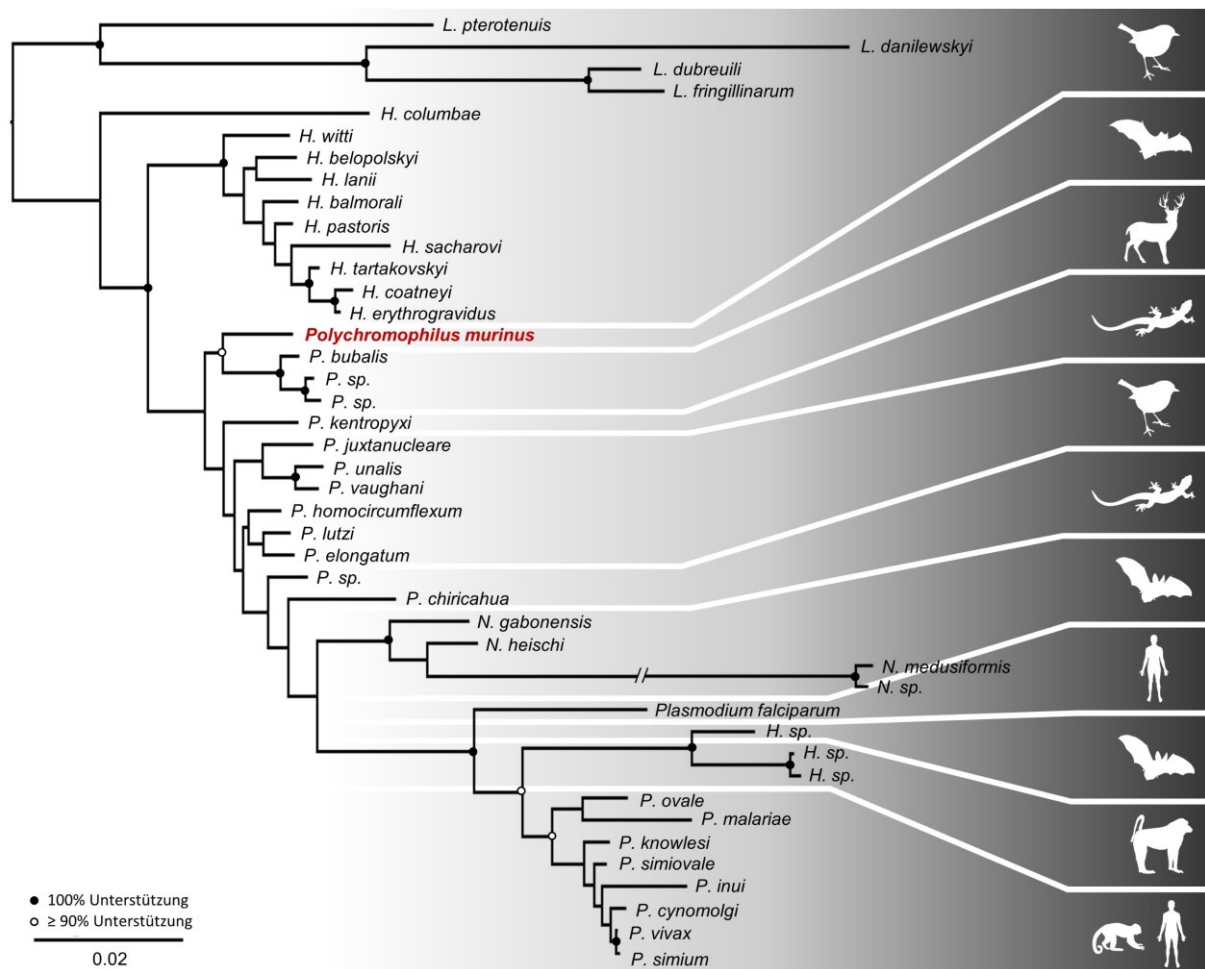


Abb. 4.23: Phylogenetische Stellung von *P. murinus* basierend auf Daten der mtDNA.

Die dargestellte Topologie basiert auf der MrBayes-Analyse der Aminosäuresequenzen der drei Strukturgene auf der mtDNA. Alle Verzweigungslängen sind proportional zur durchschnittlichen Anzahl der Substitutionen. Die Länge des Balkens entspricht 0,02 Substitutionen pro Position. Zur besseren Übersicht wurde die Astlänge des längsten Zweigs halb so lang dargestellt (gekennzeichnet durch //). Unterstützungswerte entstammen auf sechs separaten phylogenetischen Analysen. ML-, MP- und Mr-Bayes-Analysen der Aminosäuresequenzen sowie die entsprechenden Analysen der gesamten mtDNA-Sequenz. Verzweigungen mit Unterstützungswerten größer als 90 % (bzw. 0,90) sind durch einen weißen Punkt, Unterstützungen von 100 % (bzw. 1,0) durch einen schwarzen Punkt markiert. Rechts dargestellt Wirtsspektrum der untersuchten Spezies. Rot hervorgehoben: in dieser Arbeit sequenzierte mtDNA.

4.4.3 Konvergente Reduktion des Apikoplasten innerhalb der Haemosporida

Im Gegensatz zum mitochondrialen Genom lag das Genom der apikomplexen Plastide (Apikoplastmid, plDNA) auf zwei Contigs fragmentiert vor. Dabei waren proteinkodierende Bereiche (24.227 bp) vom tRNA/rRNA-Gen-Cluster (5.042 bp) getrennt. Die Coverage des tRNA/rRNA-Gen-Clusters war im Vergleich zum größeren Bereich (178x) deutlich höher (733x). Daher lag die Vermutung nahe, dass die beiden Contigs Teil einer für Apikoplastmide typischen zweiteiligen, ringförmigen Architektur sind, die sich durch eine größere proteinkodierenden Region und einem invertiertem Repeat (IR) charakterisieren lässt. Dieser IR ist wiederum aus zwei tRNA/rRNA-Gen-Cluster mit umgekehrter Orientierung zusammengesetzt (Wilson et al., 1996).

Bioinformatische Bestreben die Contigs in einer solchen Anordnung durch iteratives Mapping zu vervollständigen (vgl. 3.12.4, Abb. 4.24a, obere gestrichelte Linien), waren nur zwischen dem proteinkodierenden Bereich und den tRNA/rRNA-Gen-Clustern erfolgreich. Die Sequenz, welche die beiden tRNA/rRNA-Gen-Cluster des IR verbindet, konnte allerdings nicht aufgelöst werden. Auch anfängliche Versuche, den fehlenden Bereich über spezifische Primer zu amplifizieren schlugen fehl (vgl. 3.7.2, Tab 2.1).

Die Primer wurden deshalb im weiteren Verlauf so gewählt, dass die PCR-Produkte, ausgehend vom 16S-rRNA-Gen eines tRNA/rRNA-Gen-Clusters, das zweite tRNA/rRNA-Gen-Cluster überspannten (Abb. 4.24 a, untere gestrichelte Linien). Da beide amplifizierten Fragment ein gesamtes tRNA/rRNA-Gen-Cluster enthalten sollten, wurden Fragmentgrößen von über 7 kbp erwartet. Tatsächlich waren die Amplifikate mit 2 kbp und 1,8 kbp um etwa 5 kbp kürzer als erwartet. Mit den Sequenzen dieser Fragmente konnte die pIDNA von *P. murinus* schließlich zusammengesetzt werden. Ausgehend von dem 16S-rRNA-Gen wurde dabei entweder in das erste (Ribosomales Protein S4, *rps4*) oder in das letzte (Schwefel Assimilierungs Protein B, *sufb*) proteinkodierende Gen hinein sequenziert (Abb. 4.24 c). Wie bereits durch die Amplifikatgrößen unterstützt, fehlte das zweite tRNA/rRNA-Gen-Cluster in beiden Fragmenten. Da sowohl das 16S-rRNA-Gen als auch das 23S-rRNA-Gen damit in direkter Nachbarschaft zu den Genen *rps4* und *sufb* lokalisiert ist, müssen zwei Isoformen der pIDNA vorliegen, in denen das tRNA/rRNA-Gen-Cluster jeweils einmal in verschiedenen Orientierungen vorkommt. Abgesehen davon sind die beiden Isoformen in ihrer Sequenz identisch.

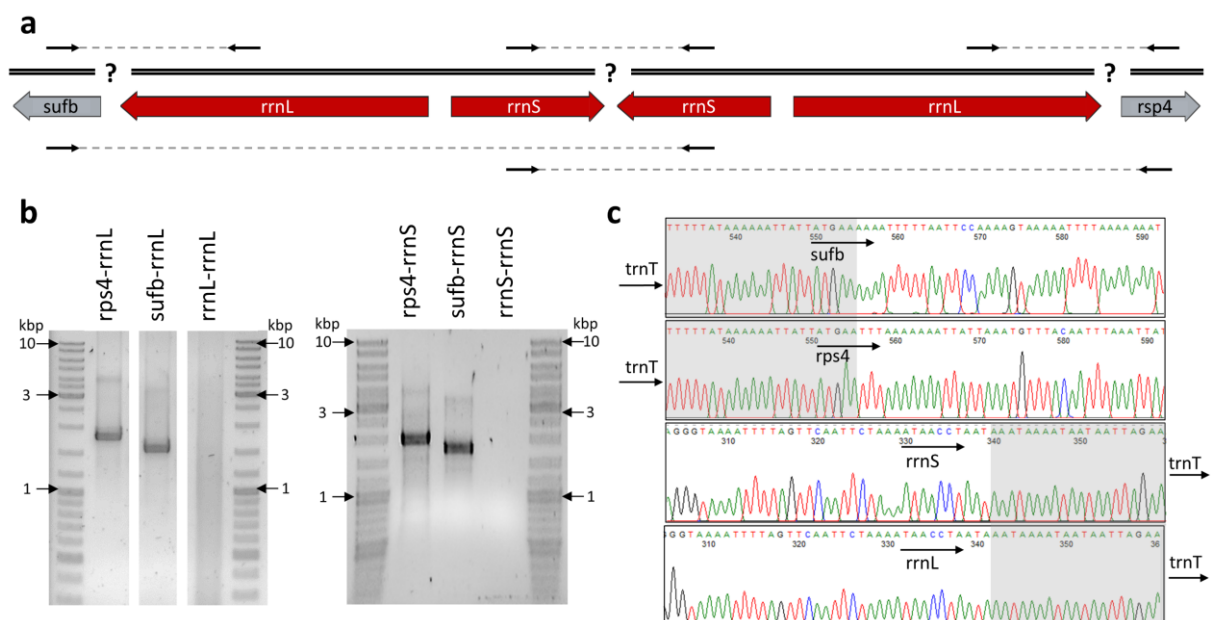


Abb. 4.24: rRNA/tRNA-Gencluster des Apikoplast-Genoms von *P. murinus*.

(a) Schematische Darstellung der Sequenzierstrategie zur Entschlüsselung der vorhandenen Lücken (?) im Genom. In der PCR- verwendete Primer sind durch Pfeile, erwartete Amplifikate durch gestrichelte Linien gekennzeichnet. Schema ist nicht maßstabsgetreu. (b) Nachweis der amplifizierten Fragmente im 0,7%igen Agarosegel. (c) Vergleich der Sequenzier-Chromatogramme der Amplifizierten Fragmente. Das tRNA/rRNA-Gen-Cluster ist von tRNA-Genen flankiert. Vor dem tRNA-Gen ist entweder das *rrnL* oder das *rrnS* kodiert, danach *sufb* oder *rps4*. Gleiche Sequenzen sind farblich hervorgehoben. Pfeile repräsentieren Transkriptionsrichtung. Abkürzungen: rps, ribosomal protein S; rrn rRNA-Gen; suf, sulphur assimilation protein.

Insgesamt erstreckt sich die plDNA über 29.336 bp und weist mit 13,3 % einen niedrigen GC-Gehalt auf. Auf dem zirkulären Genom konnten 31 Proteinkodierende, 26 tRNA- und 2 rRNA-Gene (12S und 16S) annotiert werden (Abb. 4.25, Übersicht im Anhang: Tab. 7.14). Das bereits erwähnte tRNA/rRNA Gen-Cluster besteht dabei aus zwei rRNA- sowie aus acht tRNA-Genen. Zusätzlich wird dieser Bereich auf beiden Seiten durch jeweils ein sequenzidentisches tRNA-Gen für die Aminosäure Threonin flankiert. Alle Proteinkodierenden Gene, sowie 16 weitere tRNA-Gene, sind dabei außerhalb dieses Clusters kodiert. Sowohl das Genrepertoire als auch das Arrangement dieser Region ist dabei identisch zu bereits publizierten Apikoplast-Genomen von Haemosporiden und kann grob in zwei größere Bereiche eingeteilt werden. Diese erstrecken sich über etwa 14 kbp bzw. 10 kbp und unterscheiden sich vor allem durch die Transkriptionsrichtung. Die Gene im größeren Abschnitt sind, mit Ausnahme des tRNA^{Phe}-Gens, auf dem in Abb. 4.25 dargestellten forward Strang kodiert, während alle Gene auf dem kleineren Abschnitt auf dem reversen Strang kodiert werden.

Die meisten Gene auf der plDNA kodieren für Komponenten, die bei der Transkription oder Translation im Apikoplasten beteiligt sind. Dazu zählen neben den tRNA-Genen, diverse ribosomale Proteine der großen (7 *rpl* Gene) und kleinen (10 *rps* Gene) Untereinheit, der Elongationsfaktor Tu (*tufA*), sowie die Untereinheiten der RNA-Polymerase (*rpo*). Zusätzlich kodiert das Genom die caseinolytische Protease C (*clpc*), die wahrscheinlich als Chaperon fungiert, das *sufb*, das bei der Fe/S-Proteinbiogenese beteiligt ist, sowie sieben weitere proteinkodierende Bereiche (ORF), die sich jeweils über 150-400 bp erstrecken und noch nicht näher charakterisiert werden konnten. Alle annotierten proteinkodierenden Gene beginnen dabei mit dem typischen Startkodon ATG. Als Stopkodon konnte bei jeweils 10 Genen das Kodon TAG und bei 21 Genen das Kodon TAA identifiziert werden (Übersicht in Tab. 7.15). Wie auch bei *P. falciparum* konnte bei der plDNA von *P. murinus* eine Leserasterverschiebung im *rpoC2*-Gen festgestellt werden, die zu dem Stopkodon TAG führt (Wilson et al., 1996). Während bei *P. falciparum* nach diesem Stopkodon jedoch ein Startkodon für den zweiten Bereich des *rpoC2*-Gens folgt, sind diese hier durch 31 bp getrennt. Ob diese Gene, in dieser Arbeit als *rpoC2a* und *rpoC2b* annotiert, tatsächlich für zwei unterschiedliche Genprodukte kodieren, muss noch untersucht werden.

Tatsächlich liegt auch die DNA des Apikoplasten in sehr kompakter Form vor. Es konnte jedoch ein 141 bp langes Gruppe-I-Intron innerhalb des tRNA^{Leu}-Gens identifiziert werden. Andere Introns konnten nicht identifiziert werden. Allerdings weist die plDNA mit lediglich 3,31 % (972 bp) einen recht geringen Anteil an nichtkodierender DNA auf. Eine Kontrollregion wie bei der mtDNA der Wirtsfledermaus konnte jedoch nicht identifiziert werden. Vielmehr finden sich nichtkodierende Bereiche, zumeist eher Abschnitte von etwa 5 -30 bp, zwischen den tRNA-Genen sowie in Bereichen, die tRNA-Genen von proteinkodierenden Genen trennen. Die meisten proteinkodierenden Gene sind dagegen etwas kompakter lokalisiert und lediglich von einzelnen Nukleotiden getrennt. Stellenweise überlappen sich die Leserahmen benachbarter Gene um 4 bp (*rps17* und *rpl14*, *rpl6* und *rps5*, *rpl36* und *rps11* sowie ORF129 und *clpc*). Größere Überlappungen treten lediglich in den beiden Bereichen auf, in denen die Orientierung der Transkription invertiert vorliegt (ORF78 und tRNA^{Phe}: 24 bp, ORF105 und *rps2*: 13 bp).

Das zirkuläre Molekül besteht aus 31 Genen für Proteine (graue Pfeile), zwei rRNAs (rote Pfeile) und 26 tRNAs (blaue Pfeile, Großbuchstabe steht jeweils für die entsprechende Aminosäure). Die Orientierung der Pfeile repräsentiert die Transkriptionsrichtung. Abkürzungen: clpc, caseinolytic protease C; ORF, open reading frame; rpo, RNA-Polymerase; rpl/s, large/small ribosomal protein; rrn, rRNA; sufB, sulfur mobilizing protein B, trn, tRNA.

97

beiden Fällen wird das tRNA/rRNA Gen-Cluster beidseitig von einer über 200 bp langen, invertierten Sequenz flankiert, die das rRNA^{Thr} Gen und einen kurzen Abschnitt des LSU-rRNA Gens einschließt (Abb. 7.12). Trotz der strukturellen Übereinstimmung legt der phylogenetische Kontext der beiden Arten eine konvergente Entstehung dieser Deletion nahe.

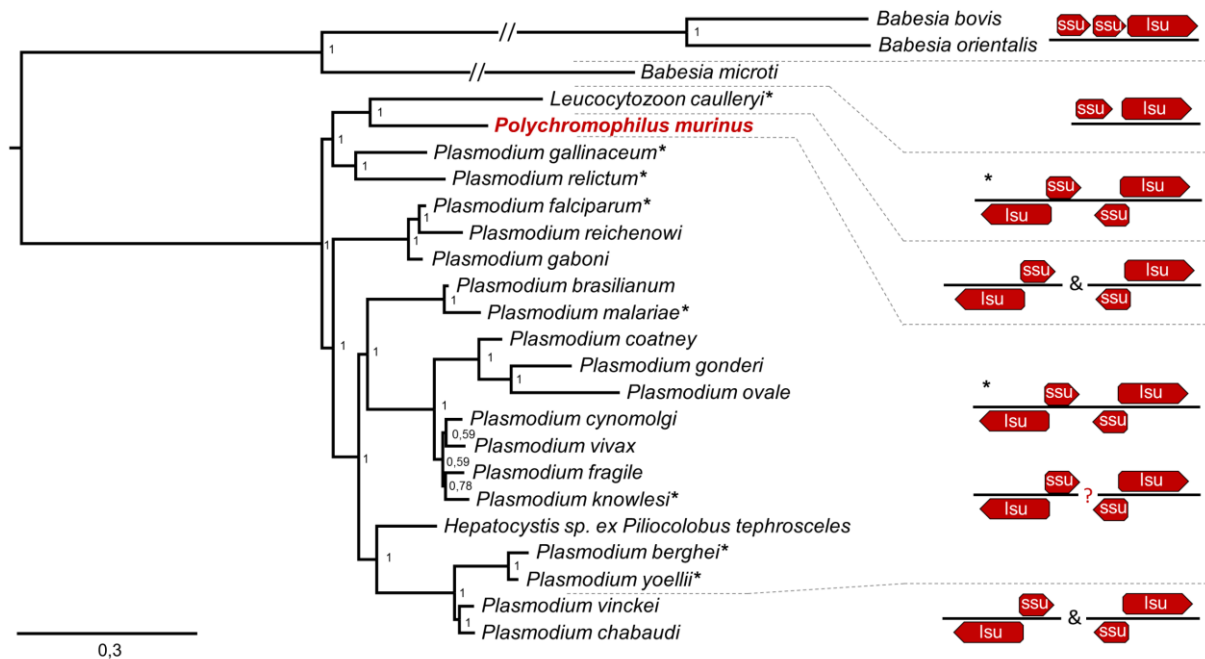


Abb. 4.26: Rearrangement des Apikoplasמידs im phylogenetischen Kontext.

Die dargestellte Topologie basiert auf der MrBayes-Analyse von Aminosäuresequenzen proteinkodierender Gene der apikomplexen Plastiden. Die Verzweigungslängen entsprechen der durchschnittlichen Anzahl von Substitutionen. Die Länge des Balkens entspricht 0,3 Substitutionen pro Position. Zur besseren Übersicht wurden die Astlängen der Piroplasmiden halb so lang dargestellt (/). Der Aufbau der einzelnen tRNA/rRNA Gen-Cluster wurde schematisch anhand der Orientierung der rRNA-Gene (rot) dargestellt. Die Orientierung der Pfeile repräsentiert die Transkriptionsrichtung. Sequenzen, die bereits vollständig sequenziert wurden und ein invers repetitives tRNA/rRNA Gen-Cluster aufweisen, sind mit einem * markiert. Bei anderen Apikoplasמידen wurde ein identischer Aufbau postuliert (?). Abkürzungen: lsu, large subunit ribosomal ribonucleic acid; ssu, small subunit ribosomal ribonucleic acid.

4.4.4 Das Kern-Genom von *P. murinus*

Zur Rekonstruktion des Kern-Genoms des Parasiten wurden sowohl die Reads der Mate-Pair-Sequenzierung der mit *P. murinus* angereicherten DNA als auch die sortierten Reads der Paired-End-Sequenzierung der infizierten Fledermaus herangezogen. In einem mehrschrittigen Prozess, der ein hohes Maß an manueller Intervention erforderte (vgl. 3.15.1), wurden 21 Sequenzen zusammengesetzt mit einer durchschnittlichen Abdeckung von 50,8 (Tab. 4.10; Abb. 4.27). Das größte Scaffold wies beispielsweise eine Größe von 1,74 Mbp auf, enthielt jedoch 629 Lücken bestehend aus 140.415 Basenpaaren.

Zusammengenommen enthielten die Sequenzen 11.354.494 Basen. 7.201 kurze Contigs mit einer Gesamtlänge von 3,32 Mbp wurden nicht in die pseudochromosom-artigen Sequenzen eingebaut. 5.516 Contigs (1,62 Mbp) waren mit 200-500 Basenpaare besonders kurz. Da sich diese Contigs während der Rekonstruktion der pseudochromosom-artigen Sequenzen als besonders wertvoll erwiesen, um weitere Reads parasitären Ursprungs zu identifizieren, wurden sie trotz ihrer geringen

Größe nicht aus dem Datensatz entfernt. Der resultierende Entwurf des Genoms von *P. murinus* hatte insgesamt eine Größe von 14,7 Mbp und wurde im Mittel 184,5fach durch die sequenzierten Reads abgedeckt (Tab. 4.10; Abb. 4.26).

Tab. 4.10: Statistiken zur genomischen Sequenz von *P. murinus* im Vergleich.

Parameter	Finales Assembly*	Assembly (>50 kbp)	<i>P. brasilianum</i>	<i>P. falciparum</i>	<i>P. gallinaceum</i>	<i>H. tartakovskyi</i>
Anzahl der Sequenzen	7.224	21	953	14	154	2.983
Anzahl der Nukleotide	14.711.870	11.354.194	30.089.869	23.332.839	25.034.007	23.209.007
Minimale Sequenzlänge	200	52.689	5.004	5.967	2042	991
Maximale Sequenzlänge	1.741.666	1.741.666	344.535	3.291.936	2.985.323	108.686
N50	679.571	868.979	52.286	1.687.656	1.343.538	17.240
Mittlerer GC-Gehalt	18,6 %	18,6 %	24,9 %	19,3 %	17,8 %	25,4 %
Anzahl der Lücken (Gaps)	4.338	3.335	53.547	0	1.840	1.484
Anzahl der Gap-Nukleotide	969.695	968.395	988	0	1.156.820	457.844

P. brasilianum (GCA_00188515.2), *P. falciparum* [3D7; PlasmoDB Release 32], *P. gallinaceum* [8A; PlasmoDB Release 32], *H. tartakovskyi* (Bensch et al., 2016), *enthält die extrachromosomalen Sequenzen des Mitochondriums und der apikomplexen Plastide.

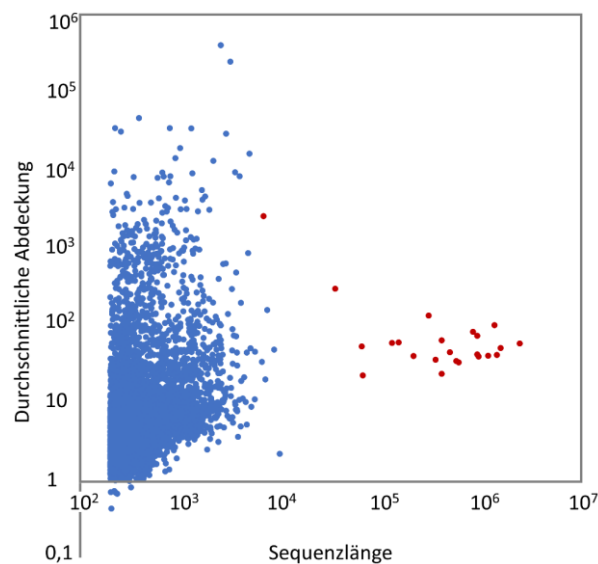


Abb. 4.27: Durchschnittliche Abdeckung der genomischen Sequenzen von *P. murinus*.

Rote Punkte: Sequenzen >50 kbp sowie die extrachromosomalen Sequenzen des Parasiten.

Bei der Rekonstruktion der pseudochromosom-artigen Sequenzen wurden während der referenzbasierten Orientierung artifiziell Lücken von 100 bp zwischen benachbarten Contigs eingeführt, die keine Aussage über die tatsächliche Größe der Lücken ermöglichen (vgl. 3.15.1). Ein Vergleich der konservierten Syntanie-Bereiche zwischen den Sequenzen von *P. murinus* und den entsprechenden Sequenzen aus dem Referenzgenom von *P. falciparum* impliziert eine Differenz in Größenordnung von etwa 7,5 Mbp (Abb. 7.13). Insgesamt ergibt sich dadurch eine Genomgröße von ungefähr 22,2 Mbp, die mit der Genomgröße verwandter Parasiten aus dem Taxon der Haemosporida vergleichbar ist. Die hier dargestellten Sequenzen repräsentieren etwa 66 % der gesamten Sequenz. Der hohe Anteil nicht vorhandener Sequenzen spiegelt sich bei der Beurteilung der Vollständigkeit universell orthologer Gene im Genom wider (vgl. 3.17). Gemessen am Datensatz für Protisten fehlen 63,3 % der konservierten Gene (Abb. 4.28). Die hohe Anzahl fehlender Gene ist mit denen aus dem Genom von *Haemoproteus tartakovskyi* vergleichbar (61 %).

Ein großer Teil der Gene, die in *P. murinus* nicht vorhanden waren, waren auch in mindestens einem der anderen Genome nicht anwesend (48,4 %).

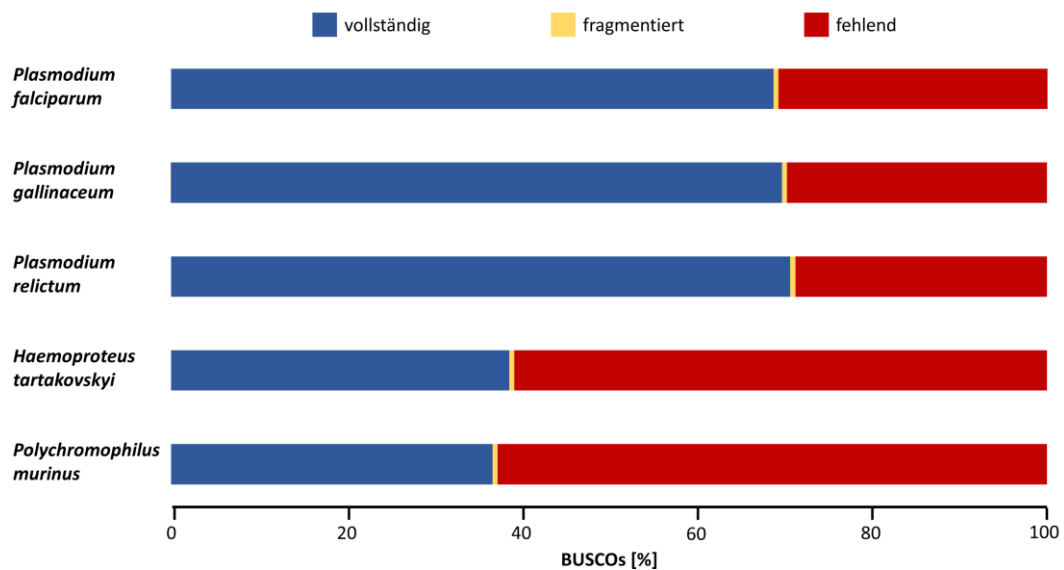


Abb. 4.28: Bewertung der Vollständigkeit des Genoms von *P. murinus* im Vergleich.

Dargestellt ist der Anteil vorhandener Gene gemessen anhand des BUSCO-Referenz-Datensatzes für Protisten zur Bewertung der Qualität einzelner Genome aus dem Taxon Haemosporida (Waterhouse et al., 2018). Blau, Anteil vollständiger Gene; Gelb, Anteil fragmentierte Gene, rot, Anteil fehlender Gene.

Mithilfe der verwendeten Maker2-Pipeline konnten in den genomischen Sequenzen insgesamt 3.366 Gene vorhergesagt werden (vgl. 3.16.1). Im Vergleich zu der Anzahl annotierter Gene verwandter Parasiten, war eine solche Anzahl, gemessen an der Größe der hier erstellten Contigs, zu erwarten. Die Gendichte der genomischen Sequenzen von *P. murinus* entspricht mit einem Wert von 229 Genen/Mbp in etwa dem von *P. falciparum* (238 Gene/Mbp). Aber auch andere Charakteristika stimmen mit denen verwandter Haemosporida überein (Tab. 4.11). Die mittlere Größe der annotierten Gene beträgt 2,2 kbp und ist damit nur etwas kleiner als die der Gene von *H. tartakovskyi* (2,3 kbp). Auch der Anteil an Genen, die größer als 4 kbp sind, ist mit 10,8 % etwas kleiner als der von *P. falciparum* oder *H. tartakovskyi* (16,3 % bzw. 15,1 %), aber immer noch deutlich größer als z. B. der von *Saccharomyces cerevisiae* (3,6 %) (Gardner et al., 2002). Auch das größte Gen war mit 24.995 bp rund 5 kbp kleiner als die entsprechenden Gene von *P. falciparum* oder *H. tartakovskyi*. Introns konnten in 62,1 % aller Gene identifiziert werden. Dies entspricht einem etwas höheren Anteil als in *P. falciparum* (55,1 %) ist jedoch etwas kleiner in *H. tartakovskyi* (69,1 %).

Die durchschnittliche Introngröße von *P. murinus* (206 bp) ist etwas größer als bei *P. falciparum* (164 bp). Exons sind mit einer durchschnittlichen Größe von 613 bp dagegen deutlich kürzer. Die Menge an Introns pro Gen ist jedoch vergleichbar und liegt bei den drei Vertretern der Gattungen *Haemoproteus*, *Plasmodium* und *Polychromophilus* zwischen zwei und drei.

Tab. 4.11: Genom-Eigenschaften von *P. murinus* im Vergleich zu zwei anderen Haemosporiden.

Parameter	<i>H. tartakovskyi</i>	<i>P. falciparum</i>	<i>P. murinus</i>
Genom			
Größe	23,2	23,3	14,7
GC-Gehalt (%)	25,4	19,3	18,6
Gene*			
Anzahl	5.988	5.542	3.354
Mittlere Größe (bp)	2.333	2.538	2.182
Gene > 4kbp (%)	15,1	16,3	10,8
Größtes Gen	29.935	30.864	24.995
Gene mit Introns (%)	69,1	55,1	62,1
Exons			
Mittlere Anzahl je Gen	2,5	2,6	2,9
GC-Gehalt (%)	28,3	23,8	21,4
Mittlere Größe (bp)	661	861	613
Introns			
GC-Gehalt (%)	23,1	12,9	14,3
Mittlere Größe (bp)	287	164	206
Intergenetische Regionen			
GC-Gehalt (%)	22,6	14,4	14,2

P. falciparum [3D7; PlasmoDB Release 32], *H. tartakovskyi* (Bensch et al., 2016), *Gene >100 bp

4.4.4.1 GC-Gehalt und Aminosäurenkomposition

Insgesamt weisen die genomischen Sequenzen von *P. murinus* mit 18,6 % einen besonders niedrigen GC-Gehalt auf. Lediglich die Genome der beiden vogelinfizierenden *Plasmodium*-Arten *P. gallinaceum* und *P. relictum* weisen mit einem GC-Gehalt von 17,8 % bzw. 18,3 % ein noch stärkeres Basenungleichgewicht auf (Böhme et al., 2018). Wie auch bei *H. tartakovskyi* und *P. falciparum* lassen sich kodierende Bereiche allein aufgrund des GC-Gehalts von intergenetischen Bereichen differenzieren. Während intergenetische Bereiche einen mittleren GC-Gehalt von lediglich 14,2 % aufweisen, zeigen die kodierenden Bereiche mit 20,6 % einen deutlich höheren GC-Gehalt. Besonders prägnant ist der Unterschied in der Basenzusammensetzung auch zwischen Exons und Introns. Introns weisen einen GC-Gehalt auf, der dem intergenetischer Bereiche gleicht. Die Basen G und C sind dagegen in Exons, zurückzuführen auf Limitationen hinsichtlich der Kodonnutzung, deutlich stärker repräsentiert (21,4 %).

Tab. 4.12: Der GC-Gehalt der Kodonpositionen

	GC1	GC2	GC3
P. mur	29,2	22,0	13,5
P. fal	38,1	27,9	16,4

P. mur, *Polychromophilus murinus*;
P. fal, *Plasmodium falciparum*

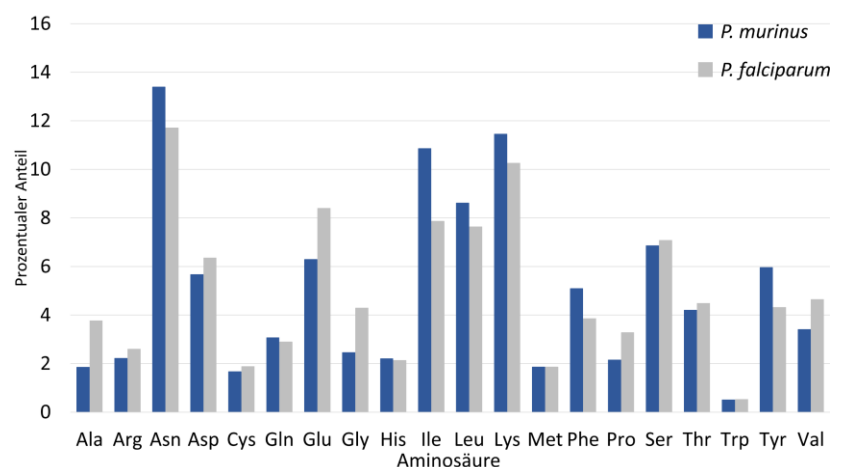


Abb. 4.29: Verwendung der Aminosäuren im Vergleich.

Dennoch ist der GC-Gehalt auch im kodierenden Bereich deutlich geringer als z. B. der von *P. falciparum*. Bemerkbar macht sich dies auch in der Basenkomposition der verwendeten Kodons. Im Vergleich werden die Basen A und T bei *P. murinus* an allen drei Nukleotidpositionen wesentlich häufiger verwendet (Tab. 4.12, vgl. auch Tab. 7.17). Der GC-Gehalt der einzelnen Positionen unterscheidet sich jedoch erheblich. Dabei sinkt der GC-Gehalt mit steigender Kodonposition ($GC1 > GC2 > GC3$), sodass vor allem die dritte Position stark hinsichtlich der Nutzung von A oder T selektioniert ist. Interessanterweise werden die Nukleotide A oder T an der dritten Position bei *P. falciparum* etwa zu gleichen Teilen verwendet, während bei *P. murinus* eine wesentlich stärkere Tendenz zur Nutzung der Base T deutlich wird (46,4 %; vgl. Tab. 7.17). Besonders häufig ist daher auch das Triplet AAT (11,7 %) vertreten, das für die Aminosäure Asn kodiert. Neben Asn spielen auch Lys und Ile, mit zusammengerechnet etwa 35,7%, eine statistisch signifikante Rolle in den Proteinen von *P. murinus* (Abb. 4.29).

Insgesamt weist das Proteom von *P. murinus* eine vergleichbare Verwendung von Aminosäuren wie *P. falciparum* und einige andere Plasmodien auf (Yadav et al., 2012). Die verstärkte Verwendung von Asn, Asp, Glu, Ile, Leu, Lys und Ser führen in den Proteinen dieser zu Regionen geringer Komplexität. Auch in dem vorliegenden Proteom kann ein erhöhter Anteil an einfachen Aminosäuren-Wiederholungen festgestellt werden. Im vorhandenen Datensatz bilden diese 1.363 homopolymerartige Regionen, die teilweise auch beträchtliche Längen erreichen können (nicht dargestellt). Besonders häufig sind dabei Regionen bestehend aus Asn (544), Lys (324), Asp (196) und Ser (119). Die putative ubiquitin-spezifische Protease (PF3D7_0403500-homolog) enthält mit 86 aufeinanderfolgenden Asn die längste homopolymerartige Aminosäuresequenz. Während diese, ähnlich wie auch bei Poly-Lys und Poly-Asp, fast ausschließlich aus Trinukleotid-Wiederholungen desselben Kodons gebildet wird, entstehen Ser-Wiederholungen in der Regel aus allen sechs möglichen Kodon-Kombinationen. Insgesamt weisen 635 (19,7 %) der vorliegenden Proteine homopolymerartige Regionen mit mindestens fünf konsekutiv identischen Aminosäuren auf. 21 (3,3 %) dieser Proteine sind zusätzlich mit einem Signalpeptid ausgestattet, während 179 (28,3 %) mindestens eine transmembrane Domäne besitzen. Für die Mehrheit der homopolymerartigen Regionen der Membranproteine (79,3 %) konnte eine nicht-zytoplasmatische Lokalisation vorhergesagt werden (nicht dargestellt). Lys- und Asn-reiche Regionen werden dabei gleichermaßen häufig in den nicht-zytoplasmatischen Raum exponiert (je 32 %). Dennoch sind homopolymerartige Regionen in Proteinen mit transmembranen Domänen statistisch nicht stärker vertreten als bei einer Zufallsverteilung erwartet.

4.4.4.2 Vorhersage zur Proteinlokalisierung

Von allen annotierten Strukturproteinen konnten 738 (22,9 %) als Membranproteine identifiziert werden (Abb. 4.30, vgl. 3.18.1). Annähernd ebenso viele Proteine weisen eine Signalsequenz für einen Transport in das Mitochondrium oder den Apikoplasten des Parasiten auf (20,4 %). Allerdings weisen vergleichsweise wenig Proteine das klassische Signalpeptid auf (11,3 %), das für die Initiierung des Proteintransports benötigt wird. Auch bei *P. falciparum* werden viele Signalpeptide, aufgrund von Sequenzverlängerungen oder abweichenden Strukturen, nicht als solche korrekt vorhergesagt (Sargeant et al., 2006). Proteine, die z. B. eine Signalsequenz für einen

Transport in das Mitochondrium von *P. murinus* enthalten, jedoch kein N-terminales Signalpeptid aufweisen, wurden daher dennoch als potenzielle Kandidaten mit einer entsprechenden Lokalisation interpretiert. 91 Proteine mit Signalpeptid weisen dagegen keine weitere Signalsequenz auf, welche die Proteine in bestimmte Zellkompartimente dirigieren, 47 davon ohne transmembrane Domänen. Die Mehrheit dieser wird mit hoher Wahrscheinlichkeit aus dem Parasiten sezerniert. Neben einigen Proteinen unbekannter Funktion handelt es sich dabei um Proteine, wie CelTOS (cell-traversal protein for *Plasmodium* ookinetes and sporozoites, Pmur00569), PSOP17 (Secreted ookinete protein, PF3D7_1216600-homolog) und GAMA (GPI-anchored micronemal antigen, PF3D7_0828800-homolog), die während der infektiösen Stadien freigesetzt werden und an der Oberfläche des Parasiten lokalisiert sind (Arumugam et al., 2011; Jimah et al., 2016). Nur zwei der putativ sezernierten Proteine weisen zusätzlich das klassische pentamere Export Element (PEXEL, RxLxE/Q/D) auf, das für einen Export der Proteine in die Wirtszelle benötigt wird. Dabei handelt es sich um ein konserviertes Membranprotein unbekannter Funktion (PF3D7_1214900-homolog, RPLIE) und um das TVN-junction protein 1 (tubovesicular membrane network junction protein 1, PF3D7_0310400-homolog, RLLID).

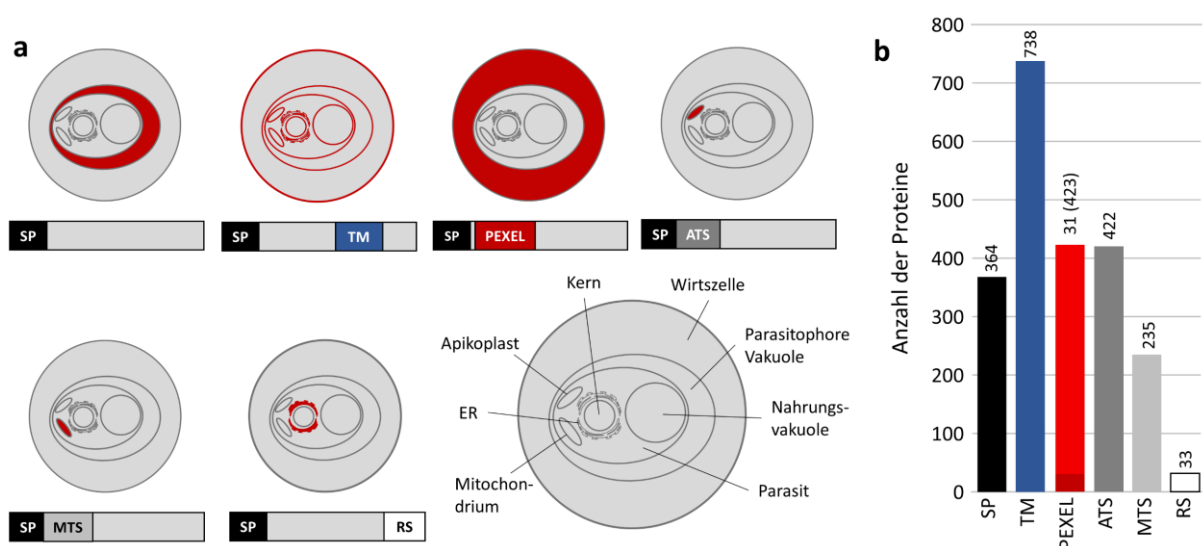


Abb. 4.30: Analyse der putativen Lokalisation annotierter Proteine von *P. murinus*.

Schematische Darstellung der Zielkompartimente (rot) der intrazellulären Form des Parasiten mit den dazugehörigen Signalsequenzen im Protein (a) und Anzahl der identifizierten Proteine (b). Hellrot: alternative PEXEL Motive. ATS, apicoplast targeting signal; MTS, mitochondrial targeting sequence; PEXEL, *Plasmodium* export element; RS, endoplasmic reticulum retention signal; SP, Signalpeptid; TM, transmembrane Domäne. (Abb. a nach Matz et al., 2018)

Tatsächlich weisen mit 31 nur relativ wenige Proteine von *P. murinus* ein klassisches PEXEL auf (1 %). Wesentlich häufiger findet man dagegen Motive, die als alternative PEXEL interpretiert werden könnten (RxLxx, R/KxL/IxE/Q/D oder R/KxL/IxxE/Q/D). 423 Proteine, die ein Signalpeptid oder eine transmembrane Domäne besitzen, weisen eine solche Sequenz auf. Allerdings ist davon auszugehen, dass mit den eingegrenzten Sequenzen, die Anzahl der putativ exportierten Proteine von *P. murinus* deutlich überschätzt wird. Lediglich 282 dieser Proteine (davon 20 mit klassischem PEXEL) weisen beispielsweise keine weitere Signalsequenz auf, die einen Transport in das Endoplasmatische Retikulum, Mitochondrium, oder den Apikoplasten initiieren. Dennoch finden sich unter den eingegrenzten Proteinen einige, die nicht in die Wirtszelle

exportiert werden, sondern vielmehr in anderen Zellkompartimenten lokalisiert sind, wie z. B. die Proteine ROM1 (rhomboid protease 1, PF3D7_1114100-homolog) oder VTC (vacuolar transporter chaperone, PF3D7_0622300-homolog).

4.5 Identifikation wirtzell-interagierender Proteine

Im Genom von *P. murinus* konnten insgesamt 3.220 proteinkodierende Gene identifiziert werden. Lediglich 74 Proteine (2,3 %) lieferten in Blast-Analysen keine signifikanten Treffer innerhalb der Haemosporida oder zu anderen hinterlegten Proteinen der nr-Datenbank des NCBI (vgl. 3.11.2). Die Mehrheit aller annotierten Proteine konnte jedoch orthologen Proteinen aus dem Phylum der Haemosporida zugeordnet werden (Abb. 4.31 a). Fast alle Proteine (3.141; 97,5 %) weisen dabei die höchste Signifikanzidentität zu Proteinen von *P. relictum* auf. Ein besonders hoher Anteil scheint zudem hochkonserviert im Taxon der Haemosporida zu sein: 2.690 Proteine (83,5 %) weisen Orthologe in den Arten *H. tartakovskyi*, *P. falciparum* und *P. relictum* auf. Dagegen konnten jedoch keine Proteine identifiziert werden, die nur in *P. murinus* und *H. tartakovskyi* konserviert sind. Dies spiegelt sich auch bei der Zuordnung der Proteine zu orthologen Gruppen (OG) wider, die sowohl homologe Proteine unterschiedlicher Arten als auch paraloge Proteinen derselben Art einschließen. Hier konnten 3.099 Proteine in 2.846 OG eingeteilt werden, die innerhalb der Aconoidasida konserviert sind (Abb. 4.31 b).

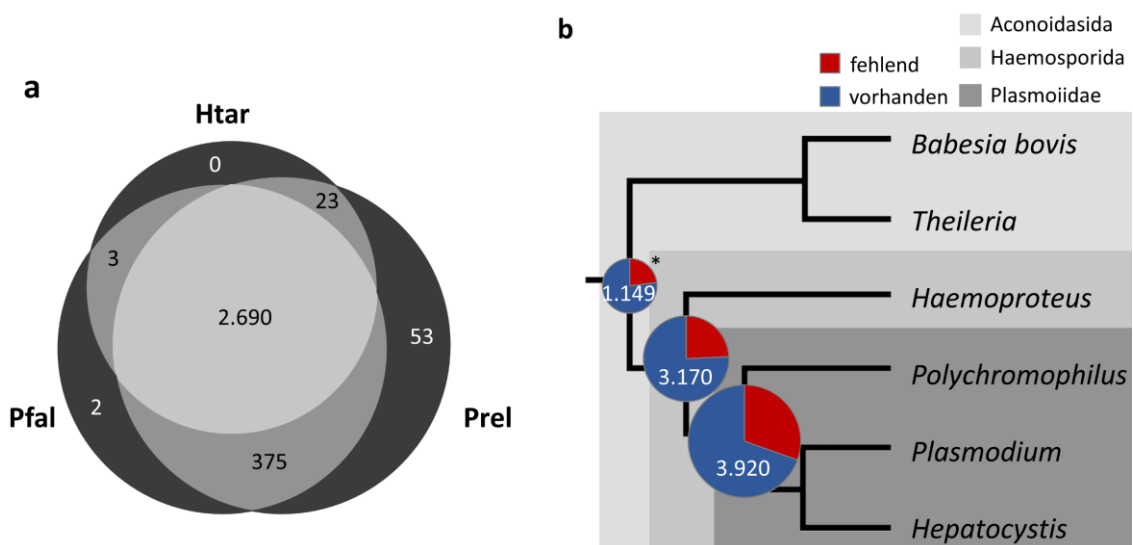


Abb. 4.31: Homologe Proteine innerhalb der Haemosporida.

(a) Schematische Darstellung der blastp-Analyse der hier identifizierten Proteine. Htar, *H. tartakovskyi*; Pfal, *P. falciparum*; Prel, *P. relictum*. (b) Evolution der Haemosporida, gekennzeichnet durch den Erwerb konservierten Proteingruppen. Tortendiagramme stellen jeweils den Anteil an diesen orthologen Gruppen, die auch in *P. murinus* nachgewiesen werden konnten (blau). Der Anteil der orthologen Gruppen, die sich andere Vertreter der jeweiligen Rangstufe teilen, allerdings nicht im vorhandenen Datensatz repräsentiert sind, wurde rot dargestellt. Diagramme sind proportional zur Anzahl der gemeinsamen orthologen Gruppen jeder Rangstufe. Zur besseren Übersicht wurde das kleinste Tortendiagramm doppelt so groß dargestellt (gekennzeichnet durch *).

1.149 OG stellen dabei das gemeinsame Kerngenom von Piroplasmiden und Haemosporiden dar. Mit Abnahme der hierarchischen Rangstufe nimmt der Konservierungsgrad des Kerngenoms und

damit auch die Anzahl an gemeinsamen OG zu. Gleichzeitig erhöht sich jedoch auch der Anteil an fehlenden Orthologen von *P. murinus*. Während lediglich 265 der OG (23,1%) fehlen, die sich andere Vertreter der Aconoidasida teilen, repräsentiert der vorliegende Datensatz nur 75,7 % (2.400 OG) des Kerngenoms der Haemosporida und 69,5% (2.728 OG) des Kerngenoms der Plasmodiidae.

4.5.1 Konservierte Proteine der Aconoidasida

Im vorliegenden Genom konnten 905 proteinkodierende Gene annotiert werden, die über alle hier untersuchten Vertreter der Haemosporida und Piroplasmida konserviert sind und 895 OG zugehören. Für rund die Hälfte dieser (467 Gene) konnten die dazugehörigen Genprodukte in Proteinklassen eingeteilt werden (Abb. 4.32). Ein großer Teil besteht dabei aus Proteinen, die eine entscheidende Rolle bei der Aufrechterhaltung grundlegender Zellfunktionen spielen. Besonders häufig konnten Proteine identifiziert werden, die an Nukleinsäuren binden (26,6 %), die Umwandlung von Stoffwechselzwischenprodukte induzieren (19,5 %) oder bei der Translation agieren (15,9 %). Aber auch konservierte proteinmodifizierende Proteine (11,4 %) und solche, die an Transportprozessen beteiligt sind (13,6 %) oder als wesentliche Bestandteile des Zytoskeletts fungieren (4,5 %), stellen einen Großteil des grundlegenden Proteinrepertoires der Aconoidasida dar.

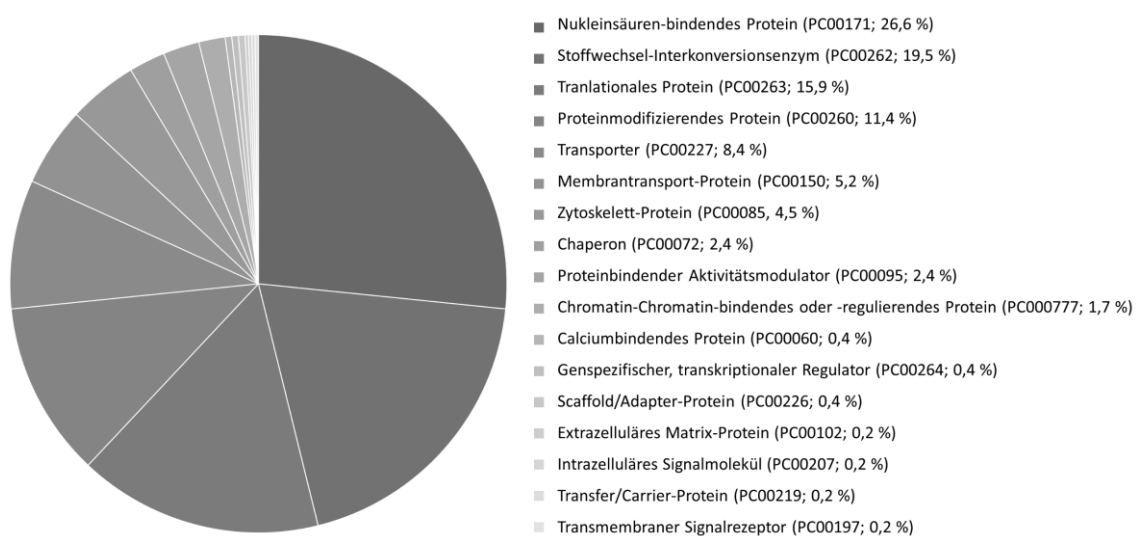


Abb. 4.32: Einteilung der konservierten Proteine der Aconoidasida in Proteinklassen.

Zur weiteren Untersuchung der funktionellen Natur des Aconoidasida-Kerngenoms wurden die identifizierten Proteine Gene Ontology (GO) Kategorien zugeordnet. Interessanterweise konnten, mit Ausnahme von lediglich zwei Proteinen, alle Proteine mindestens zu einer GO-Kategorie gruppiert werden. Wie zu erwarten sind dabei vor GO-Kategorien überrepräsentiert, die konstitutive Prozesse der Zelle widerspiegeln. Dazu zählen Prozesse der biologischen Regulation (GO:0065007) und Homöostase (GO:0042592), Kohlenhydratabhängige-Stoffwechselprozesse wie Glykolyse (GO:1901135, GO:0006096) und Oxidations-Reduktions-Prozesse (GO:0055114),

aber auch nukleinsäureabhängige Prozesse wie Translation (GO:0006412) und Replikation (GO:0006260). Genprodukte, die diesen und untergeordneten biologischen Prozessen zugeordnet werden können, sind in dem *Aconoidasida*-Kerngenom deutlich überrepräsentiert und mindestens dreimal stärker vertreten als in einem zufälligen Datensatz erwartet (nicht dargestellt).

Unter den überrepräsentierten Proteinen finden sich daher auch vermehrt Proteine, die DNA binden (GO:0003677) und z. B. eine Hydrolase-Aktivität (GO:0003678, GO:0003724) aufweisen oder als Translationsfaktoren fungieren (GO:0045182). Dazu zählen beispielsweise auch eine Reihe an Transkriptionsfaktoren die Apetala2-Domänen (AP2) aufweisen. Diese werden zu der ApiA2 Proteinfamilie gezählt und spielen bei der Genregulation diverser Stadien unterschiedlicher Apikomplexa in essenzielle Rollen (Painter et al., 2011).

Auch bei der Zuordnung zu zellulären Komponenten spiegelt sich das bisher charakterisierte *Aconoidasida*-Kerngenom wider. Viele Proteine werden als Komponenten von Zellkern (GO:0005634) und Chromosomen (GO:0005694) oder Ribonukleoproteinkomplexen (GO:1990904) gezählt. Überrepräsentiert sind zudem einzelne anatomische Einheiten, wie Bestandteile des DNA Replikation Faktor C Komplex (GO:0005663), des Proteasom-Regulationspartikel oder des mitochondrialen Respirasoms (GO:0005746). Auch andere Bestandteile von Organellen können im *Aconoidasida*-Kerngenom zugeordnet werden, sind allerdings nicht wesentlich häufiger vertreten als durch eine zufällige Verteilung zu erwarten. Anders sieht es mit Bestandteilen der für Apikomplexa typischen Pellikula aus (GO:0020039). Dazu gehören beispielsweise Myosin-A, Proteine des inneren Membrankomplexes und die glideosom-assoziierte Proteine GAP45 und GAP50. Diese sind für die Ausbildung des Motorkomplexes von Bedeutung und spielen wichtige Rollen für die Motilität und die Invasion der *Aconoidasida* (Thomas et al., 2012).

Proteine, die in der Wirtszelle lokalisiert sind, sind im *Aconoidasida*-Kerngenom dagegen unterrepräsentiert. Ausnahme dazu bilden nur wenige Proteine, wie beispielsweise ein Protein, das an der Detoxifizierung der Hämgruppe beteiligt ist. Auch molekulare Funktionen, die mit Wirtzellinteraktionen in Verbindung stehen, wie die Zelladhäsion (GO:0050839) oder die Bindung an die Wirtszelloberfläche (GO:0046812), sind unter den im Taxon *Aconoidasida* konservierten Proteine unterrepräsentiert. Lediglich die Proteine AMA-1 (apical merozoite antigen 1), RON2 und RON5 (rhoptry neck protein 2 und 5) konnten identifiziert werden, die mit Proteinen des Wirts interagieren. Dabei spielen die Proteine eine besondere Rolle als Ankerpunkt bei der Invasion, die in Form einer so genannten moving junction wahrscheinlich durch eine Bindung mit dem Zytoskelett des Wirts zustande kommt (Cao et al., 2009). Insgesamt gibt es auch nur wenige gemeinsame Proteine von Piroplasmiden und Haemosporiden, die in biologische Prozesse involviert sind, die mit dem Wirt in Verbindung stehen. Dazu zählt beispielsweise auch das HSP101 (Heat shock protein 101), das in Plasmodien als Teil des Translokalisationskomplex PTEX (*Plasmodium* Translocon of Exported Proteins) fungiert (Beck et al., 2014).

4.5.2 Konservierte Proteine der Haemosporida

Das Kerngenom der Haemosporida unterscheidet sich von dem letzten gemeinsamen Vorfahren der Aconoidasida durch einen Zugewinn einer erheblichen Anzahl neuer Genprodukte. Neben den im Taxon der Aconoidasida konservierten Proteine weisen die Haemosporida 2.102 weitere konservierte Proteine auf. 42 dieser Proteine fallen dabei in OG, die bereits durch das Kerngenom der Aconoidasida kodiert werden. Zusätzlich konnten 39 Duplikationen identifiziert werden, die alle Haemosporiden gemeinsam haben. Insgesamt sind die Haemosporiden durch den Zugewinn von 2.021 OG gekennzeichnet. Davon konnten 1.516 OG bestehend aus 1.549 Proteinen auch in dem vorliegenden Genom identifiziert werden (73,7 %). Im Gegensatz zu den konservierten Proteinen der Aconoidasida konnte ein weitaus geringerer Anteil der entsprechend konservierten Proteine der Haemosporida in Proteinklassen eingeteilt werden (37,8 %). Etwa 80 % der Proteine werden dabei in die gleichen Proteinklassen eingeteilt, die, wenn auch in einer leicht abgewandelten Reihenfolge, auch bereits im Kerngenom der Aconoidasida überrepräsentiert sind (nicht dargestellt). Obwohl auch hier am Häufigsten Proteine identifiziert werden konnten, die an Nukleinsäuren binden (21,4%), bilden proteinmodifizierende Proteine beispielsweise einen leicht höheren Anteil (12,9 %), als solche, die bei der Translation beteiligt sind (12,3 %).

Die hohe Anzahl an Proteinen, die nicht in bekannte Proteinklassen eingeteilt werden kann, steht mit dem Zugewinn einer großen Anzahl von Proteinen unbekannter Funktion im Zusammenhang: 647 Proteine (41,8 %) konnten in *P. murinus* identifiziert werden, die in allen hier untersuchte Arten der Gattungen *Plasmodium* und *Haemoproteus* konserviert sind, bislang funktionell jedoch nicht weiter eingeordnet werden konnten. Daher überrascht der hohe Anteil an Proteinen, dem mindestens eine GO-Kategorie zugeordnet werden konnte (99,3 %). Interessanterweise gibt es hier jedoch keine statistisch signifikanten Anreicherungen (nicht dargestellt). Lediglich Proteine, die in die biologischen Prozesse der Adhäsion (GO:0022610), der interspezifischen Interaktion (GO:0044419) oder der Antwort auf Stimuli (GO:0050896) involviert sind, sind im vorliegenden Datensatz unterrepräsentiert. Dazu passt auch, dass nur relativ wenige Proteine im Cytoplasma oder der Zellmembran der Wirtszelle lokalisiert sind (GO:0030430, GO:0033644). Im Parasiten selbst konnten 332 Proteine identifiziert werden, die als Membrankomponenten (GO:0031224) eingeordnet werden. Ein Fünftel dieser Proteine kann dabei funktional zu Bindungsprozessen (GO:0005488) zugeordnet werden. Die Mehrheit übt jedoch eine katalytische Aktivität aus (GO:0003824; 47 %) oder ist in Transportgeschehen involviert (GO:0005215; 28,7 %) (Abb. 4.33).

Unter den Transportproteinen finden sich einerseits essenzielle Proteine, welche die Haemosporiden mit Nährstoffen versorgen. Dazu gehören die Transporter für Vitamin B5, Hexose und Cholin (StAR-related lipid transfer protein) aber im weiteren Sinne auch das Dynamin 1, das eine wichtige Rolle bei der Aufnahme von Hämoglobin spielt. Zugleich weist das ursprüngliche Genom der Haemosporiden auch wichtige Bestandteile des zytoplasmatischen tubovesikulären Netzwerks (TVN) Das TVN ist als Erweiterung der parasitophoren Membran in der Wirtszelle verzweigt und an der Nährstoffaufnahme des Parasiten beteiligt (Lauer, 1997; Lauer et al., 2000). Wichtige TVN-Proteine, die in Vertretern der Haemosporida konserviert sind, sind beispielsweise das TVN-junction protein 1 und Rab5b (ras-related protein Rab-5B).

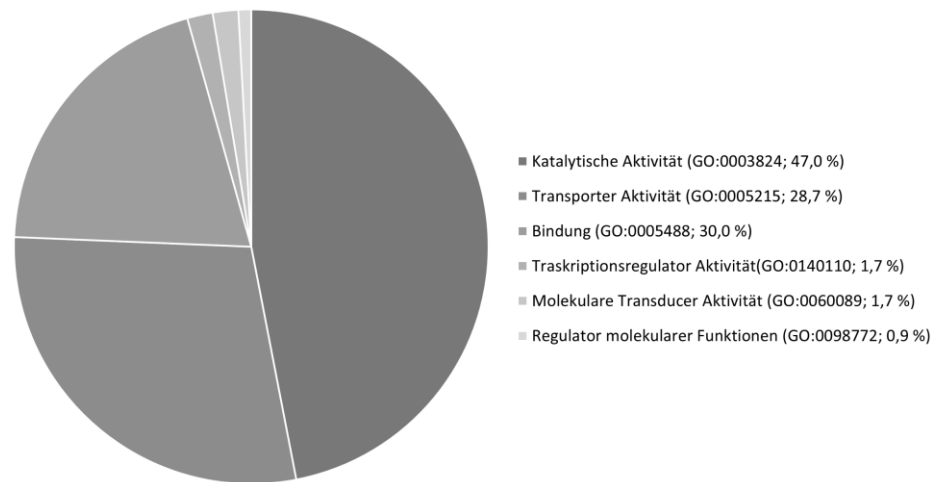


Abb. 4.33: Molekulare Funktion membranassoziierter konservierter Proteine der Haemosporida.

Im Gegensatz zu Vertretern der Piroplasmida müssen Haemosporiden bei der Nährstoffaufnahme zusätzlich auch die parasitophore Membran überwinden. Eine besondere Rolle kommt dabei dem Protein EXP2 (exported protein 2), das als Kanal für den Nährstofftransport fungiert, aber auch Teil des Translokationskomplex PTEX darstellt (Garten et al., 2018). Die parasitophore Membran stellt jedoch nicht nur für Nährstoffe und den Proteinexport eine zusätzliche Barriere dar, auch der Zellaustritt wird von dieser beeinflusst. Proteine, die dies in unterschiedlichen Stadien gewährleisten, sind in dem vorliegenden Datensatz über alle Haemosporiden konserviert. Dazu gehören: LISP1 (liver specific protein 1), SUB1 (subtilisin-like protease 1), und PLP2 (perforin like protein 2) (Ishino et al., 2009; Tawk et al., 2013; Wirth et al., 2014).

Proteine, die dagegen an der Zellinvasion oder auch der Zelltransmigration beteiligt sind, lassen sich durch eine Lokalisierung in den Mikronemen oder Rhoptrien der Parasiten charakterisieren, oder werden an der Oberfläche der invasiven Stadien präsentiert. Im vorliegenden Datensatz konnte eine Reihe an Proteinen identifiziert werden, die über alle Haemosporida hinweg konserviert sind und eine entsprechende Lokalisierung aufweisen (Tab. 4.13). Von 23 Proteinen weisen 14 bei *P. murinus* bekannte Domäne auf, die Protein-Protein-Interaktion initiieren können. Dazu gehören EGF-like Domänen (RIPR), Thrombospondin Typ-1 (TSP1) Repeats (TRAMP, CTRP), von Willebrand Typ A Faktoren (CTRP), Tyrosin-Proteinkinase Ephrin Typ A/B Rezeptor-like Domänen (CRMP3 und 4), die 6-cys-Domäne (P36), und LCCL-Domäne (CCp4).

Unter den in allen Haemosporiden konservierten Proteinen enthalten lediglich zwei weitere Proteine noch eine TSP-1 (PF3D7_0822700-homologe) oder eine Tyrosin-Proteinkinase Ephrin Typ A/B Rezeptor-like Domäne (PF3D7_0530500-homolog). Während es sich bei PF3D7_0530500 um ein konserviertes Protein mit unbekannter Funktion handelt, wurde das Protein PF3D7_0530500 bei *P. falciparum* als TRP1 (thrombospondin-related protein 1) beschrieben, das eine wichtige Rolle beim Austritt der Sporozoiten aus dem Oozysten spielt und auch bei der Invasion in die Speicheldrüsen des Endwirts beteiligt ist (Klug et al., 2017).

Tab. 4.13: Putative Invasions- und Transmigrationsabhängige Proteine.

Mikronemen Proteine	Rhoprien Proteine	Oberflächenproteine
Cell traversal protein for PF3D7_ookinetes and sporozoites 1216600 (CelTOS)	Apical merozoite protein PF3D7_ (Pf34) 0419700	Ag-1 blood stage membrane PF3D7_ protein homologue (MAg-1) 0405200
Circumsporozoite-and PF3D7_ TRAP-related protein (CTRP) 0315200 *	Apical sushi protein (ASP) PF3D7_ 0405900	Cysteine repeat modular PF3D7_ protein 3 (CRMP3) 1208200
Gamete egress and sporozoite PF3D7_ traversal protein, putative 1449000 (GEST)	Armadillo-domain containing PF3D7_ rhoptry protein (ARO) 0414900	Cysteine repeat modular PF3D7_ protein 4 (CRMP4) 14754000
GPI-anchored micronemal PF3D7_ antigen (GAMA) 0828800	Repetitive organellar protein, PF3D7_ putative (ROPE) 0203000	Sporozoite surface protein PF3D7_ P36 (P36) 0317100
Microneme associated antigen PF3D7_ (MA) 0316000	Rhoptry protein 2, putative PF3D7_ (PRP2) 1320000	LCCL domain-containing PF3D7_ protein (CCp4) 0903800
Protein tyrosine phosphatase PF3D7_ (PTP)* 1113100	Thrombospondin-related PF3D7_ apical membrane protein 1218000 (TRAMP)*	Male gamete fusion factor PF3D7_ HAP2, putative (HAP2) 1014200
Rh5 interacting protein PF3D7_ (RIPR) 0323400		Parasite-infected erythrocyte PF3D7_ surface protein (PIESP1) 0310400
RHOMBOID-like protein PF3D7_ (RBL) 1114100		Phospholipase, putative PF3D7_ 0628300
secreted ookinete protein, PF3D7_ putative (PSOP20) 0715400		
Subtilisin-like protease 2 PF3D7_ (SUB2) 1136900		

* stellen ebenfalls Oberflächenproteine dar

4.5.3 Konservierte Proteine der Plasmodiidae

979 Proteine konnten identifiziert werden, die in den Vertretern der Familie Plasmodiidae konserviert sind, jedoch nicht in *H. tartakowskyi* nachgewiesen werden können. Insgesamt gehören die Proteine 790 orthologen Gruppen an. 220 davon schließen bereits Proteine mit ein, die über das Taxon Aconoidasida (212 OG) oder über das Taxon Haemosporida (8 OG) konserviert sind. Vor allem Proteine mit einer Transferase-Aktivität (GO:0016740) liegen unter den, wahrscheinlich aus Duplikationen hervorgegangenen, Proteinen überrepräsentiert vor (nicht dargestellt). Zudem sind auch Proteine überrepräsentiert, die in den Rhoptrien lokalisiert sind. Dazu zählen RON3, RON4 und RON12, SPATR (Secreted protein with altered thrombospondin repeat domain), RAP3 (Rhoptry-associated protein 3), RhopH2 und RhopH3 (high molecular weight rhoptry protein 2 und 3) sowie RALP1 (rhoptry-associated leucine zipper-like protein 1). Keines dieser Proteine konnte auch im vorliegenden Datensatz identifiziert werden. Tatsächlich konnten im Genom von *Polychromophilus* nur recht wenige Proteine nachgewiesen werden, die auch über alle Vertreter der Gattungen *Plasmodium* und *Hepatocystis*, über das Kerngenom der Haemosporida hinaus, konserviert sind. Lediglich 332 Proteine (33,9 %) sind in dem vorliegenden Datensatz enthalten. Im Vergleich zum Genom von *P. falciparum* sind dabei Proteine, die mit Komponenten des Wirtes interagieren unterrepräsentiert (nicht dargestellt). Nur zwei Proteine konnten identifiziert werden, die in das Zytoplasma des Wirtes transportiert werden (GO:0033655). Dabei handelt es sich um TRX2 (Thioredoxin 2) und GDPD (Glycerophosphodiester phosphodiesterase), die in bei der Antwort auf Hypoxie (P00030) und oxidativen Stress (P00046)

involviert sind. TRX2 ist zudem eine wichtige Komponente des Translokationskomplex PTEX (de Koning-Ward et al., 2009).

Überrepräsentiert unter den in *P. murinus* identifizierten Proteinen sind dagegen solche, die nicht weiter klassifiziert werden konnten. Dies schließt sowohl die Gruppierung nach biologischen Prozessen und nach molekularer Funktion als auch nach zellulären Komponenten mit ein. Tatsächlich handelt es sich bei 179 Proteinen (53,9 %) um solche, die zwar über die untersuchten Gattungen hinweg konserviert sind, denen bislang jedoch noch keine Funktion zugeordnet werden konnte. Unter den restlichen Proteinen finden sich vor allem solche, die an zellulären Prozessen beteiligt (GO:0009987) sind und zelluläre Komponenten unterschiedlichster Organellen darstellen (GO:0043229). Proteine, die mit Mikronemen assoziiert sind, konnten nicht identifiziert werden. Allerdings enthielt der Datensatz zwei Proteine, die in den Rhoptrien der Parasiten lokalisiert sind. Einerseits handelt es sich dabei um Rab11a, das eine wichtige Rolle bei der Ookinetenentwicklung spielt, und andererseits um RA (rhoptry associated adhesin), das bei der Sialinsäuren-abhängigen Invasion in Erythrozyten beteiligt ist (Anand et al., 2016; Patil et al., 2019).

Auch Proteine, die Domänen aufweisen, die Protein-Protein-Interaktionen zwischen Parasit und Wirt initiieren könnten, stellen eine Seltenheit im Datensatz dar. Lediglich ein Protein mit dem Ankyrin Repeat (PF3D7_1250900-homolog), ein Protein mit EGF-like Domäne (PF3D7_0623300-homolog) und ein Protein mit 6-Cys Domäne (PF3D7_1346800-homolog, P47) konnten identifiziert werden. In fünf weiteren Proteinen konnten WD40 Repeat Domänen vorhergesagt werden. Mit Ausnahme von WDR16 (PF3D7_1348700-homolog, WD repeat-containing protein 16), das zu der Zilien- und Flagellen-assoziierten Protein 52 Familie (PTHR13720:SF14) gezählt wird, sind die Proteine bislang jedoch noch charakterisiert (PF3D7_0416900-, PF3D7_0524100-, PF3D7_1250900- und PF3D7_1459200-homologe).

Vergleichbar mit dem Anteil der Proteine unbekannter Funktionen konnte auch in einem großen Anteil an Proteinen keine funktionellen Domänen identifiziert werden. Nur 141 Proteine (42,5 %) weisen eine charakteristische Domänenarchitektur auf. Interessanterweise ist im konservierten Proteom der Haemosporida jedoch der Anteil an Proteinen mit niedriger Komplexität (42,8 %) im Vergleich zum gesamten Datensatz erniedrigt (55,8%). Auch der Anteil der Proteine, die homopolymerartige Regionen mit mindestens fünf konsekutiv identischen Aminosäuren aufweisen, ist hier erniedrigt (12,3 %). Der Anteil an Proteinen mit Transmembrandomänen (30,7 %) ist dagegen leicht erhöht, während besonders wenige Proteine mit klassischem Signalpeptid vorhanden sind (5,7 %), darunter auch das bereits erwähnte Protein P47 und das PF3D7_0623300-homolog.

4.5.4 Plastizität des Exportsignals

Im vorliegenden Genom konnten insgesamt 61 Proteine identifiziert werden, die exportiert werden und damit in den infektiösen Stadien der Parasiten auf der Plasmamembran präsentiert oder in die Wirtszelle transportiert werden. Eine Übersicht über diese Proteine ist in Tab. 7.18 dargestellt. Die Mehrheit dieser (37; 60,7 %) ist innerhalb des Taxons Haemosporida konserviert und erst nach der Abspaltung von den Piroplasmiden entstanden. Lediglich 14 Proteine (23 %), die exponiert oder

exportiert werden, sind dagegen über die Aconoidasida hinweg konserviert. Ähnlich viele Proteine (11, 18 %) sind dagegen über die Familie Plasmodiidae konserviert und nach der Abspaltung von der Gattung *Haemoproteus* entstanden. Allerdings sind sechs Proteine davon, darunter vier 6-Cys Proteine, in der Gattung *Hepatocystis* wieder verloren gegangen.

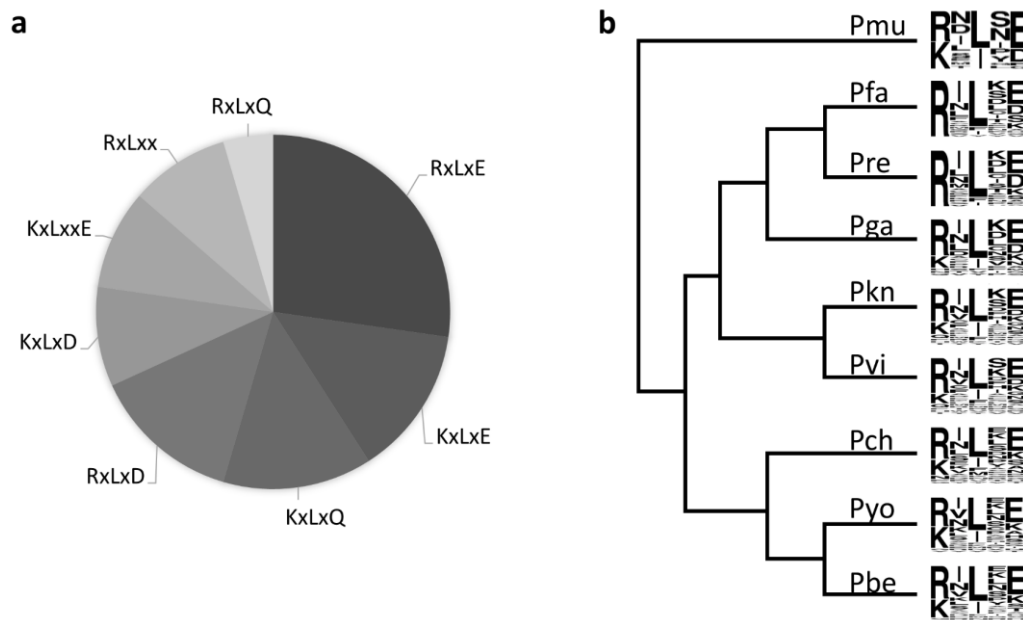


Abb. 4.34: Plastizität und Evolution des PEXEL innerhalb der Familie der Plasmodiidae.

Häufigkeit der alternativen PEXEL-Motive im Genom von *P. murinus* (a). Evolutive Entstehung unterschiedlicher PEXEL-Motive (b) Sequenzen und phylogenetischer Stammbaum wurden aus (Pick et al., 2011) übernommen. Pmu, *P. murinus*; Pfa, *P. falciparum*; Pre, *P. reichenowi*; Pga, *P. gallinaceum*; Pkn, *P. knowlesi*; Pvi, *P. vivax*; Pch, *P. chabaudi*; Pyo, *P. yoelii*; Pbe, *P. berghei*.

Rund die Hälfte dieser Proteine weist ein PEXEL oder PEXEL-ähnliches Motiv auf (36, 59 %). 18 Proteine davon, wurden bereits auch von (Pick et al., 2011) untersucht. Wie auch bei alternativen PEXEL außerhalb der Laverania, ist das Motiv im vorliegenden Genom deutlich schwächer konserviert und weist gewisse Variabilität in der Sequenz auf (Abb. 4.34). Von allen identifizierten Motiven weisen 20 ein Lysin (K) an der ersten Position auf, das damit deutlich häufiger vorhergesagt wurde als das klassische Arginin (R) (Abb. 4.34 a). Auch an der dritten Position, das in Plasmodien fast immer ein hydrophobes Leucin (L) enthält, weisen die orthologen Proteine von *P. murinus* eine deutlich höhere Variabilität auf. Mehrere Proteine weisen ein Isoleucin (I) an dritter Stelle auf. Im Vergleich zu anderen alternativen PEXEL innerhalb der Gattung *Plasmodium* ist der Anteil mit isoleucinhaltigen PEXEL sogar geradezu auffällig (41,7 %) (Abb. 4.34 b). Die fünfte konservierte PEXEL-Position, die als die am wenigsten konservierte Position angesehen wird, ist dagegen geradezu konserviert (Hiller et al., 2004; Marti et al., 2004; Pick et al., 2011). Vor allem die Glutaminsäure (E) ist dabei überrepräsentiert. Asparaginsäure (D) und Glutamin (Q) treten dagegen seltener auf. In den Vertretern der Gattung *Plasmodium* scheint die fünfte Position kaum konserviert. Hier steht teilweise Serin (S), Lysin (K), Asparagin (N), Alanin (A) oder Threonin (T) an letzter Stelle.

4.5.5 Putative Protein-Protein-Interaktionen zwischen Parasit und Wirt

Von allen in *P. murius* identifizierten Proteinen, die putativ exportiert werden und deren orthologe Proteine in den nahverwandten Plasmodien an der Zellinvasion oder der Zalltransmigration beteiligt sind, sind 11 Proteine hinsichtlich Protein-Protein-Interaktionen mit dem Wirt besonders interessant. In einer vorausgegangenen Publikation wurden in einen orthologiebasierten Ansatz für die entsprechend orthologen Proteine aus *P. falciparum* Interaktionen zu Proteinen des Menschen abgeleitet (Cuesta-Astroz et al., 2019). Vorausgesagt wurden dabei 17 Interaktionspartner (Abb. 4.35). Proteine des Zytoskeletts sind dabei überrepräsentiert (GO:00005856). Neun Proteine, die als mögliche Interaktionspartner fungieren, stellen dabei Typ II Keratine dar. Lediglich für drei der Proteine von *P. falciparum* konnte keine Interaktion mit Keratinen vorausgesagt werden. Darunter die Dipeptidyl Aminopeptidase 3 (PF3D7_0404700, DAPAP3), ein Protein mit 6-Cys Domäne (PF3D7_0404900, P41) und das Rhoptries assoziierte Adhesin (PF3D7_1012200, RA). Für die restlichen Proteine, darunter beispielsweise auch die bereits aufgeführten Proteine AMA-1, RON2 und RON11, GAMA und TRAMP, wurde dagegen eine Interaktion mit jedem der Keratine (KRT1, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT75, KRT8 und KRT85) vorausgesagt. Auch in den verwandten Parasiten *P. vivax* und *P. knowlesi* sind diese Interaktionen vorhergesagt worden (Cuesta-Astroz et al., 2019).

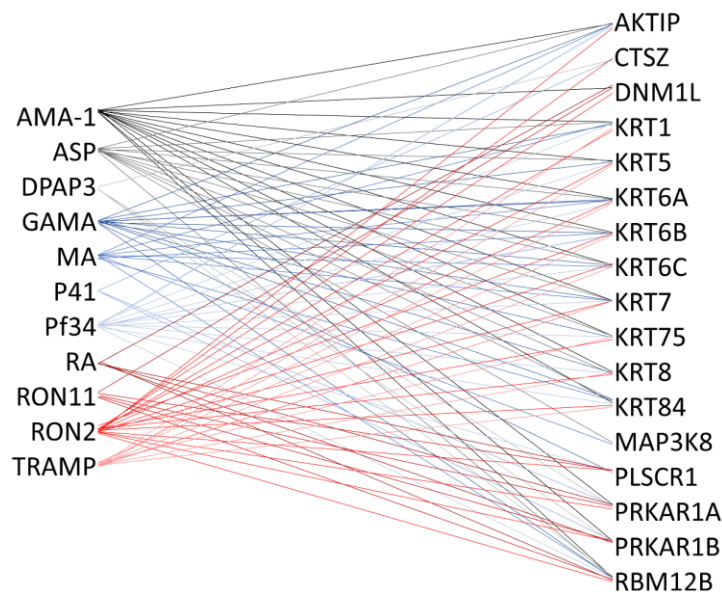


Abb. 4.35: Protein-Protein-Netzwerk zwischen parasitären, exportierten Proteinen und dem Wirt.

Darstellung der interagierenden Proteine basiert auf (Cuesta-Astroz et al., 2019). Linke Seite stellt Proteine von *P. falciparum* dar, während auf der rechten Seite die Interaktionspartner des Menschen aufgelistet sind. AMA-1, apical membrane antigen 1; ASP, apical sushi protein; DPAP3, dipeptidyl aminopeptidase 3; GAMA, GPI-anchored micronemal antigen; MA, microneme associated antigen; p41, 6-cysteine protein P41; Pf34, apical merozoite protein, RA, rhoptry associated adhesin; RON2/11, rhoptry neck protein2/11; TRAMP, thrombospondin-related apical membrane protein; AKTIP; AKT-interacting protein; CTSZ, Cathepsin Z; DNM1L, Dynamin 1 Like; KRT, Keratin; MAP3K8, Mitogen-activated protein kinase kinase kinase 8; PLSCR1, Phospholipid Scramblase 1; PRKAR1A/B, protein kinase cAMP-dependent type I regulatory subunit alpha/beta; RBM12B, RNA binding motif protein 12B.

5 Diskussion

Ein Genom enthält sämtliche vererbare genetische Information zur Ausprägung morphologischer und physiologischer Merkmale und definiert so die Beschaffenheit eines jeden Organismus. Die Analyse eines Genoms verspricht wertvolle Einblicke in die Grundlagen eines Lebewesens. Mit Einführung der Next-Generation-Sequencing-Technologien hat sich unser Verständnis parasitärer Organismen in Hinblick auf Diversität, Entwicklung von Pathogenität und funktionaler Rolle in Wechselwirkung mit Umwelt und Wirt graduell verbessert (z. B. Hupalo et al., 2015; Greenwood et al., 2016). Kein Wunder also, dass das Interesse an Genomanalysen von Parasiten seit der ersten Sequenzierung des Genoms von *Haemophilus influenzae* im Jahr 1995 rapide zugenommen hat (Fleischmann et al., 1995). Im letzten Jahrzehnt ist auch die Anzahl vollständiger Genome von Vertretern des Phylums Apikomplexa stark angestiegen. 61 Genome unterschiedlicher Organismen sind bislang auf den Seiten des NCBI öffentlich verfügbar (Stand: November 2019). Die überwiegende Mehrheit dieser stammt von Vertretern, die eine große medizinische oder veterinärmedizinische Rolle spielen. Lediglich drei Genome passen nicht in dieses Schema: die Genome zweier Gregarinen und das von *Haemoproteus tartakovskyi*, der wie alle Malariaparasiten zu den Haemosporiden gezählt wird. Die sequenzierten Erreger repräsentieren nur einen kleinen Teil der Diversität der Haemosporida. Umfangreiche Untersuchungen unterschiedlichster Wildtierbestände charakterisierten eine immense Anzahl weiterer Parasiten, die auch in anderen Vertebraten malariaähnliche Krankheiten hervorrufen können (z. B. Marzal et al., 2011; Martinsen et al., 2016; Schaer et al., 2017; Boundenga et al., 2018). Die größte Vielfalt an Haemosporiden konnte dabei in Fledermäusen identifiziert werden (Perkins et al., 2016). Umso bemerkenswerter ist daher die Tatsache, dass abgesehen von einigen wenigen Genen, bislang keine molekularen Daten von Haemosporiden aus Fledermäusen in öffentlichen Datenbanken vertreten sind (vgl. Galen et al., 2018).

5.1 *Polychromophilus murinus* – Ein herausforderndes Genom-Sequenzierungsprojekt

Unabhängig von dem zu sequenzierenden Organismus und der gewählten Sequenzierstrategie beruhen alle Genomsequenzierungen auf einer optimalen Probenvorbereitung, die eine sorgfältige Extraktion von genomischer DNA des Zielorganismus einschließt. Viele parasitäre Arten lassen sich jedoch nicht so einfach isolieren: entweder weil sie nicht kultiviert werden können, oder weil sie in einer engen Beziehung mit einem Wirt oder anderen Organismen leben (Kumar et al., 2013). Bei *P. murinus* kommt hinzu, dass vor allem die Wasserfledermaus *Myotis daubentonii* als Wirt fungiert. Dabei handelt es sich um eine Fledermaus, die zwar in Deutschland nicht auf der roten Liste der gefährdeten Arten steht, europaweit allerdings streng gesetzlich geschützt ist (EUROBATS, 1991). Durch die enge Bindung zwischen Parasit, Fledermausfliege und Wasserfledermaus sind Labor- und Feldexperimente nicht unproblematisch durchführbar.

Eine direkte Sequenzierung des Parasiten aus unbehandelten Proben von einem der beiden Wirte ist aufgrund des überwältigenden Anteils an Wirts-DNA in nachgelagerten Analysen

problematisch. Die Schwierigkeiten bei der Genomsequenzierung eines Parasiten aus kontaminierten Ausgangsproben spiegeln sich auch in der Literatur wider. So erzeugen klinische Proben, die einen höheren Anteil als 80 % menschlicher DNA enthalten, nur eine geringe Anzahl an Reads, die auf das genetische Material des Parasiten *P. falciparum* zurückzuführen sind (Oyola et al., 2013). Aus DNA-Proben, in denen dagegen nur maximal 30 % der DNA auf den Wirt zurückzuführen waren, war eine Rekonstruktion des parasitären Genoms bereits zu 85 % möglich (Auburn et al., 2011). Für optimale Voraussetzungen zur Genomsequenzierung von *Polychromobilus murinus* wurden daher Methoden etabliert, die eine effiziente Isolierung des Parasiten ermöglichen.

5.1.1 Das unzertrennliche Parasiten-Wirt-System

Der einzigartige Lebenszyklus der Haemosporida bietet gleich mehrere Angriffspunkte, die für eine Isolierung des Parasiten aus einem der beiden Wirte optimal erscheinen. Vor allem die extrazellulären, invasiven Stadien, die Merozoiten und die Sporozoiten, der Parasiten sind in diesem Zusammenhang vielversprechend. Im Gegensatz zu Merozoiten, die bislang nur aus synchronisierten Zellkulturen von Plasmodien mit intraerythrozytärer Schizogonie isoliert werden können, sind Sporozoiten frei zugänglich (Boyle et al., 2010; Kennedy et al., 2012). Die Sporozoiten des Parasiten *P. murinus* werden in Oozysten im Mitteldarm der Fledermausfliege *N. kolenatii* gebildet und persistieren anschließend im Lumen der Speicheldrüsen (Gardner et al., 1988). Anders als bei Mücken, liegen die Speicheldrüsen von *N. kolenatii* neben dem Mitteldarm im anterioren Teil des Abdomens (Gardner et al., 1988). Eine Präparation ohne Beschädigung der Speicheldrüsen war in dieser Arbeit nicht möglich und scheint auch bereits in anderen Arbeiten mit Problemen verbunden gewesen zu sein (Witsenburg et al., 2015b). Die Präparation des Mitteldarms der Fledermausfliegen war dagegen ohne größeren Aufwand erfolgreich. Dennoch konnten in den Geweben keine oozystenähnlichen Strukturen identifiziert werden.

Obwohl die Fledermausfliegen fast ihr gesamtes Leben ausschließlich auf der Wirtsfledermaus verbringen (Szentiványi et al., 2016) und ihre Wirte, wie auch in dieser Arbeit, immer wieder eine hohe Infektionsrate gezeigt haben (Megali et al., 2011; Witsenburg et al., 2014), ist die Prävalenz von *P. murinus* in *N. kolenatii* geringer als erwartet. In lediglich fünf von 28 Fledermausfliegen konnte mittels PCR genetisches Material vom Parasiten nachgewiesen werden. Statistisch mag dies nicht unbedingt signifikant erscheinen, allerdings spiegelt sich die geringe Anzahl an Fledermausfliegen, die positiv auf *P. murinus* getestet werden können, auch in der Literatur wider (Witsenburg et al., 2015b; Ramasindrazana et al., 2018). Neben der erhöhten Mortalität wird dies auch auf eine verstärkte Immunantwort der Fledermausfliege zurückzuführen sein, die eine weitere Entwicklung des Parasiten limitiert (Witsenburg et al., 2015b).

Auch für *Plasmodium falciparum* ist die Entwicklung in der *Anopheles* Mücke ein kritischer Schritt: nur etwa 38 % aller Mücken, die infiziertes Blut aufnehmen, infizieren sich auch selbst mit dem Parasiten und bilden reife Oozysten aus (Gouagna et al., 1998). Besonders deutlich wird diese kritische Lebensphase bei *P. murinus*, wenn man sich vor Augen führt, dass in der ursprünglichen Beschreibung unterschiedlicher Differenzierungsstadien des Parasiten in nur sechs von 446 Fledermausfliegen Oozysten morphologisch nachgewiesen werden konnten (Gardner et al., 1988).

Das kürzlich veröffentlichte Genom von *P. relictum* hat gezeigt, dass es sich bei der Isolierung von Oozysten dennoch um eine vielversprechende Methode zur Aufreinigung von wirtsfreier Parasiten-DNA handelt. Im Vergleich zum Genom einer nah verwandten Plasmodienart, die auf eine andere Art und Weise um Wirt getrennt worden war, konnte hier eine deutlich bessere Genom-Qualität erzielt werden (Böhme et al., 2018).

In dieser Arbeit erwies sich die Aufreinigung des Parasiten aus den Blutproben der Wirtsfledermäuse als wesentlich praktikabler. Wie andere Haemosporiden lassen sich die Vertreter der Art *P. murinus* durch ein intrazelluläres Differenzierungsstadium in Erythrozyten der Wirtsfledermäuse charakterisieren, die keine eigenen Zellkerne enthalten. Zur Isolierung von Haemosporiden aus Blutproben sind daher bereits eine Reihe an Methoden hervorgegangen, die eine Trennung von infizierten Erythrozyten von anderen Zellen ermöglichen (z. B. Auburn et al., 2011; Tao et al., 2011; Chakarov et al., 2012). Eine Problematik, welche *Polychromophilus* für molekularbiologische Untersuchungen aller Art jedoch mit sich bringt, liegt in der geringen Menge an Parasiten-befallenen Erythrozyten. Im Gegensatz zu anderen Haemosporida weisen die Blutzellen des Zwischenwirtes nur Gametozyten auf (Gardner et al., 1988). Eine Parasitämie von unter 0,02 %, wie sie in dieser Arbeit festgestellt wurde, ist also keine Seltenheit (Witsenburg et al., 2014). Hinzu kommt, dass aus Wasserfledermäusen nur wenig Blut entnommen werden kann (hier: > 30 µl), sodass insgesamt nur eine sehr geringe Menge DNA aus den Proben gewonnen werden kann, die eine weitere Aufreinigung der Probe erschwert.

Fledermäuse der Gattung *Myotis* haben mit einer durchschnittlichen haploiden Genomgröße von ca. 2,1 Gb ein deutlich größeres Genom als das haploide eines Vertreters der Haemosporida (~23 Mb) (vgl. Seim et al., 2013; Zhang et al., 2013; Bensch et al., 2016; Böhme et al., 2018). Eine unvollständige Entfernung von Leukozyten aus den Ausgangsproben ist für die weitere Analyse daher ein großer Nachteil: Aufgrund der unterschiedlichen Genomgrößen führt die Anwesenheit eines einzigen Leukozyten im Verhältnis zu 100 Parasiten bei einer Sequenzierung zu mehr als 95 % Reads, die einen Ursprung im Genom des Wirts haben. In den meisten untersuchten Proben war eine Anreicherung der Parasiten über einen Ficoll-Gradienten um ein Vielfaches möglich. Eine statistisch signifikante Veränderung der Zusammensetzung des Anteils der genomischer DNA zu Gunsten des Parasiten wurde jedoch über *in vitro* Exflagellation erreicht (vgl. 4.1.3).

Die Exflagellation männlicher Gametozyten von Vertretern der Haemosporiden erfolgt in der Regel im Darm des Vektors und beschreibt die Entwicklung der Mikrogametozyten zu mehreren hochmotilen flagellenartigen Mikrogameten. Dabei handelt es sich um einen Prozess, der *in vitro* initiiert werden kann (z. B. Billker et al., 1997; Palinauskas et al., 2013). Die *in vitro* Exflagellation wurde inzwischen in einer Reihe an Experimenten mit ganz unterschiedlichen Fragestellungen, ausgehend von der Charakterisierung morphologischer Aspekte von Mikrogameten, Zygoten und Ookineten (z. B. Gardner et al., 1988; Coral et al., 2015; Chagas et al., 2019) bis hin Untersuchungen von Artbildungsprozessen, Hybridisierungen und intraspezifischen Variationen (Valkiūnas et al., 2008, 2013; Neal, 2011), erfolgreich eingesetzt. Die Etablierung der Methode eröffnet zugleich Möglichkeiten zur Analyse von unterschiedlichen genetischen Grundlagen, die für die sexuelle Fortpflanzung der Parasiten oder die Anpassung dieser an unterschiedliche Vektoren essenziell sind (Ngwa et al., 2013). Die Exflagellation ist abhängig von vielen Faktoren, wie einem steigenden

pH-Wert, einem Temperaturabfall oder der Anwesenheit eines Stoffwechselprodukts aus dem Tryptophanstoffwechsel des Vektors, der Xanthurensäure (Carter et al., 1977; Billker et al., 1997; Garcia et al., 1998). Anders als bei vielen Plasmodienarten lässt sich die *in vitro* Exflagellation bei Arten der Gattung *Haemoproteus* jedoch auch bereits durch eine Exposition von reifen Gametozyten an der Luft induzieren (Palinauskas et al., 2013). Das Potential, das dabei zur Trennung von Parasiten und Wirtszellen entsteht, wurde auch bei der Genomsequenzierung des Parasiten *Haemoproteus tartakovskyi* genutzt. Dabei konnte gezeigt werden, dass eine effiziente Isolierung des Parasiten möglich ist, sobald die Parasitämie einer Probe 2 % übersteigt (Palinauskas et al., 2013). Im Gegensatz zu der *in vitro* Exflagellation von *P. murinus* in dieser Arbeit erfolgte die Aufreinigung von *H. tartakovskyi* zudem unter kontrollierten Bedingungen im Labor. Dennoch war rund die Hälfte der Sequenzen auf das Genom des Wirts zurückzuführen (Bensch et al., 2016).

Obwohl die Genomkopienzahl von *P. murinus* nach der *in vitro* Exflagellation deutlich höher war als die des Wirts, enthielt die sequenzierte Probe retrospektiv auch hier neben diversen anderen Kontaminationen immer noch einen nicht unerheblichen Anteil an Wirts-DNA (vgl. 4.2.2). Tatsächlich wird die Standard-WGA oft mit unvollständiger Genomabdeckung und einer Überrepräsentation GC-reicher Sequenzen in Verbindung gebracht (Woyke et al., 2010; Oyola et al., 2014; Sabina et al., 2015). Erschwerend kommt hinzu, dass auch die Illumina-Technologie Schwierigkeiten bei der Sequenzierung von Fragmenten aus homogenen Basen zeigt (Videvall, 2018). Sowohl GC-reiche als auch GC-arme genomische Sequenzen sind in der Sequenzierung daher häufig unterrepräsentiert (Hansen et al., 2010; Benjamini et al., 2012).

5.1.2 Facetten eines Metagenoms: Ko-Infektionen und andere Kontaminationen

Die Rekonstruktion eines oder mehrerer Genome aus Sequenzen, die ihren Ursprung in der DNA unterschiedlicher Organismen haben, führt, ohne weiteren bioinformatischen Aufwand, nur selten zum Erfolg. Mit ihrem hohen Anteil genkodierender Sequenzen mit nur wenig repetitiver DNA stellen allein bakterielle Kontaminationen eine Herausforderung bei der Herstellung von Contigs dar (Kumar et al., 2013). Weisen die Organismen zudem auch noch unterschiedliche Genomgrößen mit unterschiedlichem GC-Gehalt auf, entsteht bei der Sequenzierung eine ungleichmäßige Abdeckung, die wiederum die Gesamtqualität eines Assembly herabsetzt (Kumar et al., 2013). Die Folge ist die Entstehung von falsch rekonstruierten Contigs, Misassemblies, die im schlimmsten Fall artifiziell aus den Genomen unterschiedlicher Organismen zusammengesetzt sind. Beispiele für Genome, die Contigs aus Kontaminationen unterschiedlichster Art enthalten, gibt es in der Literatur viele (z. B. Laurence et al., 2014; Salter et al., 2014; Borner et al., 2017). Dazu zählen etwa Genome, die bei der Sequenzierung mit menschlicher oder bakterieller DNA kontaminiert waren oder aber auch erst bei der Sequenzierung durch das Vorhandensein von Restfragmenten aus früheren Sequenzierungsläufen mit DNA anderer Organismen vermischt wurden (Merchant et al., 2014; Kryukov et al., 2016; Souvorov et al., 2018).

Im schlimmsten Fall ergeben sich aus diesen falsche Schlussfolgerungen diverser biologischer Fragestellungen. Eines der bekanntesten Beispiele hierfür stellt das Genom eines Tardigraden dar,

dessen bakterielle Kontaminationen als Resultat eines massiven horizontalen Gentransfers interpretiert wurden (Boothby et al., 2015; Arakawa, 2016). Inzwischen sind eine Reihe unterschiedlicher Tools publiziert worden, die den Umgang mit Kontaminationen in Rohdaten erleichtern (z. B. Schmieder et al., 2011; Kumar et al., 2013; Pandey et al., 2016; Fiévet et al., 2019). Die meisten dieser sind jedoch auf die Identifizierung von Sequenzen bakteriellen Ursprungs ausgelegt. Genomsequenzierungsprojekte von eukaryotischen Parasiten, die, wie in dieser Arbeit, auf die Qualität von Referenzgenomen zur Identifizierung von eukaryotischen Kontaminationen angewiesen sind, sind daher meist benachteiligt

Auch in dieser Arbeit wurde mit *Myotis davidii* ein Referenzgenom verwendet, das Kontaminationen eines eukaryotischen Parasiten aus dem Phylum Apikomplexa enthält (Orosz, 2015; Borner et al., 2017). In den ersten Mappings waren noch rund 66 Contigs enthalten, die etwa 31 kbp des Genoms einer Art aus der Gattung *Hammondia* repräsentierten. Obwohl *Myotis davidii* auch deutlich näher mit der Wirtsfledermaus von *P. murinus* verwandt ist als die anderen Referenzfledermäuse, konnte aber kein signifikant höherer Anteil an Reads festgestellt werden, die nach dem Mapping als Reads mit putativem Ursprung in der DNA der Fledermaus identifiziert wurden (nicht dargestellt). Auch nach der Entfernung der putativen Kontamination veränderte sich das Mapping nicht wesentlich. Zurückzuführen ist dies einerseits auf den geringen Anteil an Reads, die in der ursprünglichen Sequenzierung aus der DNA von *P. murinus* resultierten, und andererseits in der genetischen Distanz zwischen den Parasiten (vgl. Parfrey et al., 2011).

Die Ergebnisse zeigen, dass Kontaminationen nicht immer ein Hindernis in Sequenzierungsprojekten sein müssen. Viel mehr kristallisiert sich immer stärker heraus, dass diese in NGS-Daten eine Ressource für viele biologische Fragestellungen sein können (z. B. Borner et al., 2017; Laine et al., 2019; Sangiovanni et al., 2019). Auch diese Arbeit konnte von den Kontaminationen eines anderen Sequenzierungsprojekts profitieren. Im Genom des Uganda-Stummelaffens (*Ptilocolobus tephrosceles*; ASM277652v1) konnten mit dem ContamFinder Contigs identifiziert werden, die aus einer Infektion mit einem Vertreter der Art *Hepaticocystis* resultierten (Borner et al., 2017, zur Verfügung gestellt von Dr. Janus Borner, Universität Hamburg). Mit diesen konnte das Apikoplasmid fast vollständig rekonstruiert werden (vgl. 4.5.2). Neben einigen wenigen Genfragmenten und drei mitochondrialen Genomen, die auf den Seiten des NCBI öffentlich verfügbar sind, stellte dies bis Dezember 2019 die größte Sequenz eines Parasiten der Gattung *Hepaticocystis* dar. Das Potential einer zielgerichteten Sequenzierung von Multi-Genomen und Multi-Transkriptomen aus Blutproben ist sogar noch viel größer: neben der Charakterisierung einzelner pathogener Arten ermöglichen NGS-Methoden Untersuchungen von Parasit-Wirt-Interaktionen oder Analysen von Infragemeinschaften von Parasiten (Pauli et al., 2015; Videvall et al., 2015; Santos et al., 2018; Toscani Field et al., 2018; Galen et al., 2019).

Obwohl Fledermäuse wegen ihrer Diversität an Viren, die sie beherbergen können, Gegenstand aktueller Forschung sind, ist über ihre Parasitenfauna wenig bekannt (Hayman, 2016; Clarke-Crespo et al., 2017). Mit denen als Kontaminationen bewerteten Contigs, der sequenzierten Genome, gibt diese Arbeit einen Einblick in die Diversität der Infragemeinschaft im Blut der Fledermäuse. Tatsächlich sind in der mit *P. murinus* infizierten (WF43) und nicht infizierten Fledermaus (WF32) eine große Anzahl an Parasiten identifiziert worden. Vor allem aber konnte in

der Wasserfledermaus WF43 ein breites Spektrum an unterschiedlichen Organismen nachgewiesen werden. Hier waren insbesondere Sequenzen von Vertretern der Bacteroidetes, Proteobacteria, Chlorophyta, Arthropoda und Nematoda überpräsentiert (vgl. 4.2.2). Die Mehrheit aller Pathogene hatte einen bakteriellen Ursprung. Neben den Bacteroidetes und Proteobacteria konnten diese in elf weitere Stämme eingeordnet werden. Da die Charakterisierung der Parasitenfauna nicht Gegenstand dieser Arbeit war, wurden die einzelnen Spezies nicht näher identifiziert. Interessant waren jedoch Vertreter der Gattung *Mycoplasma*, die intrazellulär in Erythrozyten leben und in Säugetieren wahrscheinlich auch über haematophage Vektoren übertragen werden (Mascarelli et al., 2016). Eine signifikante Korrelation zwischen Infektionen mit *Polychromophilus* und *Mycoplasma* scheint es jedoch nicht zu geben, sodass eine Rolle der Fledermausfliege bei der Übertragung fraglich erscheint (Holz et al., 2019). Anders sieht es dagegen bei der Übertragung von *Bartonella* Spezies aus, die ebenfalls in dem Genom der mit *P. murinus*-infizierten Fledermaus identifiziert werden konnten. Hier wurde bereits beschrieben, dass Fledermausfliegen eine Vielzahl an diesen enthalten und wahrscheinlich eine tragende Rolle als Reservoir spielen (Wilkinson et al., 2016). Die meisten Vertreter dieser Gattung scheinen in ihrem Wirtsspektrum nur auf eine Säugetierart eingeschränkt zu sein (Chomel et al., 2009). Allerdings wurde inzwischen auch ein fledermaus-assoziiertes zoonotisches Erreger beschrieben, der mit einer Endokarditis im Menschen in Verbindung gebracht wird (Wilkinson et al., 2016). Auch die Identifizierung von Contigs, die auf die Gattung *Pseudogymnoascus* zurückzuführen waren, waren im Assembly möglich. Der Pilz, der mit dem Massensterben von Fledermäusen durch das Weißnasen-Syndrom in Nord-Amerika in Verbindung steht, kann auch für europäische Fledermäuse pathogen sein, ist hier jedoch nicht mit einer erhöhten Sterblichkeit assoziiert (Bandouchova et al., 2015). Die exemplarisch dargestellten Erreger waren mit Ausnahme einiger vereinzelter Contigs, die *Mycoplasma* zugeordnet werden konnten, nicht in der Wasserfledermaus WF32 zu finden. Im Vergleich zur Wasserfledermaus Mg43 war die Diversität der festgestellten Kontaminationen auch deutlich geringer (vgl. 4.3.1). Die enorme Diversität, in der mit *P. murinus* infizierte Fledermaus, ist auch ein Resultat aus Umweltkontaminationen (z. B. „Chlorophyta-Kontaminationen“, vgl. 4.2.2), die methodisch im Feld nicht zu vermeiden waren. Dennoch spiegelt die hohe Anzahl an identifizierten endogenen Erregern jedoch das Bild einer immunsupprimierten Fledermaus wider.

In heterogenen Parasitengemeinschaften wird angenommen, dass Interaktionen zwischen unterschiedlichen Organismen, wie z. B. Konkurrenz um Ressourcen oder eine gegenseitige Manipulation, ein Hauptfaktor in der Evolution der einzelnen Parasiten darstellt (Alizon et al., 2013; Bashey, 2015). Inwiefern die Präsenz von unterschiedlichen Parasiten einen Einfluss auf die Entwicklung der besonders wirtsspezifischen Vertreter der Gattung *Polychromophilus* gehabt haben könnte, kann an dieser Stelle nur spekuliert werden.

5.1.3 Die zielgerichtete Anreicherung des Parasiten-Genoms

Zielführend zur Sequenzierung des Genoms von *P. murinus* war eine mehrphasige Strategie basierend auf einer bioinformatischen Isolierung der Sequenzinformationen des Parasiten und einer anschließenden Resequenzierung. In Multi-Genom- und Multi-Transkriptom-Analysen erfolgt die Trennung der Sequenzen unterschiedlicher Organismen in der Regel *in silico* über einen

Abgleich mit entsprechenden Referenzgenomen in Form eines Mappings (z. B. Humphrys et al., 2013; Videvall et al., 2015; Marsh et al., 2017; Mu et al., 2019). Eine solche Herangehensweise hat sich in dieser Arbeit als wenig erfolgsversprechend herausgestellt (vgl. 4.2.3). Obwohl bei der Positionierung der Reads an Referenzgenome anderer Apikomplexa die meisten unerwünschten Sequenzen verworfen werden konnten, wurden auch Bereiche von *P. murinus*, die eine geringe Identität zu den Referenzgenomen aufweisen, aus dem Datensatz entfernt. Ein Mapping gegen die Genome nah verwandter Fledermausarten des Wirtes entfernt dagegen nur einen kleinen Teil aller kontaminierenden Sequenzen. Dabei bleiben zwar mehr Reads aus *P. murinus* erhalten, allerdings standen diese einem immer noch größeren Anteil an Reads aus der Wirtsfledermaus und anderen Organismen gegenüber. Die Ergebnisse spiegeln ein häufiges Problem bei der Analyse eines Organismus wider, das nicht zum Modell-Organismus avanciert ist: der Umfang eines Mappings wird davon beeinflusst, wie sehr Reads und Referenz in ihrer Sequenz übereinstimmen. Bereits bei unterschiedlichen Allelen, die nur zu einem gewissen Grad an die Referenzregion kartiert werden können, kann sich eine Verzerrung der Ergebnisse einschleichen (Ballouz et al., 2019; Günther et al., 2019). Mit einer Zunahme der genetischen Distanz wird ein Mapping zudem immer unzuverlässiger, sodass Reads, die auf diese Weise extrahiert werden, nur bedingt die Zielsequenzen abbilden (Gouin et al., 2015; Price et al., 2017).

Die Alternative die Reads mit BLAST-Algorithmen abzugleichen, ist dagegen wesentlich toleranter gegenüber abweichenden Sequenzen und viel sensitiver bei einer Zuordnung von Sequenzen (Buchfink et al., 2015; Suzuki et al., 2017; Boratyn et al., 2019). Gleichzeitig geht dies jedoch mit einem enormen Zeit- und Rechenaufwand einher. Die Sequenzen wurden daher in mehreren Schritten sortiert: Auf DNA-Ebene gegen das Genom der Wirtsfledermaus und auf Proteinebene gegen einen Datensatz, der repräsentative Proteine von Säugetieren, Pilzen und Bakterien enthielt. Insgesamt 1,3 Mb aus dem Genom von *P. murinus* konnten so rekonstruiert werden (vgl. 3.14 und 4.4.4). Die Ergebnisse dieses Prozesses zeigen, dass es mit der richtigen Strategie zur Aufbereitung von NGS-Daten möglich ist, qualitativ hochwertige Contigs eines intrazellulären Parasiten zu erhalten, die frei von kontaminierenden Sequenzen, wie die des Wirtes, sind und viel Potential für weitere Herangehensweisen liefern. In dieser Arbeit dienten die Sequenzen als Basis für eine weitere Anreicherung der genomischen DNA von *P. murinus*.

Die Idee spezifische DNA-Sequenzen mithilfe von bekannten Fragmenten zu vervielfältigen ist nicht neu: seit ihrer Einführung hat sich die PCR als unersetzliche Standardmethode zur Anreicherung von bestimmten genomischen Regionen entwickelt. Auch in Kombination mit NGS-Methoden findet die PCR, meist in Form der Long-Range-PCR, gelegentlich Anwendung (z. B. Melum et al., 2010; Ozcelik et al., 2012; Deiner et al., 2017). Eine Anreicherung von genomischen Bereichen über PCR, die beispielsweise mit bestimmten Krankheiten in Verbindung stehen, kann eine kostengünstige Alternative sein, wenn es sich um wenige Zielregionen handelt, die sich auf bis zu 30 kb erstrecken (Jia et al., 2015). Mit zunehmender Anzahl und Größe der Regionen von Interesse nimmt die Anwendbarkeit der PCR aufgrund zunehmenden Zeit- und Kostenaufwands jedoch stark ab.

Andere Methoden, die für groß angelegte Projekte im Mb-Bereich besser geeignet sind, werden unter Sequence Capture bzw. Targeted Enrichment zusammengefasst. Grob unterscheidet man

hier hybridisierungs-basierte Methoden, die Zielregionen durch Biotin markieren, immobilisieren oder enzymatisch zirkularisieren (Mamanova et al., 2010). Als Referenz dienen dabei DNA- oder RNA-Proben bereits sequenzierter Genome, die auch eine gewisse evolutive Distanz zum Untersuchungsobjekt haben können (Lemmon et al., 2012). Ausgehend von dem *H. tartakovskyi* Genom wurden beispielsweise bereits 100-300 Loci in anderen vogelinfizierenden Haemosporida sequenziert. Darunter fallen diverse Vertreter der Gattungen *Leucocytozoon*, *Haemoproteus* und *Plasmodium*, die sich auf mtDNA-Ebene um bis zu 16 % unterscheiden (Huang et al., 2018; Barrow et al., 2019). Im Vergleich weisen die mitochondrialen Genome der Plasmodien *P. gallinaceum*, *P. falciparum*, *P. vivax* und *P. yoelii*, aber auch die hier sequenzierte mtDNA von *P. murinus*, einen Konservierungsgrad von etwa 90 % auf (McIntosh, 1998). Ob beispielsweise ein Targeting Enrichment basierend auf Fragmenten von *P. falciparum* bei Vertretern der Gattung *Polychromophilus* möglich ist, müsste jedoch ausgetestet werden. Die hier sequenzierten genomischen Bereiche lassen zumindest aufgrund ihrer Sequenzähnlichkeit zu den bislang sequenzierten Genomen der Vertreter der Gattung *Plasmodium*, den Schluss zu, dass viele Loci gut konserviert sind. Ein Nachteil dieser Methode ist jedoch, dass sie nur bedingt für die Charakterisierung neuer Genome geeignet ist (Gaudin et al., 2018). Dabei beruht das Targeting Enrichment auf der Anreicherung hoch konservierter Bereiche, auch über mehrere Taxa hinweg (z. B. Crawford et al., 2012; Lemmon et al., 2012). Art- oder gattungsspezifische Besonderheiten, die bislang noch nicht sequenziert worden sind und nicht in der Referenz vorhanden sind, können daher nicht erfasst werden.

Eine alternative bietet hierzu die, auch in dieser Arbeit, angewendete SWGA. Dabei handelt es sich um eine Methode, bei der, ähnlich zu der PCR, ein spezifischer Anteil einer DNA-Probe amplifiziert wird. Im Gegensatz zur herkömmlichen PCR werden dabei jedoch Primer verwendet, die besonders häufig an komplementäre Sequenzen im Zielgenom binden (Leichty et al., 2014). Im Idealfall kommen die ausgewählten Sequenzen nicht oder nur ganz selten in der kontaminierenden DNA vor (Clarke et al., 2017). In den letzten Jahren hat sich die SWGA gerade in der Diagnostik von *P. falciparum*, *P. knowlesi* und *P. vivax* als nützliches Tool erwiesen (Oyola et al., 2016; Cowell et al., 2017; Benavente et al., 2019). Die ersten sequenzierten Genome, bei der die SWGA angewendet worden ist, waren mit *P. gaboni*, *P. reichenowi* und *P. blacklocki* Vertreter der Gattung *Plasmodium* (Sundararaman et al., 2016; Otto et al., 2018). Inzwischen sind jedoch auch eine Reihe weiterer Genomprojekte veröffentlicht worden, bei denen Pathogene, darunter ein Bakterium, Viren und Nematoden, im Vorfeld einer Sequenzierung mit der SWGA angereichert worden sind (Abbas et al., 2019; Small et al., 2019; Cocking et al., 2020).

Mit Ausnahme des Bakteriums *Coxiella burnetii* (43-45 % GC) handelt es sich dabei um Parasiten, die sich durch einen niedrigen GC-Gehalt im Genom charakterisieren lassen. Auch in dieser Arbeit ist der GC-Gehalt des Parasiten (18,6 %) im Vergleich zum Wirtsgenom der Wasserfledermaus (40,4 % GC) deutlich niedriger. Daher sind bei der SWGA von *P. murinus* Primer verwendet worden, die besonders GC-arm sind (vgl. Tab. 2.4). Wie auch bereits in anderen Arbeiten beschrieben, war nach der Sequenzierung der angereicherten DNA eine deutlich ungleichmäßige Abdeckung zu erkennen (vgl. 4.4.1). Dies kann auf eine ungleichmäßige Verteilung der Primerbindungsstellen im Genom zurückzuführen sein, die zu einer sehr starken Abdeckung einiger Bereiche führt, während andere im abschließenden Genom gar nicht oder nur kaum

repräsentiert sind. Gerade in Kombination mit der Illumina-Sequenzierung, die bekanntermaßen homogene gegenüber solchen mit niedrigem GC-Gehalt bevorzugt, kann eine SWGA zu vielen schlecht abgedeckten Bereichen im Genom führen (Hansen et al., 2010; Benjamini et al., 2012). Wie auch bei dem Genom von *P. murinus* scheinen dabei vor allem Bereiche der Sub-Telomere betroffen zu sein (Oyola et al., 2016; Sundararaman et al., 2016; Otto et al., 2018). Eine methodische Alternative für die Sequenzierung GC-armer Genome stellt ein Hybridansatz basierend aus Illumina- und PacBio-Sequenzierung dar. Diese Herangehensweise war auch bei der Sequenzierung des Pilzgenoms von *Pecoramyces ruminantium* erfolgreich, der den bislang niedrigsten GC-Gehalt (17 %) unter allen sequenzierten Eukaryoten aufweist (Youssef et al., 2013; Videvall, 2018).

5.2 Evolution charakteristischer Merkmale im Genom der Haemosporiden

5.2.1 Genomreduktion und GC-Gehalt als Zeichen der Evolution

Zwischen verschiedenen Eukaryoten bestehen teilweise erhebliche Genomgrößenunterschiede. Sowohl eines der größten (*Polychaos dubium*, 670 Gbp) als auch das eines der kleinsten Genome (*Encephalitozoon intestinalis*, 2,3 Mbp) können unter den Vertretern der Protozoen gefunden werden (Gregory et al., 1999; Corradi et al., 2010). In der Regel besteht kaum ein Zusammenhang zwischen der Genomgröße und der morphologischen Komplexität einer Art (Markov et al., 2010). Parasitische Protozoen weisen mit etwa 10-80 Mbp jedoch immer relativ kleine Genome auf (Worthe et al., 2005). Mit einer durchschnittlichen Größe von etwa 23 Mbp finden sich die Genome der Haemosporida im unteren Bereich (z. B. Carlton et al., 2002; Gardner et al., 2002; Bensch et al., 2016; Böhme et al., 2018). Im Vergleich zu ihren nächsten freilebenden Verwandten lassen sich Haemosporiden durch einen massiven Verlust an Faktoren charakterisieren, die für freilebende Organismen essenziell sind (Woo et al., 2015). Dazu gehören beispielsweise viele Stoffwechselwege, wie die Photosynthese und die Fähigkeit zur Sterin-Synthese, oder Bestandteile, die für die Ausbildung von Flagellen benötigt werden (Woo et al., 2015). Dieser Verlust spiegelt sich auch in der Anzahl der kodierenden Gene wider. Während die freilebende Art *Chromera velia* ein Genom der Größe 193,4 Mbp mit 26.112 proteinkodierenden Genen aufweist, haben Haemosporiden im Schnitt zwischen 5.000 und 6.000 proteinkodierenden Genen (z. B. Carlton et al., 2002; Gardner et al., 2002; Woo et al., 2015; Bensch et al., 2016; Böhme et al., 2018). Die Evolution von Parasiten ist dabei im Allgemeinen mit einer Vereinfachung der Organismen verknüpft. Insbesondere bei Bakterien ist ein solches Phänomen gut untersucht. Hier kann der Übergang von freilebenden Bakterien zu solchen, die obligat intrazellulär leben, mit einem Verlust von bis 85 % des ursprünglichen Genrepertoires einhergehen (Moran, 2002). Neben vielen Genen, die in nicht mehr benötigten Stoffwechselwegen oder spezifischen Zellfunktionen beteiligt sind, kann vor allem auch der Verlust redundanter oder duplizierter Gene auftreten. Dabei wirkt sich der Selektionsdruck zum kleineren Genom nicht nur auf das Kerngenom, sondern auch auf extrachromosomale Sequenzen aus. Besonders deutlich wird dies auch durch den in dieser Arbeit

sequenzierten Apikoplasten (vgl. 4.4.3). Die in *P. murinus* nachgewiesene Deletion eines tRNA/rRNA Gen-Clusters, die zugleich mit einer Entstehung von zwei unterschiedlichen Isoformen des Apikoplasten einhergeht, stellt eine konvergente Entwicklung dar, die auch im nagetierinfizierenden *P. chabaudi* bereits beschrieben worden ist (Sato et al., 2013).

Im Kerngenom sind die enormen Genomgrößenunterschiede zwischen freilebenden und parasitischen Organismen auch auf das Vorhandensein von transponierbaren Elementen zurückzuführen. Etwa 30 Mbp sind im nächsten freilebenden Verwandten, *Chromera velia*, auf transponierbare Elemente zurückzuführen (Woo et al., 2015). In der Gattung *Plasmodium* wurde dagegen lange davon ausgegangen, dass transponierbare Elemente nicht vorhanden sind (Carlton et al., 2002; Gardner et al., 2002). Inzwischen konnten jedoch vereinzelte Elemente nachgewiesen werden (Böhme et al., 2018).

Zugleich wurde die Hypothese aufgestellt, dass insbesondere AT-reiche Genome eine gewisse Resistenz gegenüber der Invasion transponierbarer Elemente aufweisen (Durand et al., 2006). Ein Mechanismus, der in erster Linie als Abwehr gegen transponierbare Elemente angesehen wird und mit einer Anreicherung AT-reicher Regionen in Zusammenhang steht, ist die in Pilzen konservierte repeat-induzierte Punktmutation (Testa et al., 2016). Mechanismus in Haemosporiden nachgewiesen werden konnte, ist ein solcher auf Grund der in Plasmodien nachgewiesenen Überrepräsentation von G:C→A:T Transitionen zumindest vorstellbar (Hamilton et al., 2016). Die meisten bislang sequenzierten Vertreter des Taxons Haemosporida weisen einen beeindruckend niedrigen GC-Gehalt auf (Bensch et al., 2016; Videvall, 2018). Auch *P. murinus* hat mit 18,6 % einen besonders niedrigen GC-Gehalt, der sogar den von *P. falciparum* übertrifft (19,3% GC) (vgl. 4.4.4). Intra- und intergenetische Bereiche sind mit 13,3 % bzw. 13,2 % besonders GC-arm. Zustände kommt dies durch einfache Repeats niedriger Komplexität, die stellenweise sogar einen noch weitaus niedrigeren GC-Gehalt (< 10 %) aufweisen können.

Trotz der Limitationen hinsichtlich der Kodonnutzung sind auch die Gene bei *P. murinus* besonders GC-arm (78,6 %). Vergleichbar mit anderen Vertretern der Haemosporida sind dadurch häufig homopolymerartige Regionen in Proteinen assoziiert (Singh et al., 2004; Russell et al., 2014). Diese treten häufig in Kombination mit konventionellen Protein-Protein-, Protein-Lipid- oder Protein-DNA-Interaktionsdomänen auf (Pelassa et al., 2015). Insbesondere Proteine, die aus polaren (Ser, Gln) oder geladenen (Asp, Glu, Lys, Arg) Regionen bestehen zeigen eine größere Tendenz untereinander zu interagieren (Pelassa et al., 2015). Neben homopolymerartigen Aminosäuresequenzen aus Asn kommen besonders häufig Poly-Lys und Poly-Asp in den Proteinen von *P. murinus* vor. Während die Funktion Asn-reicher Regionen noch nicht aufgeklärt ist, wurde für exportierte Proteine mit Lys-reichen Sequenzen eine periphere Lokalisation in der Wirtszelle vorausgesagt (Davies et al., 2016). Die Abundanz dieser homopolymerartigen ist natürlicherweise mit dem niedrigen GC-Gehalt der Haemosporida assoziiert. Das Genom von *P. vivax* beispielsweise, das einen eher ausgeglichenen GC-Gehalt besitzt, weist deutlich weniger Lysin-reiche Regionen im Vergleich zu *P. falciparum* auf (Davies et al., 2016). Der niedrige GC-Gehalt und damit auch die Tendenz zu vermehrter Homopolymer-Integration in Proteinen ist ein ursprüngliches Charakteristikum in den Haemosporida. Der vogel-infizierende *H. tartakowskyi* weist mit einem GC-Gehalt von 23,1 %, ebenfalls eine starke Verschiebung in der Kodonverwendung

auf (Bensch et al., 2016). Auch hier sind Asn-, Lys- und Ser-reiche Sequenzen häufig vertreten (nicht dargestellt). Die hier untersuchten Vertreter der Piroplasmida *Babesia bovis*, *B. microti* und *Theileria parvum* weisen dagegen eine gleichmäßigere Verteilung zwischen GC und AT im Genom auf (Brayton et al., 2007).

5.2.2 Das konservierte Genrepertoire der Haemosporiden

Der Vergleich von Genomen unterschiedlicher Vertreter der Apikomplexa zeigte teilweise erhebliche Unterschiede. Syntänie Regionen, die auf einer konservierten Anordnung homologer Gene beruhen, scheinen dabei zwischen den unterschiedlichen Taxa vollständig zu fehlen und sind lediglich zwischen nahe verwandten Arten anzutreffen (DeBarry et al., 2011). Eine Ausnahme bilden hierzu die Vertreter der Aconoidasida. Hier können etwa 1.300 orthologe Gene identifiziert werden, die über 435 Syntänie-Blöcke verteilt sind (Pain, 2005; Swapna et al., 2017). Bei einem Großteil dieser Proteine handelt es sich um Eins-zu-Eins Orthologe. Dabei konnte bereits in einer anderen Arbeit gezeigt werden, dass viele dieser Gene, zumindest im Organismus *P. chabaudi*, essenziell am Wachstum des Blutstadiums beteiligt sind (Bushell et al., 2017).

Auch hier konnte gezeigt werden, dass es sich hauptsächlich um Gene handelt, die in grundlegende Zellprozesse involviert sind (vgl. 4.5.2). Einerseits konnten vor allem Gene, die an konstitutiven Stoffwechselprozessen, wie der Kohlenhydrat- oder Fettsäureumwandlung beteiligt sind, identifiziert werden. Aber auch solche, die in der Zellzyklus-Regulation, der Translation oder Transkription mitwirken, konnten als Bestandteile des konservierten Kerngenoms zwischen Piroplasmiden und Haemosporiden nachgewiesen werden. Darunter finden sich 15 Proteine mit der DNA-bindenden Domäne AP2. Ursprünglich wurden AP2-Domänen, manchmal auch ERF-Domänen (ethylen response factor), in Pflanzen beschrieben (Riechmann et al., 1998; Balaji, 2005). Proteine mit mindestens einer dieser Domänen fungieren dabei als Transkriptionsfaktoren, die an der Expression stressabhängiger Gene aber auch bei der Kontrolle diverser zellulärer Prozesse während unterschiedlicher Entwicklungsstadien beteiligt sind (Jofuku et al., 1994). Auch in *Plasmodium* spp. konnte gezeigt werden, dass Proteine der ApiAp2-Familie über alle Lebensstadien hinweg eine besondere Rolle bei der Differenzierung der Parasiten spielen (Balaji, 2005; Yuda et al., 2009, 2010; Iwanaga et al., 2012; Kafsack et al., 2014; Sinha et al., 2014).

Zugleich spiegeln auch einige andere konservierte Gene die frühe Adaptation an den weitgehend ähnlichen Lebensstil der beiden Taxa wider. Dabei teilen sich Haemosporiden und Piroplasmiden einige morphologische und entwicklungsgeschichtliche Merkmale, wie z. B. apikale komplexe Organellen, die Entwicklung innerhalb der Blutzellen von Wirbeltierwirten und die sexuelle Reproduktion in wirbellosen Vektoren (Homer et al., 2000; Jalovecka et al., 2018). Konserviert sind dabei insbesondere Zytoskelett-Proteine, die einerseits die äußere Form stabilisieren, andererseits aber auch die gleitende Beweglichkeit der Parasiten bewirken (Baum et al., 2006b; Soldati-Favre, 2008; Woo et al., 2015). Auch Bestandteile der Invasionsmaschinerie sind konserviert. Die molekularen Prozesse während der Invasion sind besonders gut in den Arten *P. falciparum* oder *T. gondii* charakterisiert (Besteiro et al., 2009; Frénil et al., 2017). Dabei bilden Parasiten eine bewegliche Verbindung zu der Wirtszelle (moving junction) auf, die sich vom apikalen zum

posterioren Ende des Parasiten bewegt. Ausgebildet wird diese durch Proteine (RON2 und RON4), die aus den Rhoptrien in die Wirtszelle injiziert werden, die Wirtsmembran überspannen und als Ankerpunkt für das Oberflächenprotein AMA-1 dienen (Alexander et al., 2005; Cao et al., 2009; Vulliez-Le Normand et al., 2012).

Die Mehrheit der Gene, die Invasions- oder Zelltransmigrationsprozessen zugeordnet werden können und innerhalb des Taxons der Haemosporida konserviert sind, fehlen jedoch in den hier untersuchten Vertretern der Piroplasmiden (vgl. 4.5.2 und 4.5.5). Im Vergleich zu den Haemosporiden weisen Piroplasmiden mit etwa 8 Mbp ein deutlich kompakteres Genom auf (Martínez-Ocampo, 2018). Dies legt die Vermutung nahe, dass ein Großteil dieser Gene im ursprünglichen Genom der Aconoidasida vorhanden war und anschließend in den Piroplasmiden verloren gegangen sind. Dafür spricht, dass der unabhängige Verlust vieler Gene in unterschiedlichen Abstammungslinien der Apikomplexa tatsächlich sogar die stärkste evolutive Kraft zu sein scheint, die auf die Genome der einzelnen Vertreter wirkt (Gardner et al., 2002; Keeling, 2004; Slamovits et al., 2004; Kissinger et al., 2011). Andererseits weisen viele haemosporida-spezifischen Gene, keine Orthologe in anderen Apikomplexa auf, was darauf hindeutet, dass sie entweder *de novo* entstanden sind oder sich bis zur Unkenntlichkeit weiterentwickelt haben (DeBarry et al., 2011; Kissinger et al., 2011). Viele konstitutive Proteine der Apikomplexa, darunter z. B. auch ribosomale Proteine, unterscheiden sich auch so sehr von ihren „homologen“ Proteinen anderer Eukaryoten, dass keine Sequenzähnlichkeit nachgewiesen werden kann (Wasmuth et al., 2009). Erklärungen für neu entstandene beruhen meist auf Duplikations- und Divergenzmechanismen, aber auch *de novo* Gene basierend auf nicht-kodierenden Bereichen sind vorgeschlagen worden (Long et al., 2003; Tautz et al., 2011; Magadum et al., 2013; Rödelberger et al., 2019). Auf diese Weise können einzelne Genome entstehen, die bis zu einem Drittel auf Genen basieren, die keine erkennbare Homologie zu anderen Taxa aufweisen (Rödelberger et al., 2019). Auch unter den Haemosporiden findet sich nur ein Anteil von etwa 60 % an homologen Genen (Carlton et al., 2001; Bensch et al., 2016). Darunter ist ein erheblicher Anteil an Genen, die bislang funktionell nicht weiter eingeteilt werden können, da sie keine konservierten Domänen und auch sonst keine Sequenzähnlichkeiten zu Proteinen anderer Arten aufweisen (vgl. 4.5.2).

Nichtsdestotrotz spielen diese Gene eine wichtige, manchmal sogar essenzielle, Rolle in unterschiedlichen Lebensstadien der Parasiten (Bushell et al., 2017). Viele dieser Proteine werden zudem eine besondere Rolle in der Parasit-Wirt-Interaktion spielen, die in unterschiedlichen Taxa konserviert sind. Eine Analyse der konservierten Proteine über unterschiedliche hierarchische Rangstufen hinweg, wie in dieser Arbeit (vgl. 4.5), ermöglicht daher auch die Identifizierung neuer Kandidatenproteine, die eine wesentliche Rolle bei Parasit-Wirt-Interaktionen spielen. Gerade in Hinblick auf biologische Gemeinsamkeiten und Unterschiede der betrachteten Rangordnungen oder Taxa können für viele Proteine Hypothesen für eine funktionelle Einordnung abgeleitet werden. Dazu zählen etwa Proteine, die in allen Vertretern der Aconoidasida konserviert sind, jedoch außerhalb dieses Taxons nicht gefunden werden können. Diese könnten beispielsweise interessante Kandidaten für die intrazelluläre Lebensweise in Blutzellen und damit auch für die Entwicklung der Gametozyten darstellen.

5.2.3 Proteine mit Relevanz für die Parasit-Wirt-Interaktion

Die überwiegende Mehrheit spezifischer Proteine in Vertretern der Plasmodiidae repräsentiert art- oder linien-spezifische Innovationen (Wasmuth et al., 2009). Dabei handelt es sich vermehrt um Proteinfamilien, die in Parasiten-Wirt-Interaktionen involviert sind und relativ junge Adaptationen an den jeweiligen Wirt darstellen (Wasmuth et al., 2009; Reid, 2015). Die entsprechenden Gene sind tandemartig in Regionen kodiert, die an Telomere oder auch an Zentromere angrenzen (Gardner et al., 2002; Kissinger et al., 2011). Leider konnten genau diese Bereiche, in denen vermutlich auch die Mehrheit der artspezifischen Gene von *P. murinus* enthalten sind, in dieser Arbeit nicht sequenziert werden. Wie bereits dargestellt enthalten allerdings auch die gut konservierten Bereiche über alle Haemosporiden hinweg ein gewisses Grundrepertoire an Genen, das für eine erfolgreiche Invasion und Entwicklung der Parasiten essenziell ist.

In ihren Wirbeltierwirten überleben die Parasiten ausschließlich intrazellulär. Daher ist die Invasion für die Vertreter der Plasmodiidae ein besonders kritischer Schritt. Grundlage der Invasion ist die spezifische Erkennung der geeigneten Wirtszelle (Iyer et al., 2007). Hier unterscheiden sich einerseits die einzelnen Arten voneinander aber auch die unterschiedlichen Stadien. Während bei *Polychromophilus* und auch bei vogelinfizierenden Plasmodienarten diverse Zellen des retikuloendothelialen Systems als Wirtszellen in Frage kommen, findet die exoerythrozytäre Schizogonie säugetierinfizierender Plasmodien ausschließlich in Hepatozyten statt (Manwell, 1946; Garnham, 1973b; Landau, Rosin, et al., 1980; Soulard et al., 2015; Valkiūnas et al., 2017). Erythrozyten stellen dagegen für alle Vertreter der Plasmodiidae essenzielle Wirtszellen dar — auch wenn diese in den fledermausinfizierenden Gattungen ausschließlich für die Entwicklung der Gametozyten genutzt werden (Garnham, 1948, 1951, 1973a; Garnham et al., 1953). Dennoch sind viele Proteine, die an den Invasionsprozessen beteiligt und in den Rhoptrien oder Mikronemen lokalisiert sind, konserviert (vgl. 4.5.2). Darunter finden sich auch einige, die, zumindest in *P. falciparum*, in der Migration der Sporozoit zur Leber und in der Hepatozyteninvasion beteiligt sind, wie CSP, TRAP, CTRP, CelTOS, SPECT1 und GEST (Garcia et al., 2006).

Die Invasion der Apikomplexa wird als ein mehrschrittiger Prozess verstanden, bei dem eine erste reversible Interaktion über Proteine der Mikronemen zu unspezifischen Bindungspartnern der Wirtszelle erfolgt (Cowman et al., 2006; Masamichi Aikawa et al., 2008). In dieser Arbeit konnte gezeigt werden, dass für viele der konservierten Proteine eine Interaktion zu unterschiedlichen Typ II Keratinen vorhergesagt werden kann (vgl. 4.5.5). Keratine stellen die größte Gruppe der intermediären Filamentproteine dar und sind zugleich die am häufigsten vorkommenden Proteine von Epithelzellen (Jacob et al., 2018). Dabei spielen sie eine Rolle bei der Aufrechterhaltung der mechanischen Stabilität der Zellen (Moll et al., 2008). Vorstellbar ist hier eine konservierte Form der Zellinvasion und -transmigration.

Aktuelle Publikationen konnten zeigen, dass Sporozoit der Gattung *Plasmodium* teilweise über mehrere Stunden an der Inokulationsstelle verbleiben (Yamauchi et al., 2007). Interessanterweise konnte auch gezeigt werden, dass sich nagetierinfizierende Sporozoit in den keratinreichen Zellen der Epidermis, Dermis und der Haarfollikel zu blutzellinfizierenden Merozoiten entwickeln können (Gueirard et al., 2010; Jacob et al., 2018). Auch vogelinfizierende Plasmodien invadieren

Zellen nahe der Inokulationsstelle, bevor die daraus entstehenden Merozoiten sich in Zellen des retikuloendothelialen Systems weiterentwickeln (Valkiūnas et al., 2017). Inwiefern eine erste Interaktion an dieser Stelle tatsächlich über Keratine erfolgt, kann an dieser Stelle nur spekuliert werden. Einen Hinweis auf eine besondere Rolle dieser liefert jedoch auch eine kürzlich veröffentlichte epidemiologische Arbeit. Hier konnte nachgewiesen werden, dass Vögel, die an einer unkontrollierten Produktion der Keratine leiden, eine bis zu 2,6fach höhere Wahrscheinlichkeit einer *Plasmodium*-Infektion aufweisen (Wilkinson et al., 2016).

Einige der Proteine, die in den Parasiten selbst Interaktionen mit dem Wirt initiieren, ob nun bei der Invasion, Zelltransmigration, beim Zellaustritt oder bei der Modifikation der Wirtszelle, weisen eine dafür deutlich ausgeprägte Domänenarchitektur auf (vgl. 4.5.2) (Carruthers et al., 2008a). Dazu zählen unter anderem Proteine mit EGF-ähnlichen Domänen (epidermal growth factor-like domains), die sich in den extrazellulären Bereichen einer Vielzahl membrangebundener oder sezernierter Proteine befinden und über die Grenzen unterschiedlichster Organismen hinweg Zell-Zell- oder Zell-Matrix-Interaktionen vermitteln (Appella et al., 1988; Balzar et al., 2001; Baton et al., 2005; Haltom et al., 2015). Auch bei VWA- (Willebrand Faktor A-) und TSR-Domänen (thrombospondin typ 1 repeat) handelt es sich um evolutionär konservierte Domänen diverser eukaryotischer Proteine, denen eine essenzielle Rolle bei der Vermittlung von Zell-Zell- und Zell-Matrix-Interaktionen nachgewiesen werden konnte (Adams et al., 2000; Whittaker et al., 2002). Beide Domänen fungieren als Rezeptoren für diverse Liganden, wie interzelluläre Adhäsionsmoleküle, Komponenten des Komplementsystems, Gerinnungsfaktoren, Glykoproteine der extrazellulären Matrix zu denen auch Kollagene, Laminin, Heparansulfat- oder Chondroitinsulfat-Glykane gehören (Naitza et al., 1998; Adams et al., 2000; Whittaker et al., 2002; Carruthers et al., 2008b; Ramakrishnan et al., 2011). Hier waren die Domänen hauptsächlich in Proteinen nachweisbar, die eine Funktion in der Motilität oder der Zellinvasionen haben, wie CSP, TRAP oder CTRP (vgl. Carruthers et al., 2008a). Eine besondere Rolle spielen auch Proteine der 6-cys Familien. Dabei handelt es sich um Proteine mit der s48/45-Domäne, die bisher nur in Vertretern der Apikomplexa gefunden werden konnte (Arredondo et al., 2017). Strukturell ist die Domäne aus sechs disulfidbindenen Cysteinen aufgebaut, die eine ephrinenähnliche Form annehmen und während der unterschiedlichsten Stadien im Lebenszyklus vermutlich an der Adhäsion an Ephrin-Rezeptoren des Wirts beteiligt sind (van Dijk et al., 2001; Eksi et al., 2006; Kaushansky et al., 2015; Arredondo et al., 2018).

Alle bislang sequenzierten Vertreter der Gattung *Plasmodium* weisen 14 Mitglieder der 6-cys Familie auf (Arredondo et al., 2017). Interessanterweise sind einige davon in *Hepatocystis* wieder verloren gegangen (vgl. 4.5.4). Insgesamt konnten hier nur sieben 6-cys Proteine identifiziert werden. Drei der in *Hepatocystis* nicht nachweisbaren Proteine, werden auch in dem hier dargestellten Datensatz nicht kodiert (Abb. 5.1, vgl. Tab. 7.19). Dabei handelt es sich um die Proteine P41, P230 und P48/45, die auf der Merozoitenoberfläche oder auf der Plasmamembran der Gametozyten exponiert werden (Williamson et al., 1996; van Dijk et al., 2001; Sanders et al., 2005).

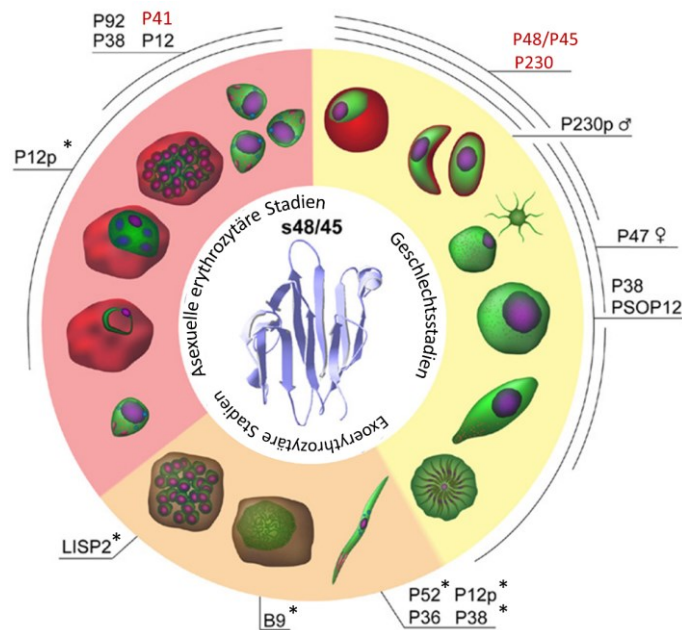


Abb. 5.1: Expression der 6-cys Proteine im Lebenszyklus von *Plasmodium*.

Dargestellt ist die Entwicklung der Vertreter der Gattung *Plasmodium* im Vektor (gelb) und im Wirbeltierwirt (orange und rosa). Proteine, die mindestens eine s48/45 Domäne aufweisen (Struktur im Zentrum) werden in den verschiedenen Stadien des Lebenszyklus wie markiert exprimiert. Proteine, die mit einem Sternchen markiert wurden, sind im Genom von *P. murinus* nicht nachweisbar, aber in *H. tartakowskyi* zu finden. Rot dargestellt sind dagegen die Proteine, die weder in *P. murinus* noch in *Hepatocystis* nachweisbar sind (Abb. modifiziert nach Arredondo et al., 2012).

Gerade in Hinblick auf die paraphyletische Verwandtschaft der Gattung *Plasmodium*, die fledermausinfizierende Parasiten der Gattungen *Nycteria*, *Hepatocystis* und *Polychromophilus* ausschließt (vgl. 4.27), kann ein weiteres Verständnis der abstammungsspezifischen Unterschiede des Kernrepertoires neue Informationen zu Wirtsanpassungen liefern. Tatsächlich haben die meisten Gene, die über alle Plasmodien hinweg konserviert sind, jedoch weder in dem hier untersuchten Genom noch in *Hepatocystis* als kodierend identifiziert wurden, eine unbekannte Funktion. Einige der wenigen Proteine, darunter die Calcium-abhängige Protein Kinase 5, einige CLAG-Proteine (cytoadherence linked asexual protein), das exportierte Protein 1 und das parasitophore Vakuolen-Protein 1, haben dagegen eine wichtige Funktion während der asexuellen Blutstadien in *Plasmodium* (Gupta et al., 2015; Absalon et al., 2018; Morita et al., 2018; Nessel et al., 2020).

Im Vergleich zu anderen Vertretern der Plasmodiidae weist *P. murinus* zudem auch besonders wenige Proteine mit einem klassischen PEXEL-Motiv auf (vgl. 4.4.4.2). Dagegen weisen die in dieser Arbeit identifizierten PEXEL-Orthologen häufig die Aminosäure Lysin anstelle eines Arginins auf. Auch das für die dritte Position typische Leucin ist häufig durch ein Isoleucin ausgetauscht. Festgestellt wurde dies bereits, wenn auch nicht in einem solchen großen Umfang wie bei *P. murinus*, in anderen Plasmodienarten. Dabei wurde davon ausgegangen, dass eine gewisse evolutive Plastizität des Motivs existiert (Pick et al., 2011). Erste Daten deuteten jedoch darauf, dass nur ein Arginin an erster Stelle des pentameren Motivs funktionsfähig sei, während eine etwas aktuellere Untersuchung die Funktionalität einiger alternativer PEXEL nachweisen konnte (Boddey et al., 2009; Sleebs et al., 2014; Schulze et al., 2015). Tatsächlich scheinen die alternativen Motive KxLxE und RxIxEx auch in *P. falciparum*, in Abhängigkeit von der umgebenden Sequenz

erfolgreich einen Export zu induzieren (Schulze et al., 2015). Die Analyse weiterer PEXEL-Orthologer, wie sie in dieser Arbeit identifiziert wurden, kann dazu beitragen, vorausgesetzt natürlich, dass die entsprechenden Proteine in *P. murinus* wirklich exportiert werden, konservierte Sequenzen, Strukturen oder Eigenschaften in den Proteinen zu identifizieren, welche für den Export der Proteine essenziell sind. Damit könnten konservierte Regeln identifiziert werden, die eine Vorhersage weiterer exportierten Proteine in Malariaparasiten erlaubt. Im Idealfall ginge eine solche über den PEXEL-abhängigen Transport hinaus und würde auch PNEPs (PEXEL-negative exported proteins) miteinschließen. Dies könnte dazu beitragen neue Signale oder etwa Sequenzcharakteristika zu identifizieren, die nur einen partialen Export, einen (teil-)stadien-abhängigen Export oder etwa eine spezifische Lokalisierung in Wirtszellkompartimenten gewährleisten. Auf diese Weise können weitere essenzielle Proteine identifiziert werden, die eine tragende Rolle in der Parasit-Wirt-Interaktion spielen.

5.3 Ausblick: Parasiten prägen das Genom ihres Wirts — und umgekehrt

Parasiten stehen unter einem stetigen Anpassungsdruck, der häufig als koevolutionäres Wettrüsten zu beobachten ist, bei dem eine Anpassung der Pathogene, die eine Manipulation des Wirtes erleichtern, Gegenanpassungen der Wirte hervorruft. Der Prozess der Koevolution zwischen Parasiten und ihren Wirten erstreckt sich über viele Generationen hinweg und ist daher nur schwer zu erfassen. Dennoch können einige molekulare Signaturen im Genom der Wirte auf parasitische Infektionen zurückgeführt werden. Insbesondere beim Menschen sind inzwischen eine Reihe an hämoglobinvererbten Störungen (Thalassämie, Sichelzellanämie, HbC und HbE), Erythrozyten-Polymorphismen (Ovalozytose und Duffy-Blutgruppe), Enzymopathien (G6PD-Mangel und PK-Mangel) und immunogenetische Varianten (HLA-Allele, Komplementrezeptor 1, NOS2, Tumor-Nekrose-Faktor- α -Promotor und Chromosom-5q31-q33-Polymorphismen) bekannt, die mit einer gewissen Resistenz gegen Plasmodien in Verbindung stehen (López et al., 2010). Die Beispiele zeigen, dass viele Parasiten zur Erhöhung der genetischen Diversität ihrer Wirte beitragen können. Andererseits können langfristige Parasit-Wirt-Assoziationen für die Parasiten selbst mit einem erheblichen Preis einhergehen: dem Verlust an genetischer Variation, der mit der Fähigkeit verbunden ist, alternative Wirte zu nutzen (Brues, 1920; Yotoko et al., 2006). Dem entgegen stehen allerdings neuere phylogenetische Analysen des Taxons Haemosporida. Dabei deuten die Ergebnisse auf einen einzigen Ursprung der säugetierinfizierenden Gattungen *Polychromophilus*, *Nycteria*, *Hepatocystis* und *Plasmodium* mit einem sekundären Wechsel zurück zu Sauropsiden (Borner et al., 2016; Galen et al., 2018).

Wie diese Arbeit gezeigt hat sind die Vertreter der Plasmodiidae dafür mit einem Grundrepertoire an Proteinen ausgestattet, das eine erfolgreiche Invasion und anschließende Modifikation der Wirte erlaubt, um für den Parasiten günstige Bedingungen zum Überleben zu schaffen. Während bei anderen Haemosporida lediglich ein Wirtswechsel zwischen nahen verwandten Wirten möglich zu sein scheint, kommen fledermausinfizierenden Gattungen bei dem Wechsel zurück zu den Sauropsiden aber auch zwischen unterschiedlichen Säugetierklassen (Huftiere, Primaten und

Nager) eine besondere Rolle zu (Schaer et al., 2013; Lutz et al., 2016; Galen et al., 2018). Mit dem Potential, Fledermäuse zu infizieren, geht sowohl ein Wechsel des Vektors als auch ein Verlust der asexuellen Schizogonie innerhalb von Blutzellen einher (Galen et al., 2018). Inwiefern der Wechsel des Vektors biologisch signifikant ist, konnte im Rahmen dieser Arbeit nicht beantwortet werden. Bei dem konvergenten Verlust der erythrozytären Schizogonie wird dagegen jedoch das Immunsystem der Fledermaus, das kein Bestandteil dieser Arbeit war, eine wesentliche Rolle spielen.

Erst seit wenigen Jahren sind einige molekularen Mechanismen gezeigt worden, durch die Fledermäuse Erkrankungen durch virulenten Krankheitserreger umgehen und auch noch bei potenziell tödlichen Pathogenen offenbar asymptomatisch verbleiben (Brook et al., 2015, 2020; Banerjee et al., 2020). Einerseits steht dies mit einer verstärkten adaptiver Immunantwort andererseits jedoch auch mit einer gewissen Erregertoleranz in Verbindung (Mandl et al., 2018; Brook et al., 2020). Untersuchung von Parasit-Wirt-Interaktionen in Fledermäusen haben daher ein beträchtliches Potenzial, um neue Einblicke in Wirts-Toleranzmechanismen zu gewinnen. Gleichzeitig eröffnet die hier untersuchte Parasit-Wirt-Konstellation neue Möglichkeiten zur Erforschung von Wirtswechseln oder etwa konservierten Mechanismen, wie Wirtszellexporte und -modifikationen. Die weitere Erforschung unterschiedlicher Vertreter der Haemosporida in Hinblick auf konservierte oder neue Pathogenitätsmechanismen wird in Zukunft noch eine wichtige Rolle einnehmen und bietet Potential vielversprechender Angriffspunkte bei der Entwicklung von neuen Therapieansätzen gegen Malaria bei Mensch und Tier.

6 Literaturverzeichnis

- Abbas, A. A., Taylor, L. J., Dothard, M. I., Leiby, J. S., Fitzgerald, A. S., Khatib, L. A., ... Bushman, F. D. (2019). Redondoviridae, a Family of Small, Circular DNA Viruses of the Human Oro-Respiratory Tract Associated with Periodontitis and Critical Illness. *Cell Host & Microbe*, 25(5), 719–729.e4. <https://doi.org/10.1016/j.chom.2019.04.001>
- Absalon, S., Blomqvist, K., Rudlaff, R. M., DeLano, T. J., Pollastri, M. P., & Dvorin, J. D. (2018). Calcium-Dependent Protein Kinase 5 Is Required for Release of Egress-Specific Organelles in *Plasmodium falciparum*. *mBio*, 9(1). <https://doi.org/10.1128/mBio.00130-18>
- Adams, J. C., & Tucker, R. P. (2000). The thrombospondin type 1 repeat (TSR) superfamily: Diverse proteins with related roles in neuronal development. *Developmental Dynamics*, 218(2), 280–299. [https://doi.org/10.1002/\(SICI\)1097-0177\(200006\)218:2<280::AID-DVDY4>3.0.CO;2-0](https://doi.org/10.1002/(SICI)1097-0177(200006)218:2<280::AID-DVDY4>3.0.CO;2-0)
- Adl, S. M., Leander, B. S., Simpson, A. G. B., Archibald, J. M., Anderson, O. R., Bass, D., ... Spiegel, F. W. (2007). Diversity, Nomenclature, and Taxonomy of Protists. *Systematic Biology*, 56(4), 684–689. <https://doi.org/10.1080/10635150701494127>
- Adl, S. M., Simpson, A. G. B., Lane, C. E., Lukeš, J., Bass, D., Bowser, S. S., ... Spiegel, F. W. (2012). The Revised Classification of Eukaryotes. *Journal of Eukaryotic Microbiology*, 59(5), 429–514. <https://doi.org/10.1111/j.1550-7408.2012.00644.x>
- Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., ... Goecks, J. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Research*, 44(W1), W3–W10. <https://doi.org/10.1093/nar/gkw343>
- Aikawa, M. (1971). Parasitological review. Plasmodium: the fine structure of malarial parasites. *Experimental parasitology*, 30(2), 284–320. Abgerufen von <http://www.ncbi.nlm.nih.gov/pubmed/4399774>
- Aikawa, M., & Miller, L. H. (2008). *Structural Alteration of the Erythrocyte Membrane During Malarial Parasite Invasion and Intraerythrocytic Development*. <https://doi.org/10.1002/9780470715444.ch4>
- Alexander, D. L., Mital, J., Ward, G. E., Bradley, P., & Boothroyd, J. C. (2005). Identification of the Moving Junction Complex of *Toxoplasma gondii*: A Collaboration between Distinct Secretory Organelles. *PLoS Pathogens*, 1(2), e17. <https://doi.org/10.1371/journal.ppat.0010017>
- Alizon, S., de Roode, J. C., & Michalakis, Y. (2013). Multiple infections and the evolution of virulence. *Ecology Letters*, 16(4), 556–567. <https://doi.org/10.1111/ele.12076>
- Alonge, M., Soyk, S., Ramakrishnan, S., Wang, X., Goodwin, S., Sedlazeck, F. J., ... Schatz, M. C. (2019). RaGOO: fast and accurate reference-guided scaffolding of draft genomes. *Genome Biology*, 20(1), 224. <https://doi.org/10.1186/s13059-019-1829-6>
- Alters, S. (1999). *Biology: Understanding Life* (Third Edit; J. H. Hauck, B. L. McKean, M. Hill, & J. Angel, Hrsg.). Sudbury: Jones and Barlett Publishers.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Amino, R., Thiberge, S., Martin, B., Celli, S., Shorte, S., Frischknecht, F., & Ménard, R. (2006). Quantitative imaging of *Plasmodium* transmission from mosquito to mammal. *Nature Medicine*, 12(2), 220–224. <https://doi.org/10.1038/nm1350>
- Anand, G., Reddy, K. S., Pandey, A. K., Mian, S. Y., Singh, H., Mittal, S. A., ... Gaur, D. (2016). A novel *Plasmodium falciparum* rhoptry associated adhesin mediates erythrocyte invasion through the sialic-acid dependent pathway. *Scientific Reports*, 6(1), 29185. <https://doi.org/10.1038/srep29185>
- Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Abgerufen 14. Mai 2016, von <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Angell, R. L., Butlin, R. K., & Altringham, J. D. (2013). Sexual Segregation and Flexible Mating Patterns in Temperate Bats. *PLoS ONE*, 8(1), e54194. <https://doi.org/10.1371/journal.pone.0054194>
- Appella, E., Weber, I. T., & Blasi, F. (1988). Structure and function of epidermal growth factor-like regions in proteins. *FEBS Letters*, 231(1), 1–4. [https://doi.org/10.1016/0014-5793\(88\)80690-2](https://doi.org/10.1016/0014-5793(88)80690-2)
- Arakawa, K. (2016). No evidence for extensive horizontal gene transfer from the draft genome of a tardigrade. *Proceedings of the National Academy of Sciences*, 113(22), E3057–E3057. <https://doi.org/10.1073/pnas.1602711113>
- Arredondo, S. A., & Kappe, S. H. I. (2017). The s48/45 six-cysteine proteins: mediators of interaction throughout the *Plasmodium* life cycle. *International Journal for Parasitology*, 47(7), 409–423. <https://doi.org/10.1016/j.ijpara.2016.10.002>
- Arredondo, S. A., Swearingen, K. E., Martinson, T., Steel, R., Dankwa, D. A., Harupa, A., ... Kappe, S. H. I. (2018). The Micronemal *Plasmodium* Proteins P36 and P52 Act in Concert to Establish the Replication-Permissive Compartment Within Infected Hepatocytes. *Frontiers in Cellular and Infection Microbiology*, 8. <https://doi.org/10.3389/fcimb.2018.00413>
- Arredondo, S. A., Cai, M., Takayama, Y., MacDonald, N. J., Anderson, D. E., Aravind, L., ... Miller, L. H. (2012). Structure of the *Plasmodium* 6-cysteine s48/45 domain. *Proceedings of the National Academy of Sciences*, 109(17), 6692–6697. <https://doi.org/10.1073/pnas.1204363109>
- Arumugam, T. U., Takeo, S., Yamasaki, T., Thonkukiatkul, A., Miura, K., Otsuki, H., ... Tsuboi, T. (2011). Discovery of GAMA,

- a Plasmodium falciparum Merozoite Micronemal Protein, as a Novel Blood-Stage Vaccine Candidate Antigen. *Infection and Immunity*, 79(11), 4523–4532. <https://doi.org/10.1128/IAI.05412-11>
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., ... Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1), 25–29. <https://doi.org/10.1038/75556>
- Auburn, S., Campino, S., Clark, T. G., Djimde, A. A., Zongo, I., Pinches, R., ... Kwiatkowski, D. P. (2011). An Effective Method to Purify Plasmodium falciparum DNA Directly from Clinical Blood Samples for Whole Genome High-Throughput Sequencing. *PLoS ONE*, 6(7), e22213. <https://doi.org/10.1371/journal.pone.0022213>
- August, T. A., Nunn, M. A., Fensome, A. G., Linton, D. M., & Mathews, F. (2014). Sympatric Woodland Myotis Bats Form Tight-Knit Social Groups with Exclusive Roost Home Ranges. *PLoS ONE*, 9(10), e112225. <https://doi.org/10.1371/journal.pone.0112225>
- Aurrecoechea, C., Brestelli, J., Brunk, B. P., Dommer, J., Fischer, S., Gajria, B., ... Wang, H. (2009). PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Research*, 37(Database), D539–D543. <https://doi.org/10.1093/nar/gkn814>
- Aurrecoechea, C., Barreto, A., Basenko, E. Y., Brestelli, J., Brunk, B. P., Cade, S., ... Zheng, J. (2017). EuPathDB: the eukaryotic pathogen genomics database resource. *Nucleic Acids Research*, 45(D1), D581–D591. <https://doi.org/10.1093/nar/gkw1105>
- Balaji, S. (2005). Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Research*, 33(13), 3994–4006. <https://doi.org/10.1093/nar/gki709>
- Ballouz, S., Dobin, A., & Gillis, J. A. (2019). Is it time to change the reference genome? *Genome Biology*, 20(1), 159. <https://doi.org/10.1186/s13059-019-1774-4>
- Balzar, M., Briaire-de Bruijn, I. H., Rees-Bakker, H. A. M., Prins, F. A., Helfrich, W., de Leij, L., ... Litvinov, S. V. (2001). Epidermal Growth Factor-Like Repeats Mediate Lateral and Reciprocal Interactions of Ep-CAM Molecules in Homophilic Adhesions. *Molecular and Cellular Biology*, 21(7), 2570–2580. <https://doi.org/10.1128/MCB.21.7.2570-2580.2001>
- Bandouchova, H., Bartonicka, T., Berkova, H., Brichta, J., Cerny, J., Kovacova, V., ... Pikula, J. (2015). Pseudogymnoascus destructans: Evidence of Virulent Skin Invasion for Bats Under Natural Conditions, Europe. *Transboundary and Emerging Diseases*, 62(1), 1–5. <https://doi.org/10.1111/tbed.12282>
- Banerjee, A., Baker, M. L., Kulcsar, K., Misra, V., Plowright, R., & Mossman, K. (2020). Novel Insights Into Immune Systems of Bats. *Frontiers in Immunology*, 11. <https://doi.org/10.3389/fimmu.2020.00026>
- Bannister, L. H., Hopkins, J. M., Fowler, R. E., Krishna, S., & Mitchell, G. H. (2000). A brief illustrated guide to the ultrastructure of Plasmodium falciparum asexual blood stages. *Parasitology today*, 16(10), 427–433. Abgerufen von <http://www.ncbi.nlm.nih.gov/pubmed/11006474>
- Bao, W., Kojima, K. K., & Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA*, 6(1), 11. <https://doi.org/10.1186/s13100-015-0041-9>
- Barnett, D. W., Garrison, E. K., Quinlan, A. R., Stromberg, M. P., & Marth, G. T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics*, 27(12), 1691–1692. <https://doi.org/10.1093/bioinformatics/btr174>
- Barrow, L. N., Allen, J. M., Huang, X., Bensch, S., & Witt, C. C. (2019). Genomic sequence capture of haemosporidian parasites: Methods and prospects for enhanced study of host–parasite evolution. *Molecular Ecology Resources*, 19(2), 400–410. <https://doi.org/10.1111/1755-0998.12977>
- Bashey, F. (2015). Within-host competitive interactions as a mechanism for the maintenance of parasite diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1675), 20140301. <https://doi.org/10.1098/rstb.2014.0301>
- Baton, L. A., & Ranford-Cartwright, L. C. (2005). Do malaria ookinete surface proteins P25 and P28 mediate parasite entry into mosquito midgut epithelial cells? *Malaria journal*, 4, 15. <https://doi.org/10.1186/1475-2875-4-15>
- Baum, J., Richard, D., Healer, J., Rug, M., Krnajska, Z., Gilberger, T.-W., ... Cowman, A. F. (2006). A Conserved Molecular Motor Drives Cell Invasion and Gliding Motility across Malaria Life Cycle Stages and Other Apicomplexan Parasites. *Journal of Biological Chemistry*, 281(8), 5197–5208. <https://doi.org/10.1074/jbc.M509807200>
- Baum, J., Papenfuss, A. T., Baum, B., Speed, T. P., & Cowman, A. F. (2006). Regulation of apicomplexan actin-based motility. *Nature Reviews Microbiology*, 4(8), 621–628. <https://doi.org/10.1038/nrmicro1465>
- Beck, J. R., Muralidharan, V., Oksman, A., & Goldberg, D. E. (2014). PTEX component HSP101 mediates export of diverse malaria effectors into host erythrocytes. *Nature*, 511(7511), 592–595. <https://doi.org/10.1038/nature13574>
- Beeson, J. G., Drew, D. R., Boyle, M. J., Feng, G., Fowkes, F. J. I., & Richards, J. S. (2016). Merozoite surface proteins in red blood cell invasion, immunity and vaccines against malaria. *FEMS Microbiology Reviews*, 40(3), 343–372. <https://doi.org/10.1093/femsre/fuw001>
- Belachew, E. B. (2018). Immune Response and Evasion Mechanisms of Plasmodium falciparum Parasites. *Journal of Immunology Research*, 2018, 1–6. <https://doi.org/10.1155/2018/6529681>
- Benavente, E. D., Gomes, A. R., De Silva, J. R., Grigg, M., Walker, H., Barber, B. E., ... Campino, S. (2019). Whole genome sequencing of amplified Plasmodium knowlesi DNA from unprocessed blood reveals genetic exchange events between Malaysian Peninsular and Borneo subpopulations. *Scientific Reports*, 9(1), 9873. <https://doi.org/10.1038/s41598-019-46398-z>
- Benjamini, Y., & Speed, T. P. (2012). Summarizing and correcting the GC content bias in high-throughput sequencing. *Nucleic Acids*

- Research*, 40(10), e72–e72. <https://doi.org/10.1093/nar/gks001>
- Bensch, S., Canbäck, B., DeBarry, J. D., Johansson, T., Hellgren, O., Kissinger, J. C., ... Valkiūnas, G. (2016). The Genome of *Haemoproteus tartakovskyi* and Its Relationship to Human Malaria Parasites. *Genome Biology and Evolution*, 8(5), 1361–1373. <https://doi.org/10.1093/gbe/evw081>
- Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research*, 27(2), 573–580. <https://doi.org/10.1093/nar/27.2.573>
- Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsche, G., ... Stadler, P. F. (2013). MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, 69(2), 313–319. <https://doi.org/10.1016/j.ympev.2012.08.023>
- Besteiro, S., Michelin, A., Poncet, J., Dubremetz, J.-F., & Lebrun, M. (2009). Export of a *Toxoplasma gondii* Rhoptry Neck Protein Complex at the Host Cell Membrane to Form the Moving Junction during Invasion. *PLoS Pathogens*, 5(2), e1000309. <https://doi.org/10.1371/journal.ppat.1000309>
- Bhak, Y., Jeon, Y., Jeon, S., Chung, O., Jho, S., Jun, J., ... Bhak, J. (2017). *Myotis rufoniger* genome sequence and analyses: *M. rufoniger*'s genomic feature and the decreasing effective population size of *Myotis* bats. *PLOS ONE*, 12(7), e0180418. <https://doi.org/10.1371/journal.pone.0180418>
- Billker, O., Shaw, M. K., Margos, G., & Sinden, R. E. (1997). The roles of temperature, pH and mosquito factors as triggers of male and female gametogenesis of *Plasmodium berghei* in vitro. *Parasitology*, 115 (Pt 1), 1–7. Abgerufen von <http://www.ncbi.nlm.nih.gov/pubmed/9280891>
- Boddey, J. A., Moritz, R. L., Simpson, R. J., & Cowman, A. F. (2009). Role of the *Plasmodium* Export Element in Trafficking Parasite Proteins to the Infected Erythrocyte. *Traffic*, 10(3), 285–299. <https://doi.org/10.1111/j.1600-0854.2008.00864.x>
- Boetzer, M., & Pirovano, W. (2012). Toward almost closed genomes with GapFiller. *Genome Biology*, 13(6), R56. <https://doi.org/10.1186/gb-2012-13-6-r56>
- Bogdanowicz, W. (1994). *NoMyotis daubentonii*. *Mammalian Species*, 475, 1–9. <https://doi.org/10.2307/3504215>
- Böhme, U., Otto, T. D., Cotton, J. A., Steinbiss, S., Sanders, M., Oyola, S. O., ... Berriman, M. (2018). Complete avian malaria parasite genomes reveal features associated with lineage-specific evolution in birds and mammals. *Genome Research*, 28(4), 547–560. <https://doi.org/10.1101/gr.218123.116>
- Boothby, T. C., Tenlen, J. R., Smith, F. W., Wang, J. R., Patanella, K. A., Osborne Nishimura, E., ... Goldstein, B. (2015). Evidence for extensive horizontal gene transfer from the draft genome of a tardigrade. *Proceedings of the National Academy of Sciences*, 112(52), 15976–15981. <https://doi.org/10.1073/pnas.1510461112>
- Boratyn, G. M., Thierry-Mieg, J., Thierry-Mieg, D., Busby, B., & Madden, T. L. (2019). Magic-BLAST, an accurate RNA-seq aligner for long and short reads. *BMC Bioinformatics*, 20(1), 405. <https://doi.org/10.1186/s12859-019-2996-x>
- Borner, J., Pick, C., Thiede, J., Kolawole, O. M., Kingsley, M. T., Schulze, J., ... Burmester, T. (2016). Phylogeny of haemosporidian blood parasites revealed by a multi-gene approach. *Molecular Phylogenetics and Evolution*, 94(Pt A), 221–231. <https://doi.org/10.1016/j.ympev.2015.09.003>
- Borner, J., & Burmester, T. (2017). Parasite infection of public databases: a data mining approach to identify apicomplexan contaminations in animal genome and transcriptome assemblies. *BMC Genomics*, 18(1), 100. <https://doi.org/10.1186/s12864-017-3504-1>
- Boundenga, L., Makanga, B., Ollomo, B., Gilabert, A., Rougeron, V., Mve-Ondo, B., ... Paupy, C. (2016). Haemosporidian Parasites of Antelopes and Other Vertebrates from Gabon, Central Africa. *PLOS ONE*, 11(2), e0148958. <https://doi.org/10.1371/journal.pone.0148958>
- Boundenga, L., Ngoubangoye, B., Mombo, I. M., Tsubmou, T. A., Renaud, F., Rougeron, V., & Prugnolle, F. (2018). Extensive diversity of malaria parasites circulating in Central African bats and monkeys. *Ecology and evolution*, 8(21), 10578–10586. <https://doi.org/10.1002/ece3.4539>
- Boyle, M. J., Wilson, D. W., Richards, J. S., Riglar, D. T., Tetteh, K. K. A., Conway, D. J., ... Beeson, J. G. (2010). Isolation of viable *Plasmodium falciparum* merozoites to define erythrocyte invasion events and advance vaccine and drug development. *Proceedings of the National Academy of Sciences*, 107(32), 14378–14383. <https://doi.org/10.1073/pnas.1009198107>
- Brayton, K. A., Lau, A. O. T., Herndon, D. R., Hannick, L., Kappmeyer, L. S., Berens, S. J., ... Nene, V. M. (2007). Genome Sequence of *Babesia bovis* and Comparative Analysis of Apicomplexan Hemoprotozoa. *PLoS Pathogens*, 3(10), e148. <https://doi.org/10.1371/journal.ppat.0030148>
- Bright, A., Tewhey, R., Abeles, S., Chuquiyauri, R., Llanos-Cuentas, A., Ferreira, M. U., ... Winzeler, E. A. (2012). Whole genome sequencing analysis of *Plasmodium vivax* using whole genome capture. *BMC Genomics*, 13(1), 262. <https://doi.org/10.1186/1471-2164-13-262>
- Brook, C. E., & Dobson, A. P. (2015). Bats as „special“ reservoirs for emerging zoonotic pathogens. *Trends in Parasitology*, 23(3), 172–180. <https://doi.org/10.1016/j.tim.2014.12.004>
- Brook, C. E., Boots, M., Chandran, K., Dobson, A. P., Drosten, C., Graham, A. L., ... van Leeuwen, A. (2020). Accelerated viral dynamics in bat cell lines, with implications for zoonotic emergence. *eLife*, 9. <https://doi.org/10.7554/eLife.48401>
- Brues, C. T. (1920). The selection of food-plants by insects, with special preference to lepidopterous larvae. *The American Naturalist*,

- 54(633), 313–332.
- Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, 12(1), 59–60. <https://doi.org/10.1038/nmeth.3176>
- Buchmann, K. (2014). Evolution of Innate Immunity: Clues from Invertebrates via Fish to Mammals. *Frontiers in Immunology*, 5. <https://doi.org/10.3389/fimmu.2014.00459>
- Burgin, C. J., Colella, J. P., Kahn, P. L., & Upham, N. S. (2018). How many species of mammals are there? *Journal of Mammalogy*, 99(1), 1–14. <https://doi.org/10.1093/jmammal/gyx147>
- Bushell, E., Gomes, A. R., Sanderson, T., Anar, B., Girling, G., Herd, C., ... Billker, O. (2017). Functional Profiling of a Plasmodium Genome Reveals an Abundance of Essential Genes. *Cell*, 170(2), 260–272.e8. <https://doi.org/10.1016/j.cell.2017.06.030>
- Calisher, C. H., Childs, J. E., Field, H. E., Holmes, K. V., & Schountz, T. (2006). Bats: Important Reservoir Hosts of Emerging Viruses. *Clinical microbiology reviews*, 19(3), 531–545. <https://doi.org/10.1128/CMR.00017-06>
- Cantarel, B. L., Korf, I., Robb, S. M. C., Parra, G., Ross, E., Moore, B., ... Yandell, M. (2007). MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research*, 18(1), 188–196. <https://doi.org/10.1101/gr.6743907>
- Cao, J., Kaneko, O., Thongkukiattul, A., Tachibana, M., Otsuki, H., Gao, Q., ... Torii, M. (2009). Rhoptry neck protein RON2 forms a complex with microneme protein AMA1 in Plasmodium falciparum merozoites. *Parasitology International*, 58(1), 29–35. <https://doi.org/10.1016/j.parint.2008.09.005>
- Carlton, J. M.-R. (2003). The Plasmodium vivax genome sequencing project. *Trends in Parasitology*, 19, 227–231. [https://doi.org/10.1016/S1471-4922\(03\)00066-7](https://doi.org/10.1016/S1471-4922(03)00066-7)
- Carlton, J. M.-R., Angiuoli, S. V., Suh, B. B., Kooij, T. W., Pertea, M., Silva, J. C., ... Carucci, D. J. (2002). Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii. *Nature*, 419(6906), 512–519. <https://doi.org/10.1038/nature01099>
- Carlton, J. M.-R., Muller, R., Yowell, C. A., Fluegge, M. R., Sturrock, K. A., Pritt, J. R., ... Dame, J. B. (2001). Profiling the malaria genome: a gene survey of three species of malaria parasite with comparison to other apicomplexan species. *Molecular and Biochemical Parasitology*, 118(2), 201–210. [https://doi.org/10.1016/S0166-6851\(01\)00371-1](https://doi.org/10.1016/S0166-6851(01)00371-1)
- Carlton, J. M.-R., Silva, J., & Hall, N. (2005). The genome of model malaria parasites, and comparative genomics. *Current issues in molecular biology*, 7(1), 23–37. <https://doi.org/10.15580778>
- Carlton, J. M., Adams, J. H., Silva, J. C., Bidwell, S. L., Lorenzi, H., Caler, E., ... Fraser-Liggett, C. M. (2008). Comparative genomics of the neglected human malaria parasite Plasmodium vivax. *Nature*, 455(7214), 757–763. <https://doi.org/10.1038/nature07327>
- Carreno, R. A., Kissinger, J. C., McCutchan, T. F., & Barta, J. R. (1997). Phylogenetic analysis of haemosporinid parasites (apicomplexa: Haemosporina) and their coevolution with vectors and intermediate hosts. *Archiv für Protistenkunde*, 148(3), 245–252. [https://doi.org/10.1016/S0003-9365\(97\)80005-X](https://doi.org/10.1016/S0003-9365(97)80005-X)
- Carruthers, V. B., & Tomley, F. M. (2008a). Microneme proteins in apicomplexans. *Sub-cellular biochemistry*, 47, 33–45. https://doi.org/10.1007/978-0-387-78267-6_2
- Carruthers, V. B., & Tomley, F. M. (2008b). Microneme Proteins in Apicomplexans. In *Molecular Mechanisms of Parasite Invasion* (S. 33–45). https://doi.org/10.1007/978-0-387-78267-6_2
- Carter, R., & Nijhout, M. (1977). Control of gamete formation (exflagellation) in malaria parasites. *Science*, 195(4276), 407–409. <https://doi.org/10.1126/science.12566>
- Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M.-A., Barrell, B. G., & Parkhill, J. (2005). ACT: the Artemis comparison tool. *Bioinformatics*, 21(16), 3422–3423. <https://doi.org/10.1093/bioinformatics/bti553>
- Castresana, J. (2000). Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Molecular Biology and Evolution*, 17(4), 540–552. <https://doi.org/10.1093/oxfordjournals.molbev.a026334>
- Castro, G. A., & Olson, L. J. (1996). Introduction to Parasitology. In S. Baron (Hrsg.), *Medical Microbiology* (4. Aufl.). Abgerufen von <https://www.ncbi.nlm.nih.gov/books/NBK8262/>
- Centers for Disease Control and Prevention. (2018). Malaria Lifecycle. Abgerufen 4. August 2019, von <https://www.cdc.gov/malaria/about/biology/index.html>
- Cézilly, F., Thomas, F., Médoc, V., & Perrmont-Minot, M. J. (2010). Host-manipulation by parasites with complex life cycles: adaptive or not? *Trends in Parasitology*, 26(6), 311–317. <https://doi.org/10.1016/j.pt.2010.03.009>
- Chagas, C. R. F., Bukauskaitė, D., Ilgūnas, M., Bernotienė, R., Iezhova, T., & Valkiūnas, G. (2019). Sporogony of four Haemoproteus species (Haemosporida: Haemoproteidae), with report of in vitro ookinetes of Haemoproteus hirundinis: phylogenetic inference indicates patterns of haemosporidian parasite ookinete development. *Parasites & Vectors*, 12(1), 422. <https://doi.org/10.1186/s13071-019-3679-1>
- Chakarov, N., Greiner, J. F. W., Hauser, S., Schützmann, D., Krüger, O., Börner, M., & Widera, D. (2012). Label-free enrichment of avian Leucocytozoon using flow cytometric sorting. *Parasitology*, 139(12), 1547–1552. <https://doi.org/10.1017/S0031182012000972>

- Chang, H. H., Falick, A. M., Carlton, P. M., Sedat, J. W., DeRisi, J. L., & Marletta, M. A. (2008). N-terminal processing of proteins exported by malaria parasites. *Molecular and Biochemical Parasitology*, 160(2), 107–115. <https://doi.org/10.1016/j.molbiopara.2008.04.011>
- Chomel, B. B., Boulouis, H.-J., Breitschwerdt, E. B., Kasten, R. W., Vayssier-Taussat, M., Birtles, R. J., ... Dehio, C. (2009). Ecological fitness and strategies of adaptation of Bartonella species to their hosts and vectors. *Veterinary Research*, 40(2), 29. <https://doi.org/10.1051/vetres/2009011>
- Christe, P., Glaizot, O., Evanno, G., Bruyndonckx, Nadia Devevey, G., Yannic, G., Patthey, P., ... Arlettau, R. (2007). Host sex and ectoparasites choice: preference for, and higher survival on female hosts. *Journal of Animal Ecology*, 76(4), 703–710. <https://doi.org/10.1111/j.1365-2656.2007.01255.x>
- Clarke-Crespo, E., de León, G. P.-P., Montiel-Ortega, S., & Rubio-Godoy, M. (2017). Helminth Fauna Associated with Three Neotropical Bat Species (Chiroptera: Mormoopidae) in Veracruz, México. *Journal of Parasitology*, 103(4), 338–342. <https://doi.org/10.1645/16-59>
- Clarke, E. L., Sundararaman, S. A., Seifert, S. N., Bushman, F. D., Hahn, B. H., & Brisson, D. (2017). swga: a primer design toolkit for selective whole genome amplification. *Bioinformatics*, 33(14), 2071–2077. <https://doi.org/10.1093/bioinformatics/btx118>
- Claros, M. G., & Vincens, P. (1996). Computational Method to Predict Mitochondrially Imported Proteins and their Targeting Sequences. *European Journal of Biochemistry*, 241(3), 779–786. <https://doi.org/10.1111/j.1432-1033.1996.00779.x>
- Clayton, A. M., Dong, Y., & Dimopoulos, G. (2014). The Anopheles Innate Immune System in the Defense against Malaria Infection. *Journal of Innate Immunity*, 6(2), 169–181. <https://doi.org/10.1159/000353602>
- Cocking, J. H., Deberg, M., Schupp, J., Sahl, J., Wiggins, K., Porty, A., ... Pearson, T. (2020). Selective whole genome amplification and sequencing of Coxiella burnetii directly from environmental samples. *Genomics*, 112(2), 1872–1878. <https://doi.org/10.1016/j.ygeno.2019.10.022>
- Coral, A. A., Valkiūnas, G., González, A. D., & Matta, N. E. (2015). In vitro development of Haemoproteus columbae (Haemosporida: Haemoproteidae), with perspectives for genomic studies of avian haemosporidian parasites. *Experimental Parasitology*, 157, 163–169. <https://doi.org/10.1016/j.exppara.2015.08.003>
- Corradi, N., Pombert, J.-F., Farinelli, L., Didier, E. S., & Keeling, P. J. (2010). The complete sequence of the smallest known nuclear genome from the microsporidian Encephalitozoon intestinalis. *Nature Communications*, 1(1), 77. <https://doi.org/10.1038/ncomms1082>
- Cowell, A. N., Loy, D. E., Sundararaman, S. A., Valdivia, H., Fisch, K., Lescano, A. G., ... Winzeler, E. A. (2017). Selective Whole-Genome Amplification Is a Robust Method That Enables Scalable Whole-Genome Sequencing of Plasmodium vivax from Unprocessed Clinical Samples. *mBio*, 8(1). <https://doi.org/10.1128/mBio.02257-16>
- Cowman, A. F., & Crabb, B. S. (2006). Invasion of Red Blood Cells by Malaria Parasites. *Cell*, 124(4), 755–766. <https://doi.org/10.1016/j.cell.2006.02.006>
- Cox, F. E. . (2002). Systematics of the parasitic Protozoa. *Trends in Parasitology*, 18(3), 108. [https://doi.org/10.1016/S1471-4922\(01\)02196-1](https://doi.org/10.1016/S1471-4922(01)02196-1)
- Crawford, N. G., Faircloth, B. C., McCormack, J. E., Brumfield, R. T., Winker, K., & Glenn, T. C. (2012). More than 1000 ultraconserved elements provide evidence that turtles are the sister group of archosaurs. *Biology Letters*, 8(5), 783–786. <https://doi.org/10.1098/rsbl.2012.0331>
- Crofton, H. D. (1971). A quantitative approach to parasitism. *Parasitology*, 62(02), 179. <https://doi.org/10.1017/S0031182000071420>
- Cuesta-Astroz, Y., Santos, A., Oliveira, G., & Jensen, L. J. (2019). Analysis of Predicted Host–Parasite Interactomes Reveals Commonalities and Specificities Related to Parasitic Lifestyle and Tissues Tropism. *Frontiers in Immunology*, 10. <https://doi.org/10.3389/fimmu.2019.00212>
- Dame, J. B., & McCutchan, T. F. (1987). Plasmodium falciparum: Hoechst dye 33258-CsCl ultracentrifugation for separating parasite and host DNAs. *Experimental Parasitology*, 64(2), 264–266. [https://doi.org/10.1016/0014-4894\(87\)90152-4](https://doi.org/10.1016/0014-4894(87)90152-4)
- Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics*, 27(8), 1164–1165. <https://doi.org/10.1093/bioinformatics/btr088>
- Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9(8), 772–772. <https://doi.org/10.1038/nmeth.2109>
- Davies, H. M., Thalassinou, K., & Osborne, A. R. (2016). Expansion of Lysine-rich Repeats in Plasmodium Proteins Generates Novel Localization Sequences That Target the Periphery of the Host Erythrocyte. *Journal of Biological Chemistry*, 291(50), 26188–26207. <https://doi.org/10.1074/jbc.M116.761213>
- de Koning-Ward, T. F., Gilson, P. R., Boddey, J. A., Rug, M., Smith, B. J., Papenfuss, A. T., ... Crabb, B. S. (2009). A newly discovered protein export machine in malaria parasites. *Nature*, 459(7249), 945–949. <https://doi.org/10.1038/nature08104>
- Deane, L. M., & Deane, M. P. (1961). Sobre dois hemocitozoários encontrados em mamíferos silvestres da Região Amazônica. *Revista do Instituto de Medicina Tropical de São*, 3(3), 107–110.
- DeBarry, J. D., & Kissinger, J. C. (2011). Jumbled Genomes: Missing Apicomplexan Synteny. *Molecular Biology and Evolution*, 28(10), 2855–2871. <https://doi.org/10.1093/molbev/msr103>

- Deiner, K., Renshaw, M. A., Li, Y., Olds, B. P., Lodge, D. M., & Pfrender, M. E. (2017). Long-range PCR allows sequencing of mitochondrial genomes from environmental DNA. *Methods in Ecology and Evolution*, 8(12), 1888–1898. <https://doi.org/10.1111/2041-210X.12836>
- Dick, C. W., & Patterson, B. D. (2006). Bat flies: Obligate ectoparasites of bats. In *Micromammals and Macroparasites* (S. 179–194). https://doi.org/10.1007/978-4-431-36025-4_11
- Dick, C. W., & Pospischil, R. (2015). Nycteribiidae (Bat Flies). In *Encyclopedia of Parasitology* (S. 1–4). https://doi.org/10.1007/978-3-642-27769-6_3462-1
- Dietz, M., & Kalko, E. K. V. (2006). Seasonal changes in daily torpor patterns of free-ranging female and male Daubenton's bats (*Myotis daubentonii*). *Journal of Comparative Physiology B*, 176(3), 223–231. <https://doi.org/10.1007/s00360-005-0043-x>
- Dietz, M., Encarnação, J. A., & Kalko, E. K. V. (2006). Small scale distribution patterns of female and male Daubenton's bats (*Myotis daubentonii*). *Acta Chiropterologica*, 8(2), 403–415. [https://doi.org/https://doi.org/10.3161/1733-5329\(2006\)8\[403:SSDPOF\]2.0.CO;2](https://doi.org/https://doi.org/10.3161/1733-5329(2006)8[403:SSDPOF]2.0.CO;2)
- Dionisi, A. (1898a). Les parasites endoglobulaires des chauvessouris. *Atti della Reale Accademia Lincei*, 7, 153–156.
- Dionisi, A. (1898b). Un parassita del globulo rosso in una specie di pipistrello (*Miniopterus Schreibersii* Kuhl.). *Atti della Reale Accademia Lincei*, 7, 214–215.
- Dong, D., Lei, M., Hua, P., Pan, Y.-H., Mu, S., Zheng, G., ... Zhang, S. (2017). The Genomes of Two Bat Species with Long Constant Frequency Echolocation Calls. *Molecular Biology and Evolution*, 34(1), 20–34. <https://doi.org/10.1093/molbev/msw231>
- Douzery, E. J. P., Snell, E. A., Baptiste, E., Delsuc, F., & Philippe, H. (2004). The timing of eukaryotic evolution: Does a relaxed molecular clock reconcile proteins and fossils? *Proceedings of the National Academy of Sciences*, 101(43), 15386–15391. <https://doi.org/10.1073/pnas.0403984101>
- Dragu, A., Csorba, G., & Bancila, R. (2019). Ecological character displacement in mandibular morphology of three sympatric horseshoe bats. *Hystrix the Italian Journal of Mammalogy*, 30(1), 51–58. <https://doi.org/https://doi.org/10.4404/hystrix-00104-2018>
- Durand, P. M., Oelofse, A. J., & Coetzer, T. L. (2006). An analysis of mobile genetic elements in three *Plasmodium* species and their potential impact on the nucleotide composition of the *P. falciparum* genome. *BMC genomics*, 7, 282. <https://doi.org/10.1186/1471-2164-7-282>
- Duval, L., Robert, V., Csorba, G., Hassanin, A., Randrianarivelojosia, M., Walston, J., ... Arieu, F. (2007). Multiple host-switching of Haemosporidia parasites in bats. *Malaria Journal*, 6(1), 157. <https://doi.org/10.1186/1475-2875-6-157>
- Duval, L., Mejean, C., Maganga, G. D., Makanga, B. K., Mangama Koumba, L. B., Peirce, M. A., ... Bourgarel, M. (2012). The chiropteran haemosporidian *Polychromophilus melanipherus*: A worldwide species complex restricted to the family Miniopteridae. *Infection, Genetics and Evolution*, 12(7), 1558–1566. <https://doi.org/10.1016/j.meegid.2012.06.006>
- Eddy, S. R. (2011). Accelerated Profile HMM Searches. *PLoS Computational Biology*, 7(10), e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>
- Eksi, S., Czesny, B., van Gemert, G.-J., Sauerwein, R. W., Eling, W., & Williamson, K. C. (2006). Malaria transmission-blocking antigen, Pfs230, mediates human red blood cell binding to exflagellating male parasites and oocyst production. *Molecular Microbiology*, 61(4), 991–998. <https://doi.org/10.1111/j.1365-2958.2006.05284.x>
- Encarnação, J. A., & Dietz, M. (2006). Estimation of food intake and ingested energy in Daubenton's bats (*Myotis daubentonii*) during pregnancy and spermatogenesis. *European Journal of Wildlife Research*, 52(4), 221–227. <https://doi.org/10.1007/s10344-006-0046-2>
- Encarnação, J. A., Becker, N. I., & Ekschmitt, K. (2010). When do Daubenton's bats (*Myotis daubentonii*) fly far for dinner? *Canadian Journal of Zoology*, 88(12), 1192–1201. <https://doi.org/10.1139/Z10-085>
- Encarnação, J. A., Kierdorf, U., Holweg, D., Jasnoch, U., & Wolters, V. (2005). Sex-related differences in roost-site selection by Daubenton's bats *Myotis daubentonii* during the nursery period. *Mammal Review*, 35(3–4), 285–294. <https://doi.org/10.1111/j.1365-2907.2005.00066.x>
- Escalante, A. A., & Ayala, F. J. (1994). Phylogeny of the malarial genus *Plasmodium*, derived from rRNA gene sequences. *Proceedings of the National Academy of Sciences*, 91(24), 11373–11377. <https://doi.org/10.1073/pnas.91.24.11373>
- Escalante, A. A., & Ayala, F. J. (1995). Evolutionary origin of *Plasmodium* and other Apicomplexa based on rRNA genes. *Proceedings of the National Academy of Sciences*, 92(13), 5793–5797. <https://doi.org/10.1073/pnas.92.13.5793>
- EUROBATS. (1991). Agreement on the Conservation of Populations of European Bats. Abgerufen 12. August 2019, von https://www.eurobats.org/about_eurobats/introduction_to_agreement
- Faust, C., & Dobson, A. P. (2015). Primate malarias: Diversity, distribution and insights for zoonotic *Plasmodium*. *One Health*, 1, 66–75. <https://doi.org/10.1016/j.onehlt.2015.10.001>
- Feagin, J. E., Harrell, M. I., Lee, J. C., Coe, K. J., Sands, B. H., Cannone, J. J., ... Gutell, R. R. (2012). The Fragmented Mitochondrial Ribosomal RNAs of *Plasmodium falciparum*. *PLoS ONE*, 7(6), e38320. <https://doi.org/10.1371/journal.pone.0038320>
- Feehery, G. R., Yigit, E., Oyola, S. O., Langhorst, B. W., Schmidt, V. T., Stewart, F. J., ... Pradhan, S. (2013). A Method for

- Selectively Enriching Microbial DNA from Contaminating Vertebrate Host DNA. *PLoS ONE*, 8(10), e76096. <https://doi.org/10.1371/journal.pone.0076096>
- Felsenstein, J. (1989). PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics*, (5), 164–166.
- Fiévet, A., Bernard, V., Tenreiro, H., Dehainault, C., Girard, E., Deshaies, V., ... Houdayer, C. (2019). ART-DeCo: easy tool for detection and characterization of cross-contamination of DNA samples in diagnostic next-generation sequencing analysis. *European Journal of Human Genetics*, 27(5), 792–800. <https://doi.org/10.1038/s41431-018-0317-x>
- Fleischmann, R., Adams, M., White, O., Clayton, R., Kirkness, E., Kerlavage, A., ... Al, E. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*, 269(5223), 496–512. <https://doi.org/10.1126/science.7542800>
- Foster, G. W. (1979). Polychromophilus from Southeastern Brown Bats (*Myotis austroriparius*) in North-central Florida. *The Journal of Parasitology*, 65(3), 465–466. Abgerufen von <http://www.jstor.org/stable/3280300>
- Frénal, K., & Soldati-Favre, D. (2009). Role of the Parasite and Host Cytoskeleton in Apicomplexa Parasitism. *Cell Host & Microbe*, 5(6), 602–611. <https://doi.org/10.1016/j.chom.2009.05.013>
- Frénal, K., Dubremetz, J.-F., Lebrun, M., & Soldati-Favre, D. (2017). Gliding motility powers invasion and egress in Apicomplexa. *Nature Reviews Microbiology*, 15(11), 645–660. <https://doi.org/10.1038/nrmicro.2017.86>
- Galen, S. C., Borner, J., Martinsen, E. S., Schaer, J., Austin, C. C., West, C. J., & Perkins, S. L. (2018). The polyphyly of *Plasmodium*: comprehensive phylogenetic analyses of the malaria parasites (order Haemosporida) reveal widespread taxonomic conflict. *Royal Society open science*, 5(5), 171780. <https://doi.org/10.1098/rsos.171780>
- Galen, S. C., Borner, J., Williamson, J. L., Witt, C. C., & Perkins, S. L. (2019). Metatranscriptomics yields new genomic resources and sensitive detection of infections for diverse blood parasites. *Molecular Ecology Resources*, 1755–0998.13091. <https://doi.org/10.1111/1755-0998.13091>
- Galinski, M. R. (2019). Functional genomics of simian malaria parasites and host–parasite interactions. *Briefings in Functional Genomics*, 18(5), 270–280. <https://doi.org/10.1093/bfpg/ely013>
- Garcia, G. E., Wirtz, R. A., Barr, J. R., Woolfitt, A., & Rosenberg, R. (1998). Xanthurenic Acid Induces Gametogenesis in *Plasmodium*, the Malaria Parasite. *Journal of Biological Chemistry*, 273(20), 12003–12005. <https://doi.org/10.1074/jbc.273.20.12003>
- Garcia, J. E., Puentes, A., & Patarroyo, M. E. (2006). Developmental Biology of Sporozoite-Host Interactions in *Plasmodium falciparum* Malaria: Implications for Vaccine Design. *Clinical Microbiology Reviews*, 19(4), 686–707. <https://doi.org/10.1128/CMR.00063-05>
- Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., ... Barrell, B. (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*, 419(6906), 498–511. <https://doi.org/10.1038/nature01097>
- Gardner, R. A., & Molyneux, D. H. (1988). Polychromophilus murinus: a malarial parasite of bats: life-history and ultrastructural studies. *Parasitology*, 96 (Pt 3), 591–605. Abgerufen von <http://www.ncbi.nlm.nih.gov/pubmed/3136420>
- Gardner, R. A., Molyneux, D. H., & Stebbings, R. E. (1987). Studies on the prevalence of haematozoa of British bats. *Mammal Review*, 17(2–3), 75–80. <https://doi.org/10.1111/j.1365-2907.1987.tb00051.x>
- Garnham, P. C. C. (1948). The developmental cycle of *Hepatocystes (plasmodium) kochi* in the monkey host. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 41(5), 601–616. [https://doi.org/10.1016/S0035-9203\(48\)90418-0](https://doi.org/10.1016/S0035-9203(48)90418-0)
- Garnham, P. C. C. (1951). An attempt to find the vector of *Hepatocystis (=Plasmodium) kochi* (Levaditi and Schoen). *Experimental Parasitology*, 1(1), 94–107. [https://doi.org/10.1016/0014-4894\(51\)90010-0](https://doi.org/10.1016/0014-4894(51)90010-0)
- Garnham, P. C. C. (1973a). The zoogeography of *Polychromophilus* and description of a new species of a Gregarine (*Lankesteria galliardii*). *Annales de Parasitologie Humaine et Comparée*, 48(2), 231–242. <https://doi.org/10.1051/parasite/1973482231>
- Garnham, P. C. C. (1973b). Polychromophilus species in insectivorous bats. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 67(1), 2–3. [https://doi.org/10.1016/0035-9203\(73\)90253-8](https://doi.org/10.1016/0035-9203(73)90253-8)
- Garnham, P. C. C., Bird, R. G., & Baker, J. R. (1967). Electron microscope studies of motile stages of malaria parasites V. Exflagellation in *Plasmodium*, *Hepatocystis* and *Leucocytozoon*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 61(1), 58–68. [https://doi.org/10.1016/0035-9203\(67\)90054-5](https://doi.org/10.1016/0035-9203(67)90054-5)
- Garnham, P. C. C., & Heisch, R. B. (1953). On a new blood parasite of insectivorous bats. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 47(5), 357–363. [https://doi.org/10.1016/S0035-9203\(53\)80016-7](https://doi.org/10.1016/S0035-9203(53)80016-7)
- Garnham, P. C. C., Lainson, R., & Shaw, J. J. (1971). A contribution to the study of the haematozoon parasites of bats. A new mammalian haemoproteid, *Polychromophilus deanei* n. sp. *Memórias do Instituto Oswaldo Cruz*, 69(1), 119–125. <https://doi.org/10.1590/S0074-02761971000100009>
- Garten, M., Nasamu, A. S., Niles, J. C., Zimmerberg, J., Goldberg, D. E., & Beck, J. R. (2018). EXP2 is a nutrient-permeable channel in the vacuolar membrane of *Plasmodium* and is essential for protein export via PTEX. *Nature Microbiology*, 3(10), 1090–1098. <https://doi.org/10.1038/s41564-018-0222-7>
- Gasteiger, E. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, 31(13), 3784–3788. <https://doi.org/10.1093/nar/gkg563>

- Gaudin, M., & Desnues, C. (2018). Hybrid Capture-Based Next Generation Sequencing and Its Application to Human Infectious Diseases. *Frontiers in Microbiology*, 9. <https://doi.org/10.3389/fmicb.2018.02924>
- Giegerich, R., Meyer, F., & Schleiermacher, C. (1996). GeneFisher--software support for the detection of postulated genes. *Proceedings. International Conference on Intelligent Systems for Molecular Biology*, 4, 68–77. Abgerufen von <http://www.ncbi.nlm.nih.gov/pubmed/8877506>
- Gomes, P. S., Bhardwaj, J., Rivera-Correa, J., Freire-De-Lima, C. G., & Morrot, A. (2016). Immune Escape Strategies of Malaria Parasites. *Frontiers in Microbiology*, 7. <https://doi.org/10.3389/fmicb.2016.01617>
- Gordon, A., & Hannon, G. J. (2010). FASTX-Toolkit. Abgerufen 8. Februar 2017, von http://hannonlab.cshl.edu/fastx_toolkit
- Gouagna, L. C., Gouagna, L. C., Mulder, B., Mulder, B., Noubissi, E., Noubissi, E., ... Boudin, C. (1998). The early sporogonic cycle of *Plasmodium falciparum* in laboratory-infected *Anopheles gambiae*: an estimation of parasite efficacy. *Tropical Medicine & International Health*, 3(1), 21–28. <https://doi.org/10.1046/j.1365-3156.1998.00156.x>
- Gouin, A., Legeai, F., Nouhaud, P., Whibley, A., Simon, J.-C., & Lemaitre, C. (2015). Whole-genome re-sequencing of non-model organisms: lessons from unmapped reads. *Heredity*, 114(5), 494–501. <https://doi.org/10.1038/hdy.2014.85>
- Gouy, M., Guindon, S., & Gascuel, O. (2010). SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Molecular Biology and Evolution*, 27(2), 221–224. <https://doi.org/10.1093/molbev/msp259>
- Greenwood, J. M., Ezquerro, A. L., Behrens, S., Branca, A., & Mallet, L. (2016). Current analysis of host–parasite interactions with a focus on next generation sequencing data. *Zoology*, 119(4), 298–306. <https://doi.org/10.1016/j.zool.2016.06.010>
- Gregory, T. R., & Hebert, P. D. N. (1999). The Modulation of DNA Content: Proximate Causes and Ultimate Consequences. *Genome Research*, 9(9), 317–324. <https://doi.org/10.1101/gr.9.4.317>
- Gubbels, M.-J., & Duraisingh, M. T. (2012). Evolution of apicomplexan secretory organelles. *International Journal for Parasitology*, 42(12), 1071–1081. <https://doi.org/10.1016/j.ijpara.2012.09.009>
- Gueirard, P., Tavares, J., Thiberge, S., Bernex, F., Ishino, T., Milon, G., ... Amino, R. (2010). Development of the malaria parasite in the skin of the mammalian host. *Proceedings of the National Academy of Sciences*, 107(43), 18640–18645. <https://doi.org/10.1073/pnas.1009346107>
- Guindon, S., Lethiec, F., Duroux, P., & Gascuel, O. (2005). PHYML Online--a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Research*, 33(Web Server), W557–W559. <https://doi.org/10.1093/nar/gki352>
- Günther, T., & Nettelblad, C. (2019). The presence and impact of reference bias on population genomic studies of prehistoric human populations. *PLOS Genetics*, 15(7), e1008302. <https://doi.org/10.1371/journal.pgen.1008302>
- Gupta, A., Thiruvengadam, G., & Desai, S. A. (2015). The conserved clag multigene family of malaria parasites: Essential roles in host–pathogen interaction. *Drug Resistance Updates*, 18, 47–54. <https://doi.org/10.1016/j.drug.2014.10.004>
- Gužvić, M. (2013). The History of DNA Sequencing. *Journal of Medical Biochemistry*, 32(4), 301–312. <https://doi.org/10.2478/jomb-2014-0004>
- Haelewaters, D., Pfliegler, W. P., Szentiványi, T., Földvári, M., Sándor, A. D., Barti, L., ... Pfister, D. H. (2017). Parasites of parasites of bats: Laboulbeniales (Fungi: Ascomycota) on bat flies (Diptera: Nycteribiidae) in central Europe. *Parasites & Vectors*, 10(1), 96. <https://doi.org/10.1186/s13071-017-2022-y>
- Hall, T. A. (1999). BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, (41), 95–98.
- Haltom, A. R., & Jafar-Nejad, H. (2015). The multiple roles of epidermal growth factor repeat O -glycans in animal development. *Glycobiology*, 25(10), 1027–1042. <https://doi.org/10.1093/glycob/cwv052>
- Hamilton, W. L., Claessens, A., Otto, T. D., Kekre, M., Fairhurst, R. M., Rayner, J. C., & Kwiatkowski, D. (2016). Extreme mutation bias and high AT content in *Plasmodium falciparum*. *Nucleic Acids Research*, gkw1259. <https://doi.org/10.1093/nar/gkw1259>
- Hansen, K. D., Brenner, S. E., & Dudoit, S. (2010). Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Research*, 38(12), e131–e131. <https://doi.org/10.1093/nar/gkq224>
- Hayman, D. T. S. (2016). Bats as Viral Reservoirs. *Annual review of virology*, 3(1), 77–99. <https://doi.org/10.1146/annurev-virology-110615-042203>
- Henry, N. B., Sermé, S. S., Siciliano, G., Sombié, S., Diarra, A., Sagnon, N., ... Alano, P. (2019). Biology of *Plasmodium falciparum* gametocyte sex ratio and implications in malaria parasite transmission. *Malaria Journal*, 18(70). <https://doi.org/10.1186/s12936-019-2707-0>
- Hiller, N. L., Bhattacharjee, S., van Ooij, C., Liolios, K., Harrison, T., Lopez-Estraño, C., & Haldar, K. (2004). A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science*, 306(5703), 1934–1937. <https://doi.org/10.1126/science.1102737>
- Holz, P. H., Lumsden, L. F., Legione, A. R., & Hufschmid, J. (2019). Polychromophilus melanipherus and haemoplasma infections not associated with clinical signs in southern bent-winged bats (*Miniopterus orianae bassanii*) and eastern bent-winged bats (*Miniopterus orianae oceanensis*). *International Journal for Parasitology: Parasites and Wildlife*, 8, 10–18. <https://doi.org/10.1016/j.ijppaw.2018.11.008>

- Homer, M. J., Aguilar-Delfin, I., Telford, S. R., Krause, P. J., & Persing, D. H. (2000). Babesiosis. *Clinical Microbiology Reviews*, 13(3), 451–469. <https://doi.org/10.1128/CMR.13.3.451-469.2000>
- Huang, X., Hansson, R., Palinauskas, V., Valkiūnas, G., Hellgren, O., & Bensch, S. (2018). The success of sequence capture in relation to phylogenetic distance from a reference genome: a case study of avian haemosporidian parasites. *International Journal for Parasitology*, 48(12), 947–954. <https://doi.org/10.1016/j.ijpara.2018.05.009>
- Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, 17(8), 754–755. <https://doi.org/10.1093/bioinformatics/17.8.754>
- Humphrys, M. S., Creasy, T., Sun, Y., Shetty, A. C., Chibucos, M. C., Drabek, E. F., ... Myers, G. S. A. (2013). Simultaneous Transcriptional Profiling of Bacteria and Their Host Cells. *PLoS ONE*, 8(12), e80597. <https://doi.org/10.1371/journal.pone.0080597>
- Hunt, M., Kikuchi, T., Sanders, M., Newbold, C., Berriman, M., & Otto, T. D. (2013). REAPR: a universal tool for genome assembly evaluation. *Genome Biology*, 14(5), R47. <https://doi.org/10.1186/gb-2013-14-5-r47>
- Hupalo, D. N., Bradic, M., & Carlton, J. M.-R. (2015). The impact of genomics on population genetics of parasitic diseases. *Current Opinion in Microbiology*, 23, 49–54. <https://doi.org/10.1016/j.mib.2014.11.001>
- Hutson, A. M. (1984). *Keds, Flat-flies and Bat-flies (Diptera, Hippoboscidae and Nycteribiidae) Handbooks for the Identification of British Insects*. (10(7); M. G. Fitton, Hrsg.). London: Royal Entomological Society of London.
- Ishino, T., Boisson, B., Orito, Y., Lacroix, C., Bischoff, E., Loussert, C., ... Baldacci, P. (2009). LISP1 is important for the egress of Plasmodium berghei parasites from liver cells. *Cellular Microbiology*, 11(9), 1329–1339. <https://doi.org/10.1111/j.1462-5822.2009.01333.x>
- Iwanaga, S., Kaneko, I., Kato, T., & Yuda, M. (2012). Identification of an AP2-family Protein That Is Critical for Malaria Liver Stage Development. *PLoS ONE*, 7(11), e47557. <https://doi.org/10.1371/journal.pone.0047557>
- Iyer, J., Grüner, A. C., Rénia, L., Snounou, G., & Preiser, P. R. (2007). Invasion of host cells by malaria parasites: a tale of two protein families. *Molecular Microbiology*, 65(2), 231–249. <https://doi.org/10.1111/j.1365-2958.2007.05791.x>
- Jacob, J. T., Coulombe, P. A., Kwan, R., & Omary, M. B. (2018). Types I and II Keratin Intermediate Filaments. *Cold Spring Harbor Perspectives in Biology*, 10(4), a018275. <https://doi.org/10.1101/cshperspect.a018275>
- Jalovecka, M., Hajdusek, O., Sojka, D., Kopacek, P., & Malandrin, L. (2018). The Complexity of Piroplasms Life Cycles. *Frontiers in Cellular and Infection Microbiology*, 8. <https://doi.org/10.3389/fcimb.2018.00248>
- Jia, H., Guo, Y., Zhao, W., & Wang, K. (2015). Long-range PCR in next-generation sequencing: comparison of six enzymes and evaluation on the MiSeq sequencer. *Scientific Reports*, 4(1), 5737. <https://doi.org/10.1038/srep05737>
- Jimah, J. R., Salinas, N. D., Sala-Rabanal, M., Jones, N. G., Sibley, L. D., Nichols, C. G., ... Tolia, N. H. (2016). Malaria parasite CelTOS targets the inner leaflet of cell membranes for pore-dependent disruption. *eLife*, 5. <https://doi.org/10.7554/eLife.20621>
- Jimenez-Ruiz, E., Morlon-Guyot, J., Daher, W., & Meissner, M. (2016). Vacuolar protein sorting mechanisms in apicomplexan parasites. *Molecular and Biochemical Parasitology*, 209(1–2), 18–25. <https://doi.org/10.1016/j.molbiopara.2016.01.007>
- Jofuku, K. D., den Boer, B. G., Van Montagu, M., & Okamoto, J. K. (1994). Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. *The Plant Cell*, 6(9), 1211–1225. <https://doi.org/10.1105/tpc.6.9.1211>
- Johnston, D. A., Blaxter, M. L., Degraeve, W. M., Foster, J., Ivens, A. C., & Melville, S. E. (1999). Genomics and the biology of parasites. *BioEssays*, 21(2), 131–147. [https://doi.org/10.1002/\(SICI\)1521-1878\(199902\)21:2<131::AID-BIES7>3.0.CO;2-I](https://doi.org/10.1002/(SICI)1521-1878(199902)21:2<131::AID-BIES7>3.0.CO;2-I)
- Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., ... Hunter, S. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics*, 30(9), 1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>
- Joop, G., & Vilcinskas, A. (2016). Coevolution of parasitic fungi and insect hosts. *Zoology*, 119(4), 350–358. <https://doi.org/10.1016/j.zool.2016.06.005>
- Kafsack, B. F. C., Rovira-Graells, N., Clark, T. G., Bancells, C., Crowley, V. M., Campino, S. G., ... Llinás, M. (2014). A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature*, 507(7491), 248–252. <https://doi.org/10.1038/nature12920>
- Käll, L., Krogh, A., & Sonnhammer, E. L. (2004). A Combined Transmembrane Topology and Signal Peptide Prediction Method. *Journal of Molecular Biology*, 338(5), 1027–1036. <https://doi.org/10.1016/j.jmb.2004.03.016>
- Kaltungo, B. Y., & Musa, I. W. (2013). A Review of Some Protozoan Parasites Causing Infertility in Farm Animals. *ISRN Tropical Medicine*, 2013, 1–6. <https://doi.org/10.1155/2013/782609>
- Karadjian, G., Hassanin, A., Saintpierre, B., Gembu Tung'aluna, G.-C., Aricy, F., Ayala, F. J., ... Duval, L. (2016). Highly rearranged mitochondrial genome in Nycteria parasites (Haemosporidia) from bats. *Proceedings of the National Academy of Sciences*, 113(35), 9834–9839. <https://doi.org/10.1073/pnas.1610643113>
- Kasai, F., O'Brien, P. C. M., & Ferguson-Smith, M. A. (2013). The bat genome: GC-biased small chromosomes associated with reduction in genome size. *Chromosoma*, 122(6), 535–540. <https://doi.org/10.1007/s00412-013-0426-9>
- Katoh, K. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids*

- Research*, 30(14), 3059–3066. <https://doi.org/10.1093/nar/gkf436>
- Kats, L. M., Black, C. G., Proellocks, N. I., & Coppel, R. L. (2006). Plasmodium rhoptries: how things went pear-shaped. *Trends in Parasitology*, 22(6), 269–276. <https://doi.org/10.1016/j.pt.2006.04.001>
- Kaushansky, A., Douglass, A. N., Arang, N., Vigdorovich, V., Dambrauskas, N., Kain, H. S., ... Kappe, S. H. I. (2015). Malaria parasites target the hepatocyte receptor EphA2 for successful host infection. *Science*, 350(6264), 1089–1092. <https://doi.org/10.1126/science.aad3318>
- Keeling, P. J. (2004). Reduction and Compaction in the Genome of the Apicomplexan Parasite *Cryptosporidium parvum*. *Developmental Cell*, 6(5), 614–616. [https://doi.org/10.1016/S1534-5807\(04\)00135-2](https://doi.org/10.1016/S1534-5807(04)00135-2)
- Kennedy, M., Fishbaugher, M. E., Vaughan, A. M., Patrapuvich, R., Boonhok, R., Yimamnuaychok, N., ... Lindner, S. E. (2012). A rapid and scalable density gradient purification method for Plasmodium sporozoites. *Malaria Journal*, 11(1), 421. <https://doi.org/10.1186/1475-2875-11-421>
- Kibbe, W. A. (2007). OligoCalc: an online oligonucleotide properties calculator. *Nucleic acids research*, 35(Web Server issue), W43-6. <https://doi.org/10.1093/nar/gkm234>
- Kim, K. (2004). Role of proteases in host cell invasion by *Toxoplasma gondii* and other Apicomplexa. *Acta tropica*, 91(1), 69–81. <https://doi.org/10.1016/j.actatropica.2003.11.016>
- Kim, Y.-K., Park, S.-G., Kim, T.-W., Park, J.-H., Adhikari, P., Kim, G., ... Oh, H.-S. (2017). Complete mitochondrial genome of the far eastern *Myotis*, *Myotis bombinus* (Chiroptera, Vespertilionidae). *Mitochondrial DNA Part A*, 28(2), 267–268. <https://doi.org/10.3109/19401736.2015.1118072>
- Kissinger, J. C., & DeBarry, J. (2011). Genome cartography: charting the apicomplexan genome. *Trends in Parasitology*, 27(8), 345–354. <https://doi.org/10.1016/j.pt.2011.03.006>
- Kissinger, J. C., Souza, P. C. A., Soarest, C. O., Paul, R., Wahl, A. M., Rathore, D., ... Kretzli, A. U. (2002). Molecular phylogenetic analysis of the avian malarial parasite *Plasmodium* (*Novyella*) *juxtannucleare*. *The Journal of parasitology*, 88(4), 769–773. [https://doi.org/10.1645/0022-3395\(2002\)088\[0769:MPAOTA\]2.0.CO;2](https://doi.org/10.1645/0022-3395(2002)088[0769:MPAOTA]2.0.CO;2)
- Klug, D., & Frischknecht, F. (2017). Motility precedes egress of malaria parasites from oocysts. *eLife*, 6. <https://doi.org/10.7554/eLife.19157>
- Koch, M., & Baum, J. (2016). The mechanics of malaria parasite invasion of the human erythrocyte – towards a reassessment of the host cell contribution. *Cellular Microbiology*, 18(3), 319–329. <https://doi.org/10.1111/cmi.12557>
- Korf, I. (2004). Gene finding in novel genomes. *BMC bioinformatics*, 5, 59. <https://doi.org/10.1186/1471-2105-5-59>
- Kryukov, K., & Imanishi, T. (2016). Human Contamination in Public Genome Assemblies. *PLOS ONE*, 11(9), e0162424. <https://doi.org/10.1371/journal.pone.0162424>
- Kuhl, H. (1817). Die Daubenton'sche Fledermaus. *Vespertilio Daubentonii* Leisler. In *Die deutschen Fledermäuse* (S. 51–52). Hanau.
- Kumar, S., Jones, M., Koutsovoulos, G., Clarke, M., & Blaxter, M. (2013). Blobology: exploring raw genome data for contaminants, symbionts and parasites using taxon-annotated GC-coverage plots. *Frontiers in Genetics*, 4. <https://doi.org/10.3389/fgene.2013.00237>
- Kuo, C.-H., Wares, J. P., & Kissinger, J. C. (2008). The Apicomplexan Whole-Genome Phylogeny: An Analysis of Incongruence among Gene Trees. *Molecular Biology and Evolution*, 25(12), 2689–2698. <https://doi.org/10.1093/molbev/msn213>
- Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., & Salzberg, S. L. (2004). Versatile and open software for comparing large genomes. *Genome biology*, 5(2), R12. <https://doi.org/10.1186/gb-2004-5-2-r12>
- Laetsch, D. R., & Blaxter, M. L. (2017). BlobTools: Interrogation of genome assemblies. *F1000Research*, 6, 1287. <https://doi.org/10.12688/f1000research.12232.1>
- Lage, J. M., Leamon, J. H., Pejovic, T., Hamann, S., Lacey, M., Dillon, D., ... Lizardi, P. M. (2003). Whole genome analysis of genetic alterations in small DNA samples using hyperbranched strand displacement amplification and array-CGH. *Genome research*, 13(2), 294–307. <https://doi.org/10.1101/gr.377203>
- Lagesen, K., Hallin, P., Rødland, E. A., Stærfeldt, H.-H., Rognes, T., & Ussery, D. W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research*, 35(9), 3100–3108. <https://doi.org/10.1093/nar/gkm160>
- Landau, I., Chabaud, A. G., Miltgen, F., & Baccam, D. (1980). [Dionisia bunoi n. g. n. sp., Haemoproteidae parasite of the microchiropteran bat *Hipposideros cyclops* in Gabon (author's transl)]. *Annales de parasitologie humaine et comparée*, 55(3), 271–280. <https://doi.org/10.1051/parasite/1980553271>
- Landau, I., Rosin, G., Miltgen, F., Hugot, J.-P., Leger, N., Beveridge, I., & Baccam, D. (1980). Sur le genre *Polycbromophilus*. *Annales de Parasitologie Humaine et Comparée*, 55(1), 13–32. <https://doi.org/10.1051/parasite/1980551013>
- Landau, I., Chavatte, J. M., Karadjian, G., Chabaud, A. G., & Beveridge, I. (2012). The haemosporidian parasites of bats with description of *Sprattiella alecto* gen. nov., sp. nov. *Parasite*, 19(2), 137–146. <https://doi.org/10.1051/parasite/2012192137>
- Landau, I., Chavatte, J. M., & Beveridge, I. (2012). *Johnsprentia copemani* gen. nov., sp. nov. (Haemoproteidae), a parasite of the Flying-fox, *Pteropus alecto* (Pteropodidae) from Queensland. *Memoirs of the Queensland Museum*, 56, 61–66.
- Landau, I., & Chabaud, A. G. (1978). Description de *Plasmodium cyclopsi* n. sp. parasite du Microchiroptère *Hipposideros cyclops*

- à Makokou (Gabon). *Annales de Parasitologie Humaine et Comparée*, 53(3), 247–253. <https://doi.org/10.1051/parasite/1978533247>
- Laslett, D. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Research*, 32(1), 11–16. <https://doi.org/10.1093/nar/gkh152>
- Lauer, S. A. (1997). A Membrane Network for Nutrient Import in Red Cells Infected with the Malaria Parasite. *Science*, 276(5315), 1122–1125. <https://doi.org/10.1126/science.276.5315.1122>
- Lauer, S. A., VanWye, J., Harrison, T., McManus, H., Samuel, B. U., Hiller, N. L., ... Haldar, K. (2000). Vacuolar uptake of host components, and a role for cholesterol and sphingomyelin in malarial infection. *The EMBO Journal*, 19(14), 3556–3564. <https://doi.org/10.1093/emboj/19.14.3556>
- Laurence, M., Hatzis, C., & Brash, D. E. (2014). Common Contaminants in Next-Generation Sequencing That Hinder Discovery of Low-Abundance Microbes. *PLoS ONE*, 9(5), e97876. <https://doi.org/10.1371/journal.pone.0097876>
- Le Roch, G., K., Chung, D.-W. D., & Ponts, N. (2012). Genomics and integrated systems biology in *Plasmodium falciparum*: a path to malaria control and eradication. *Parasite Immunology*, 34(2–3), 50–60. <https://doi.org/10.1111/j.1365-3024.2011.01340.x>
- Leichty, A. R., & Brisson, D. (2014). Selective Whole Genome Amplification for Resequencing Target Microbial Species from Complex Natural Samples. *Genetics*, 198(2), 473–481. <https://doi.org/10.1534/genetics.114.165498>
- Lemmon, A. R., Emme, S. A., & Lemmon, E. M. (2012). Anchored Hybrid Enrichment for Massively High-Throughput Phylogenomics. *Systematic Biology*, 61(5), 727–744. <https://doi.org/10.1093/sysbio/sys049>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li, L. (2003). OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Research*, 13(9), 2178–2189. <https://doi.org/10.1101/gr.1224503>
- Liu, W., Sundararaman, S. A., Loy, D. E., Learn, G. H., Li, Y., Plenderleith, L. J., ... Sharp, P. M. (2016). Multigenomic Delineation of *Plasmodium* Species of the *Laverania* Subgenus Infecting Wild-Living Chimpanzees and Gorillas. *Genome Biology and Evolution*, 8(6), 1929–1939. <https://doi.org/10.1093/gbe/evw128>
- Lochmiller, R. L., & Deerenberg, C. (2000). Trade-offs in evolutionary immunology: just what is the cost of immunity? *Oikos*, 88(1), 87–98. <https://doi.org/10.1034/j.1600-0706.2000.880110.x>
- Long, M., Betrán, E., Thornton, K., & Wang, W. (2003). The origin of new genes: glimpses from the young and old. *Nature Reviews Genetics*, 4(11), 865–875. <https://doi.org/10.1038/nrg1204>
- López, C., Saravia, C., Gomez, A., Hoebeke, J., & Patarroyo, M. A. (2010). Mechanisms of genetically-based resistance to malaria. *Gene*, 467(1–2), 1–12. <https://doi.org/10.1016/j.gene.2010.07.008>
- Lourenço, S., & Palmeirim, J. M. (2008). Which factors regulate the reproduction of ectoparasites of temperate-zone cave-dwelling bats? *Parasitology Research*, 104(1), 127–134. <https://doi.org/10.1007/s00436-008-1170-6>
- Lowe, T. M., & Chan, P. P. (2016). tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Research*, 44(W1), W54–W57. <https://doi.org/10.1093/nar/gkw413>
- Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., ... Murray, C. J. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, 380(9859), 2095–2128. [https://doi.org/10.1016/S0140-6736\(12\)61728-0](https://doi.org/10.1016/S0140-6736(12)61728-0)
- Lučan, R. K., & Hanák, V. (2011). Population ecology of *Myotis daubentonii* (Mammalia: Chiroptera) in the South Bohemia: summary of two long-term studies: 1968–1984 and 1999–2009. *Acta Societatis Zoologicae Bohemicae*, 75, 67–85.
- Lutz, H. L., Patterson, B. D., Kerbis Peterhans, J. C., Stanley, W. T., Webala, P. W., Gnoske, T. P., ... Stanhope, M. J. (2016). Diverse sampling of East African haemosporidians reveals chiropteran origin of malaria parasites in primates and rodents. *Molecular Phylogenetics and Evolution*, 99, 7–15. <https://doi.org/10.1016/j.ympev.2016.03.004>
- Magadum, S., Banerjee, U., Murugan, P., Gangapur, D., & Ravikesavan, R. (2013). Gene duplication as a major force in evolution. *Journal of Genetics*, 92(1), 155–161. <https://doi.org/10.1007/s12041-013-0212-8>
- Mamanova, L., Coffey, A. J., Scott, C. E., Kozarewa, I., Turner, E. H., Kumar, A., ... Turner, D. J. (2010). Target-enrichment strategies for next-generation sequencing. *Nature Methods*, 7(2), 111–118. <https://doi.org/10.1038/nmeth.1419>
- Mandl, J. N., Schneider, C., Schneider, D. S., & Baker, M. L. (2018). Going to Bat(s) for Studies of Disease Tolerance. *Frontiers in Immunology*, 9. <https://doi.org/10.3389/fimmu.2018.02112>
- Manwell, R. D. (1946). Bat Malaria. *American Journal of Epidemiology*, 43(1), 1–12. <https://doi.org/10.1093/oxfordjournals.aje.a119047>
- Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., ... Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057), 376–380. <https://doi.org/10.1038/nature03959>
- Marinkelle, C. J. (1995). The haemoproteid parasite, *Bioccala deanei*, from a Colombian bat, *Eptesicus fuscus* (Vespertilionidae). *Annals of Tropical Medicine & Parasitology*, 89(1), 89–91. <https://doi.org/10.1080/00034983.1995.11812935>

- Markov, A. V., Anisimov, V. A., & Korotayev, A. V. (2010). Relationship between genome size and organismal complexity in the lineage leading from prokaryotes to mammals. *Paleontological Journal*, 44(4), 363–373. <https://doi.org/10.1134/S0031030110040015>
- Marsh, J. W., Hayward, R. J., Shetty, A. C., Mahurkar, A., Humphrys, M. S., & Myers, G. S. A. (2017). Bioinformatic analysis of bacteria and host cell dual RNA-sequencing experiments. *Briefings in Bioinformatics*. <https://doi.org/10.1093/bib/bbx043>
- Marti, M., Good, R. T., Rug, M., Knuepfer, E., & Cowman, A. F. (2004). Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science*, 306(5703), 1930–1933. <https://doi.org/10.1126/science.1102452>
- Martínez-Ocampo, F. (2018). Genomics of Apicomplexa. In *Farm Animals Diseases, Recent Omic Trends and New Strategies of Treatment*. <https://doi.org/10.5772/intechopen.72633>
- Martinsen, E. S., McInerney, N., Brightman, H., Ferebee, K., Walsh, T., McShea, W. J., ... Fleischer, R. C. (2016). Hidden in plain sight: Cryptic and endemic malaria parasites in North American white-tailed deer (*Odocoileus virginianus*). *Science Advances*, 2(2), e1501486. <https://doi.org/10.1126/sciadv.1501486>
- Martinsen, E. S., & Perkins, S. L. (2013). The diversity of Plasmodium and other Haemosporidians: The intersection of taxonomy, phylogenetics and genomics. In J. M. Carlton, S. L. Perkins, & K. W. Deitsch (Hrsg.), *Malaria parasites: Comparative genomics, evolution and molecular biology* (S. 1–15). Norfolk, UK: Caister Academic Press.
- Martinsen, E. S., Perkins, S. L., & Schall, J. J. (2008). A three-genome phylogeny of malaria parasites (Plasmodium and closely related genera): Evolution of life-history traits and host switches. *Molecular Phylogenetics and Evolution*, 47(1), 261–273. <https://doi.org/10.1016/j.ympev.2007.11.012>
- Marzal, A., Ricklefs, R. E., Valkiūnas, G., Albayrak, T., Arriero, E., Bonneaud, C., ... Bensch, S. (2011). Diversity, Loss, and Gain of Malaria Parasites in a Globally Invasive Bird. *PLoS ONE*, 6(7), e21905. <https://doi.org/10.1371/journal.pone.0021905>
- Mascarelli, P. E., Tartara, G. P., Pereyra, N. B., & Maggi, R. G. (2016). Detection of Mycoplasma haemocanis, Mycoplasma haematoparvum, Mycoplasma suis and other vector-borne pathogens in dogs from Córdoba and Santa Fé, Argentina. *Parasites & Vectors*, 9(1), 642. <https://doi.org/10.1186/s13071-016-1920-8>
- Matthews, K., Kalanon, M., Chisholm, S. A., Sturm, A., Goodman, C. D., Dixon, M. W. A., ... de Koning-Ward, T. F. (2013). The Plasmodium translocon of exported proteins (PTEx) component thioredoxin-2 is important for maintaining normal blood-stage growth. *Molecular Microbiology*, 89(6), 1167–1186. <https://doi.org/10.1111/mmi.12334>
- Matveev, V. A., Kruskop, S. V., & Kramerov, D. A. (2005). Revalidation of Myotis petax Hollister, 1912 and its new status in connection with M. daubentonii (Kuhl, 1817) (Vespertilionidae, Chiroptera). *Acta Chiropterologica*, 7(1), 23–37. [https://doi.org/https://doi.org/10.3161/1733-5329\(2005\)7\[23:ROMPHA\]2.0.CO;2](https://doi.org/https://doi.org/10.3161/1733-5329(2005)7[23:ROMPHA]2.0.CO;2)
- Matz, J. M., & Matuschewski, K. (2018). An in silico down-scaling approach uncovers novel constituents of the Plasmodium-containing vacuole. *Scientific Reports*, 8(1), 14055. <https://doi.org/10.1038/s41598-018-32471-6>
- McIntosh, M. (1998). Divergent evolutionary constraints on mitochondrial and nuclear genomes of malaria parasites. *Molecular and Biochemical Parasitology*, 95(1), 69–80. [https://doi.org/10.1016/S0166-6851\(98\)00093-0](https://doi.org/10.1016/S0166-6851(98)00093-0)
- Megali, A., Yannic, G., & Christe, P. (2011). Disease in the dark: Molecular characterization of Polychromophilus murinus in temperate zone bats revealed a worldwide distribution of this malaria-like disease. *Molecular Ecology*, 20(5), 1039–1048. <https://doi.org/10.1111/j.1365-294X.2010.04905.x>
- Melnikov, A., Galinsky, K., Rogov, P., Fennell, T., Tyne, D., Russ, C., ... Neafsey, D. E. (2011). Hybrid selection for sequencing pathogen genomes from clinical samples. *Genome Biology*, 12(8), R73. <https://doi.org/10.1186/gb-2011-12-8-r73>
- Melum, E., May, S., Schilhabel, M. B., Thomsen, I., Karlsen, T. H., Rosenstiel, P., ... Franke, A. (2010). SNP discovery performance of two second-generation sequencing platforms in the NOD2 gene region. *Human Mutation*, 31(7), 875–885. <https://doi.org/10.1002/humu.21276>
- Mer, G. G., & Goldblum, N. (1947). A Haemosporidian of Bats. *Nature*, 159(4039), 444–444. <https://doi.org/10.1038/159444a0>
- Merchant, S., Wood, D. E., & Salzberg, S. L. (2014). Unexpected cross-species contamination in genome sequencing projects. *PeerJ*, 2, e675. <https://doi.org/10.7717/peerj.675>
- Mercier, C., Adjogble, K. D. Z., Däubener, W., & Delauw, M.-F.-C. (2005). Dense granules: Are they key organelles to help understand the parasitophorous vacuole of all apicomplexa parasites? *International Journal for Parasitology*, 35(8), 829–849. <https://doi.org/10.1016/j.ijpara.2005.03.011>
- Mi, H., Muruganujan, A., Ebert, D., Huang, X., & Thomas, P. D. (2019). PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Research*, 47(D1), D419–D426. <https://doi.org/10.1093/nar/gky1038>
- Mideo, N. (2009). Parasite adaptations to within-host competition. *Trends in Parasitology*, 25(6), 261–268. <https://doi.org/10.1016/j.pt.2009.03.001>
- Miller-Butterworth, C. M., Murphy, W. J., O'Brien, S. J., Jacobs, D. S., Springer, M. S., & Teeling, E. C. (2007). A Family Matter: Conclusive Resolution of the Taxonomic Position of the Long-Fingered Bats, Miniopterus. *Molecular Biology and Evolution*, 24(7), 1553–1561. <https://doi.org/10.1093/molbev/msm076>
- Miller, M. R., White, A., & Boots, M. (2007). The evolution of parasites in response to tolerance in their hosts: the good, the bad, and apparent commensalism. *Evolution*, 60(5), 945–956. <https://doi.org/10.1111/j.0014-3820.2006.tb01173.x>

- Moll, R., Divo, M., & Langbein, L. (2008). The human keratins: biology and pathology. *Histochemistry and Cell Biology*, 129(6), 705–733. <https://doi.org/10.1007/s00418-008-0435-6>
- Molyneux, D. H. (2012). Blood Parasites of bats, birds and mice-biodiversity in blood. Abgerufen 6. August 2019, von http://runemerge.univ-reunion.fr/fileadmin/Fichiers/RUNEMERGE/exchange_fellowships/La_Reunion_lecture_FINALDHM_march_16th.pdf
- Montoya, J., & Liesenfeld, O. (2004). Toxoplasmosis. *The Lancet*, 363(9425), 1965–1976. [https://doi.org/10.1016/S0140-6736\(04\)16412-X](https://doi.org/10.1016/S0140-6736(04)16412-X)
- Moran, N. A. (2002). Microbial Minimalism. *Cell*, 108(5), 583–586. [https://doi.org/10.1016/S0092-8674\(02\)00665-7](https://doi.org/10.1016/S0092-8674(02)00665-7)
- Morita, M., Nagaoka, H., Ntege, E. H., Kanoi, B. N., Ito, D., Nakata, T., ... Takashima, E. (2018). PV1, a novel Plasmodium falciparum merozoite dense granule protein, interacts with exported protein in infected erythrocytes. *Scientific Reports*, 8(1), 3696. <https://doi.org/10.1038/s41598-018-22026-0>
- Morrison, D. A. (2009). Evolution of the Apicomplexa: where are we now? *Trends in Parasitology*, 25(8), 375–382. <https://doi.org/10.1016/j.pt.2009.05.010>
- Morrisette, N. S., & Sibley, L. D. (2002). Cytoskeleton of Apicomplexan Parasites. *Microbiology and Molecular Biology Reviews*, 66(1), 21–38. <https://doi.org/10.1128/MMBR.66.1.21-38.2002>
- Mu, A., Kwong, J. C., Isles, N. S., Gonçalves da Silva, A., Schultz, M. B., Ballard, S. A., ... Howden, B. P. (2019). Reconstruction of the Genomes of Drug-Resistant Pathogens for Outbreak Investigation through Metagenomic Sequencing. *mSphere*, 4(1). <https://doi.org/10.1128/mSphere.00529-18>
- Naitza, S., Spano, F., Robson, K. J., & Crisanti, A. (1998). The Thrombospondin-related Protein Family of Apicomplexan Parasites: The Gears of the Cell Invasion Machinery. *Parasitology Today*, 14(12), 479–484. [https://doi.org/10.1016/S0169-4758\(98\)01346-5](https://doi.org/10.1016/S0169-4758(98)01346-5)
- Neal, A. T. (2011). Male gametocyte fecundity and sex ratio of a malaria parasite, Plasmodium mexicanum. *Parasitology*, 138(10), 1203–1210. <https://doi.org/10.1017/S0031182011000941>
- Nessel, T., Beck, J. M., Rayatpisheh, S., Jami-Alahmadi, Y., Wohlschlegel, J. A., Goldberg, D. E., & Beck, J. R. (2020). EXP1 is required for organisation of EXP2 in the intraerythrocytic malaria parasite vacuole. *Cellular Microbiology*. <https://doi.org/10.1111/cmi.13168>
- Ngotho, P., Soares, A. B., Hentzschel, F., Achcar, F., Bertuccini, L., & Marti, M. (2019). Revisiting gametocyte biology in malaria parasites. *FEMS Microbiology Reviews*, 43(4), 401–414. <https://doi.org/10.1093/femsre/fuz010>
- Ngwa, C. J., Scheuermayer, M., Mair, G. R., Kern, S., Brühl, T., Wirth, C. C., ... Pradel, G. (2013). Changes in the transcriptome of the malaria parasite Plasmodium falciparum during the initial phase of transmission from the human to the mosquito. *BMC Genomics*, 14(1), 256. <https://doi.org/10.1186/1471-2164-14-256>
- Nicholas, K. B., & Nicholas, H. B. J. (1997). GeneDoc: Analysis and Visualization of Genetic Variation. *EMBNET NEWS*, 4(414).
- Nikbakht, H., Xia, X., Hickey, D. A., & Golding, B. (2014). The evolution of genomic GC content undergoes a rapid reversal within the genus Plasmodium. *Genome*, 57(9), 507–511. <https://doi.org/10.1139/gen-2014-0158>
- O'Connell, J., Schulz-Trieglaff, O., Carlson, E., Hims, M. M., Gormley, N. A., & Cox, A. J. (2015). NxTrim: optimized trimming of Illumina mate pair reads: Table 1. *Bioinformatics*, 31(12), 2035–2037. <https://doi.org/10.1093/bioinformatics/btv057>
- O'Donoghue, P. (2017). Haemoprotozoa: Making biological sense of molecular phylogenies. *International journal for parasitology. Parasites and wildlife*, 6(3), 241–256. <https://doi.org/10.1016/j.ijppaw.2017.08.007>
- Orosz, F. (2015). Two recently sequenced vertebrate genomes are contaminated with apicomplexan species of the Sarcocystidae family. *International Journal for Parasitology*, 45(13), 871–878. <https://doi.org/10.1016/j.ijpara.2015.07.002>
- Orr, T. J., & Zuk, M. (2014). Reproductive delays in mammals: an unexplored avenue for post-copulatory sexual selection. *Biological Reviews*, 89(4), 889–912. <https://doi.org/10.1111/brv.12085>
- Otto, T. D., Rayner, J. C., Böhme, U., Pain, A., Spottiswoode, N., Sanders, M., ... Berriman, M. (2014). Genome sequencing of chimpanzee malaria parasites reveals possible pathways of adaptation to human hosts. *Nature Communications*, 5(1), 4754. <https://doi.org/10.1038/ncomms5754>
- Otto, T. D., Böhme, U., Jackson, A. P., Hunt, M., Franke-Fayard, B., Hoeymakers, W. A. M., ... Janse, C. J. (2014). A comprehensive evaluation of rodent malaria parasite genomes and gene expression. *BMC Biology*, 12(1), 86. <https://doi.org/10.1186/s12915-014-0086-0>
- Otto, T. D., Gilabert, A., Crellen, T., Böhme, U., Arnathau, C., Sanders, M., ... Prugnolle, F. (2018). Genomes of all known members of a Plasmodium subgenus reveal paths to virulent human malaria. *Nature Microbiology*, 3(6), 687–697. <https://doi.org/10.1038/s41564-018-0162-2>
- Outlaw, D. C., & Ricklefs, R. E. (2011). Rerooting the evolutionary tree of malaria parasites. *Proceedings of the National Academy of Sciences*, 108(32), 13183–13187. <https://doi.org/10.1073/pnas.1109153108>
- Outlaw, D. C., & Ricklefs, R. E. (2014). Species limits in avian malaria parasites (Haemosporida): how to move forward in the molecular era. *Parasitology*, 141(10), 1223–1232. <https://doi.org/10.1017/S0031182014000560>

- Oyola, S. O., Manske, M., Campino, S., Claessens, A., Hamilton, W. L., Kekre, M., ... Kwiatkowski, D. P. (2014). Optimized Whole-Genome Amplification Strategy for Extremely AT-Biased Template. *DNA Research*, 21(6), 661–671. <https://doi.org/10.1093/dnares/dsu028>
- Oyola, S. O., Ariani, C. V., Hamilton, W. L., Kekre, M., Amenga-Etego, L. N., Ghansah, A., ... Kwiatkowski, D. P. (2016). Whole genome sequencing of *Plasmodium falciparum* from dried blood spots using selective whole genome amplification. *Malaria Journal*, 15(1), 597. <https://doi.org/10.1186/s12936-016-1641-7>
- Oyola, S. O., Gu, Y., Manske, M., Otto, T. D., O'Brien, J., Alcock, D., ... Quail, M. A. (2013). Efficient depletion of host DNA contamination in malaria clinical sequencing. *Journal of Clinical Microbiology*, 51(3), 745–751. <https://doi.org/10.1128/JCM.02507-12>
- Ozcelik, H., Shi, X., Chang, M. C., Tram, E., Vlasschaert, M., Di Nicola, N., ... Siminovitch, K. (2012). Long-Range PCR and Next-Generation Sequencing of BRCA1 and BRCA2 in Breast Cancer. *The Journal of Molecular Diagnostics*, 14(5), 467–475. <https://doi.org/10.1016/j.jmoldx.2012.03.006>
- Pacheco, M. A., Matta, N. E., Valkiūnas, G., Parker, P. G., Mello, B., Stanley, C. E., ... Escalante, A. A. (2018). Mode and Rate of Evolution of Haemosporidian Mitochondrial Genomes: Timing the Radiation of Avian Parasites. *Molecular biology and evolution*, 35(2), 383–403. <https://doi.org/10.1093/molbev/msx285>
- Page, R. D. M. (1996). Tree View: An application to display phylogenetic trees on personal computers. *Bioinformatics*, 12(4), 357–358. <https://doi.org/10.1093/bioinformatics/12.4.357>
- Pain, A. (2005). Genome of the Host-Cell Transforming Parasite *Theileria annulata* Compared with *T. parva*. *Science*, 309(5731), 131–133. <https://doi.org/10.1126/science.1110418>
- Pain, A., Böhme, U., Berry, A. E., Mungall, K., Finn, R. D., Jackson, A. P., ... Berriman, M. (2008). The genome of the simian and human malaria parasite *Plasmodium knowlesi*. *Nature*, 455(7214), 799–803. <https://doi.org/10.1038/nature07306>
- Painter, H. J., Campbell, T. L., & Llinás, M. (2011). The Apicomplexan AP2 family: Integral factors regulating *Plasmodium* development. *Molecular and Biochemical Parasitology*, 176(1), 1–7. <https://doi.org/10.1016/j.molbiopara.2010.11.014>
- Palinauskas, V., Krizanauskiene, A., Iezhova, T. A., Bolshakov, C. V., Jönsson, J., Bensch, S., & Valkiūnas, G. (2013). A new method for isolation of purified genomic DNA from haemosporidian parasites inhabiting nucleated red blood cells. *Experimental Parasitology*, 133(3), 275–280. <https://doi.org/10.1016/j.exppara.2012.12.003>
- Pandey, R. V., Pabinger, S., Kriegner, A., & Weinhäusel, A. (2016). ClinQC: a tool for quality control and cleaning of Sanger and NGS data in clinical research. *BMC Bioinformatics*, 17(1), 56. <https://doi.org/10.1186/s12859-016-0915-y>
- Parfrey, L. W., Lahr, D. J. G., Knoll, A. H., & Katz, L. A. (2011). Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proceedings of the National Academy of Sciences*, 108(33), 13624–13629. <https://doi.org/10.1073/pnas.1110633108>
- Parussini, F., Coppens, I., Shah, P. P., Diamond, S. L., & Carruthers, V. B. (2010). Cathepsin L occupies a vacuolar compartment and is a protein maturase within the endo/exocytic system of *Toxoplasma gondii*. *Molecular Microbiology*, 76(6), 1340–1357. <https://doi.org/10.1111/j.1365-2958.2010.07181.x>
- Patil, H., Hughes, K. R., Lemgruber, L., Philip, N., Dickens, N., & Waters, A. P. (2019). Zygote morphogenesis but not the establishment of cell polarity in *Plasmodium berghei* is controlled by the small GTPase, RAB11A. *bioRxiv* 811802. <https://doi.org/10.1101/811802>
- Pauli, M., Chakarov, N., Rupp, O., Kalinowski, J., Goesmann, A., Sorenson, M. D., ... Hoffman, J. I. (2015). De novo assembly of the dual transcriptomes of a polymorphic raptor species and its malarial parasite. *BMC Genomics*, 16(1), 1038. <https://doi.org/10.1186/s12864-015-2254-1>
- Pelassa, I., & Fiumara, F. (2015). Differential Occurrence of Interactions and Interaction Domains in Proteins Containing Homopolymeric Amino Acid Repeats. *Frontiers in Genetics*, 6. <https://doi.org/10.3389/fgene.2015.00345>
- Pereira, S. L. (2000). Mitochondrial genome organization and vertebrate phylogenetics. *Genetics and Molecular Biology*, 23(4), 745–752. <https://doi.org/10.1590/S1415-47572000000400008>
- Pérez-Tris, J., Hasselquist, D., Hellgren, O., Krizanauskiene, A., Waldenström, J., & Bensch, S. (2005). What are malaria parasites? *Trends in Parasitology*, 21(5), 209–211. <https://doi.org/10.1016/j.pt.2005.03.001>
- Perkins, S. L. (2014). Malaria's Many Mates: Past, Present, and Future of the Systematics of the Order Haemosporida. *Journal of Parasitology*, 100(1), 11–25. <https://doi.org/10.1645/13-362.1>
- Perkins, S. L., & Schaer, J. (2016). A Modern Menagerie of Mammalian Malaria. *Trends in Parasitology*, 32(10), 772–782. <https://doi.org/10.1016/j.pt.2016.06.001>
- Perkins, S. L., & Schall, J. J. (2002). A molecular phylogeny of malarial parasites recovered from cytochrome b gene sequences. *The Journal of parasitology*, 88(5), 972–978. [https://doi.org/10.1645/0022-3395\(2002\)088\[0972:AMPOMP\]2.0.CO;2](https://doi.org/10.1645/0022-3395(2002)088[0972:AMPOMP]2.0.CO;2)
- Petersen, F. T., Meier, R., Kutty, S. N., & Wiegmann, B. M. (2007). The phylogeny and evolution of host choice in the Hippoboscoidea (Diptera) as reconstructed using four molecular markers. *Molecular Phylogenetics and Evolution*, 45(1), 111–122. <https://doi.org/10.1016/j.ympev.2007.04.023>
- Petersen, T. N., Brunak, S., von Heijne, G., & Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, 8(10), 785–786. <https://doi.org/10.1038/nmeth.1701>

- Pick, C., Ebersberger, I., Spielmann, T., Bruchhaus, I., & Burmester, T. (2011). Phylogenomic analyses of malaria parasites and evolution of their exported proteins. *BMC Evolutionary Biology*, 11(1), 167. <https://doi.org/10.1186/1471-2148-11-167>
- Pineda-Catalan, O., Perkins, S. L., Peirce, M. A., Engstrand, R., Garcia-Davila, C., Pinedo-Vasquez, M., & Aguirre, A. A. (2013). Revision of Hemoproteid Genera and Description and Redescription of Two Species of Chelonian Hemoproteid Parasites. *Journal of Parasitology*, 99(6), 1089–1098. <https://doi.org/10.1645/13-296.1>
- Price, A., & Gibas, C. (2017). The quantitative impact of read mapping to non-native reference genomes in comparative RNA-Seq studies. *PLOS ONE*, 12(7), e0180904. <https://doi.org/10.1371/journal.pone.0180904>
- Ramakrishnan, C., Dessens, J. T., Armson, R., Pinto, S. B., Talman, A. M., Blagborough, A. M., & Sinden, R. E. (2011). Vital functions of the malarial ookinete protein, CTRP, reside in the A domains. *International Journal for Parasitology*, 41(10), 1029–1039. <https://doi.org/10.1016/j.ijpara.2011.05.007>
- Ramasindrazana, B., Goodman, S. M., Dsouli, N., Gomard, Y., Lagadec, E., Randrianarivelojosia, M., ... Tortosa, P. (2018). Polychromophilus spp. (Haemosporida) in Malagasy bats: host specificity and insights on invertebrate vectors. *Malaria Journal*, 17(1), 318. <https://doi.org/10.1186/s12936-018-2461-8>
- Rambaut, A. (2004). FigTree v1.4.3. Abgerufen 17. Februar 2018, von [p://tree.bio.ed.ac.uk/software/figtree/](http://tree.bio.ed.ac.uk/software/figtree/)
- Reid, A. J. (2015). Large, rapidly evolving gene families are at the forefront of host-parasite interactions in Apicomplexa. *Parasitology*, 142 Suppl, S57-70. <https://doi.org/10.1017/S0031182014001528>
- Riechmann, J. L., & Meyerowitz, E. M. (1998). The AP2/EREBP family of plant transcription factors. *Biological Chemistry*, 379(6). <https://doi.org/10.1515/bchm.1998.379.6.633>
- Rödelsperger, C., Prabh, N., & Sommer, R. J. (2019). New Gene Origin and Deep Taxon Phylogenomics: Opportunities and Challenges. *Trends in Genetics*, 35(12), 914–922. <https://doi.org/10.1016/j.tig.2019.08.007>
- Rogowska, K., & Kokurewicz, T. (2007). The longest migrations of three bat species to the „Nietoperek“ bat reserve (Western Poland). *Berichte der Naturforschenden Gesellschaft der Oberlausitz Supplement zu Band 15*, 53–60. Abgerufen von 3rd International Conference “Bats of the Sudety mountains
- Russell, K., Cheng, C.-H., Bizzaro, J. W., Ponts, N., Emes, R. D., Le Roch, K., ... Horrocks, P. (2014). Homopolymer tract organization in the human malarial parasite Plasmodium falciparum and related Apicomplexan parasites. *BMC Genomics*, 15(1), 848. <https://doi.org/10.1186/1471-2164-15-848>
- Russo, I., Babbitt, S., Muralidharan, V., Butler, T., Oksman, A., & Goldberg, D. E. (2010). Plasmepsin V licenses Plasmodium proteins for export into the host erythrocyte. *Nature*, 463(7281), 632–636. <https://doi.org/10.1038/nature08726>
- Ryan, U., Paparini, A., Monis, P., & Hijawi, N. (2016). It's official – Cryptosporidium is a gregarine: What are the implications for the water industry? *Water Research*, 105, 305–313. <https://doi.org/10.1016/j.watres.2016.09.013>
- Sabina, J., & Leamon, J. H. (2015). *Bias in Whole Genome Amplification: Causes and Considerations*. https://doi.org/10.1007/978-1-4939-2990-0_2
- Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., ... Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC biology*, 12, 87. <https://doi.org/10.1186/s12915-014-0087-z>
- San Martin, G. (2008). Myotis daubentonii [Foto]. Abgerufen 25. August 2019, von <https://www.flickr.com/photos/sanmartin/2862372383/in/photolist-5mWpZ-6XHhYy-6XHfXm-5mWzDX-5n1Aew-5mPMBt-5mPLSx-5mW7ox/>
- San Martin, G. (2010). Nycteribiidae (parasite fly living on bats) [Foto]. Abgerufen 25. August 2019, von <https://www.flickr.com/photos/sanmartin/5021769088/in/photolist-8DKSZE-8DKRXY-5mPNyx>
- Sanders, P. R., Gilson, P. R., Cantin, G. T., Greenbaum, D. C., Nebl, T., Carucci, D. J., ... Crabb, B. S. (2005). Distinct Protein Classes Including Novel Merozoite Surface Antigens in Raft-like Membranes of Plasmodium falciparum. *Journal of Biological Chemistry*, 280(48), 40169–40176. <https://doi.org/10.1074/jbc.M509631200>
- Santos, J. C., Tarvin, R. D., O'Connell, L. A., Blackburn, D. C., & Coloma, L. A. (2018). Diversity within diversity: Parasite species richness in poison frogs assessed by transcriptomics. *Molecular Phylogenetics and Evolution*, 125, 40–50. <https://doi.org/10.1016/j.ympev.2018.03.015>
- Sargeant, T. J., Marti, M., Caler, E., Carlton, J. M.-R., Simpson, K., Speed, T. P., & Cowman, A. F. (2006). Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome biology*, 7(2), R12. <https://doi.org/10.1186/gb-2006-7-2-r12>
- Sato, S., Sesay, A. K., & Holder, A. A. (2013). The Unique Structure of the Apicoplast Genome of the Rodent Malaria Parasite Plasmodium chabaudi chabaudi. *PLOS ONE*, 8(4), e61778. <https://doi.org/10.1371/journal.pone.0061778>
- Schaer, J., Perkins, S. L., Decher, J., Leendertz, F. H., Fahr, J., Weber, N., & Matuschewski, K. (2013). High diversity of West African bat malaria parasites and a tight link with rodent Plasmodium taxa. *Proceedings of the National Academy of Sciences*, 110(43), 17415–17419. <https://doi.org/10.1073/pnas.1311061110>
- Schaer, J., Reeder, D. M., Vodzak, M. E., Olival, K. J., Weber, N., Mayer, F., ... Perkins, S. L. (2015). Nycteridia parasites of Afrotropical insectivorous bats. *International Journal for Parasitology*, 45(6), 375–384. <https://doi.org/10.1016/j.ijpara.2015.01.008>

- Schaer, J., McMichael, L., Gordon, A. N., Russell, D., Matuschewski, K., Perkins, S. L., ... Power, M. (2018). Phylogeny of Hepatocystis parasites of Australian flying foxes reveals distinct parasite clade. *International Journal for Parasitology: Parasites and Wildlife*, 7(2), 207–212. <https://doi.org/10.1016/j.ijppaw.2018.06.001>
- Schaer, J., Perkins, S. L., Ejotre, I., Vodzak, M. E., Matuschewski, K., & Reeder, D. M. (2017). Epauletted fruit bats display exceptionally high infections with a Hepatocystis species complex in South Sudan. *Scientific Reports*, 7(1), 6928. <https://doi.org/10.1038/s41598-017-07093-z>
- Scheffler, I. (2008). Zur Fähigkeit von Ektoparasiten der Fledermäuse ihre Wirte aktiv aufzusuchen. *Nyctalus (N.F.)*, 13(2–3), 177–189.
- Schmid-Hempel, P. (2008). Parasite immune evasion: a momentous molecular war. *Trends in Ecology & Evolution*, 23(6), 318–326. <https://doi.org/10.1016/j.tree.2008.02.011>
- Schmieder, R., & Edwards, R. (2011). Fast Identification and Removal of Sequence Contamination from Genomic and Metagenomic Datasets. *PLoS ONE*, 6(3), e17288. <https://doi.org/10.1371/journal.pone.0017288>
- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7), 671–675. <https://doi.org/10.1038/nmeth.2089>
- Schulze, J., Kwiatkowski, M., Borner, J., Schlüter, H., Bruchhaus, I., Burmester, T., ... Pick, C. (2015). The Plasmodium falciparum exportome contains non-canonical PEXEL/HT proteins. *Molecular Microbiology*, 97(2), 301–314. <https://doi.org/10.1111/mmi.13024>
- Schuster, F. L. (2002). Cultivation of plasmodium spp. *Clinical microbiology reviews*, 15(3), 355–364. <https://doi.org/10.1128/cmr.15.3.355-364.2002>
- Seim, I., Fang, X., Xiong, Z., Lobanov, A. V., Huang, Z., Ma, S., ... Gladyshev, V. N. (2013). Genome analysis reveals insights into physiology and longevity of the Brandt's bat Myotis brandtii. *Nature Communications*, 4(1), 2212. <https://doi.org/10.1038/ncomms3212>
- Shermann, I. W., & Hull, R. W. (1960). The Pigment (hemozoin) and Proteins of the Avian Malaria Parasite Plasmodium lophurae *. *The Journal of Protozoology*, 7(4), 409–416. <https://doi.org/10.1111/j.1550-7408.1960.tb05989.x>
- Shrivastava, A. K., Kumar, S., Smith, W. A., & Sahu, P. S. (2017). Revisiting the global problem of cryptosporidiosis and recommendations. *Tropical parasitology*, 7(1), 8–17. <https://doi.org/10.4103/2229-5070.202290>
- Sinden, R. E., Canning, E. U., & Spain, B. (1976). Gametogenesis and fertilization in Plasmodium yoelii nigeriensis : a transmission electron microscope study. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 193(1110), 55–76. <https://doi.org/10.1098/rspb.1976.0031>
- Singh, G. P., Chandra, B. R., Bhattacharya, A., Akhouri, R. R., Singh, S. K., & Sharma, A. (2004). Hyper-expansion of asparagines correlates with an abundance of proteins with prion-like domains in Plasmodium falciparum. *Molecular and Biochemical Parasitology*, 137(2), 307–319. <https://doi.org/10.1016/j.molbiopara.2004.05.016>
- Sinha, A., Hughes, K. R., Modrzynska, K. K., Otto, T. D., Pfander, C., Dickens, N. J., ... Waters, A. P. (2014). A cascade of DNA-binding proteins for sexual commitment and development in Plasmodium. *Nature*, 507(7491), 253–257. <https://doi.org/10.1038/nature12970>
- Slamovits, C. H., Fast, N. M., Law, J. S., & Keeling, P. J. (2004). Genome Compaction and Stability in Microsporidian Intracellular Parasites. *Current Biology*, 14(10), 891–896. <https://doi.org/10.1016/j.cub.2004.04.041>
- Slater, G. S. C., & Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC bioinformatics*, 6, 31. <https://doi.org/10.1186/1471-2105-6-31>
- Sleebs, B. E., Lopaticki, S., Marapana, D. S., O'Neill, M. T., Rajasekaran, P., Gazdik, M., ... Boddey, J. A. (2014). Inhibition of Plasmepsin V Activity Demonstrates Its Essential Role in Protein Export, PfEMP1 Display, and Survival of Malaria Parasites. *PLoS Biology*, 12(7), e1001897. <https://doi.org/10.1371/journal.pbio.1001897>
- Small, S. T., Labbé, F., Coulibaly, Y. I., Nutman, T. B., King, C. L., Serre, D., & Zimmerman, P. A. (2019). Human Migration and the Spread of the Nematode Parasite Wuchereria bancrofti. *Molecular Biology and Evolution*, 36(9), 1931–1941. <https://doi.org/10.1093/molbev/msz116>
- Smit, A., Hubley, R., & Green, P. (2013). RepeatMasker Open-4.0. Abgerufen von <http://www.repeatmasker.org>
- Soldati-Favre, D. (2008). Molecular dissection of host cell invasion by the Apicomplexans: the glideosome. *Parasite*, 15(3), 197–205. <https://doi.org/10.1051/parasite/2008153197>
- Sonnhammer, E. L., von Heijne, G., & Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. *Proceedings. International Conference on Intelligent Systems for Molecular Biology*, 6, 175–182. Abgerufen von <http://www.ncbi.nlm.nih.gov/pubmed/9783223>
- Soulard, V., Bosson-Vanga, H., Lorthiois, A., Roucher, C., Franetich, J.-F., Zanghi, G., ... Mazier, D. (2015). Plasmodium falciparum full life cycle and Plasmodium ovale liver stages in humanized mice. *Nature Communications*, 6(1), 7690. <https://doi.org/10.1038/ncomms8690>
- Souvorov, A., Agarwala, R., & Lipman, D. J. (2018). SKESA: strategic k-mer extension for scrupulous assemblies. *Genome Biology*, 19(1), 153. <https://doi.org/10.1186/s13059-018-1540-z>

- Spielmann, T., & Gilberger, T.-W. (2015). Critical Steps in Protein Export of *Plasmodium falciparum* Blood Stages. *Trends in Parasitology*, 31(10), 514–525. <https://doi.org/10.1016/j.pt.2015.06.010>
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
- Stanke, M., & Waack, S. (2003). Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics*, 19(Suppl 2), ii215–ii225. <https://doi.org/10.1093/bioinformatics/btg1080>
- Steinbiss, S., Silva-Franco, F., Brunk, B., Foth, B., Hertz-Fowler, C., Berriman, M., & Otto, T. D. (2016). Companion : a web server for annotation and analysis of parasite genomes. *Nucleic Acids Research*, 44(W1), W29–W34. <https://doi.org/10.1093/nar/gkw292>
- Su, X., Lane, K. D., Xia, L., Sá, J. M., & Wellem, T. E. (2019). Plasmodium Genomics and Genetics: New Insights into Malaria Pathogenesis, Drug Resistance, Epidemiology, and Evolution. *Clinical Microbiology Reviews*, 32(4). <https://doi.org/10.1128/CMR.00019-19>
- Sukumaran, J., & Holder, M. T. (2010). DendroPy: a Python library for phylogenetic computing. *Bioinformatics*, 26(12), 1569–1571. <https://doi.org/10.1093/bioinformatics/btq228>
- Sundararaman, S. A., Plenderleith, L. J., Liu, W., Loy, D. E., Learn, G. H., Li, Y., ... Hahn, B. H. (2016). Genomes of cryptic chimpanzee *Plasmodium* species reveal key evolutionary events leading to human malaria. *Nature Communications*, 7(1), 11078. <https://doi.org/10.1038/ncomms11078>
- Suzuki, S., Kakuta, M., Ishida, T., & Akiyama, Y. (2014). GHOSTX: An Improved Sequence Homology Search Algorithm Using a Query Suffix Array and a Database Suffix Array. *PLoS ONE*, 9(8), e103833. <https://doi.org/10.1371/journal.pone.0103833>
- Suzuki, S., Ishida, T., Ohue, M., Kakuta, M., & Akiyama, Y. (2017). GHOSTX: A Fast Sequence Homology Search Tool for Functional Annotation of Metagenomic Data. In D. Kihara (Hrsg.), *Protein Function Prediction* (S. 15–25). https://doi.org/10.1007/978-1-4939-7015-5_2
- Swain, M. T., Tsai, I. J., Assefa, S. A., Newbold, C., Berriman, M., & Otto, T. D. (2012). A post-assembly genome-improvement toolkit (PAGIT) to obtain annotated genomes from contigs. *Nature Protocols*, 7(7), 1260–1284. <https://doi.org/10.1038/nprot.2012.068>
- Swapna, L. S., & Parkinson, J. (2017). Genomics of apicomplexan parasites. *Critical Reviews in Biochemistry and Molecular Biology*, 52(3), 254–273. <https://doi.org/10.1080/10409238.2017.1290043>
- Szentiványi, T., Estók, P., & Földvári, M. (2016). Checklist of host associations of European bat flies (Diptera: Nycteribiidae, Streblidae). *Zootaxa*, 4205(2), 101. <https://doi.org/10.11646/zootaxa.4205.2.1>
- Tachibana, S.-I., Sullivan, S. A., Kawai, S., Nakamura, S., Kim, H. R., Goto, N., ... Tanabe, K. (2012). *Plasmodium cynomolgi* genome sequences provide insight into *Plasmodium vivax* and the monkey malaria clade. *Nature Genetics*, 44(9), 1051–1055. <https://doi.org/10.1038/ng.2375>
- Talman, A. M., Domarle, O., McKenzie, F. E., Arie, F., & Robert, V. (2004). Gametocytogenesis: the puberty of *Plasmodium falciparum*. *Malaria Journal*, 3, 24. <https://doi.org/10.1186/1475-2875-3-24>
- Tao, Z.-Y., Xia, H., Cao, J., & Gao, Q. (2011). Development and evaluation of a prototype non-woven fabric filter for purification of malaria-infected blood. *Malaria Journal*, 10(1), 251. <https://doi.org/10.1186/1475-2875-10-251>
- Tautz, D., & Domazet-Lošo, T. (2011). The evolutionary origin of orphan genes. *Nature Reviews Genetics*, 12(10), 692–702. <https://doi.org/10.1038/nrg3053>
- Tawk, L., Lacroix, C., Gueirard, P., Kent, R., Gorgette, O., Thiberge, S., ... Barale, J.-C. (2013). A Key Role for *Plasmodium* Subtilisin-like SUB1 Protease in Egress of Malaria Parasites from Host Hepatocytes. *Journal of Biological Chemistry*, 288(46), 33336–33346. <https://doi.org/10.1074/jbc.M113.513234>
- Team, S. T. D. (2015). SRA Toolkit. Abgerufen 1. Februar 2018, von <http://ncbi.github.io/sra-tools/>
- Templeton, T. J., Martinsen, E. S., Kaewthamasorn, M., & Kaneko, O. (2016). The rediscovery of malaria parasites of ungulates. *Parasitology*, 143(12), 1501–1508. <https://doi.org/10.1017/S0031182016001141>
- Testa, A. C., Oliver, R. P., & Hane, J. K. (2016). OcculterCut: A Comprehensive Survey of AT-Rich Regions in Fungal Genomes. *Genome Biology and Evolution*, 8(6), 2044–2064. <https://doi.org/10.1093/gbe/evw121>
- Thakur, A., Mikkelsen, H., & Jungersen, G. (2019). Intracellular Pathogens: Host Immunity and Microbial Persistence Strategies. *Journal of Immunology Research*, 2019, 1–24. <https://doi.org/10.1155/2019/1356540>
- The Gene Ontology Consortium. (2019). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*, 47(D1), D330–D338. <https://doi.org/10.1093/nar/gky1055>
- Theodor, O., & Moscona, A. (1954). On bat parasites in Palestine. I. Nycteribiidae, Streblidae, Hemiptera, Siphonaptera. *Parasitology*, 44, 157–245.
- Thomas, D. C., Ahmed, A., Gilberger, T. W., & Sharma, P. (2012). Regulation of *Plasmodium falciparum* Glideosome Associated Protein 45 (PfGAP45) Phosphorylation. *PLoS ONE*, 7(4), e35855. <https://doi.org/10.1371/journal.pone.0035855>
- Toscani Field, J., Weinberg, J., Bensch, S., Matta, N. E., Valkiūnas, G., & Sehgal, R. N. M. (2018). Delineation of the Genera

- Haemoproteus and Plasmodium Using RNA-Seq and Multi-gene Phylogenetics. *Journal of Molecular Evolution*, 86(9), 646–654. <https://doi.org/10.1007/s00239-018-9875-3>
- Trang, D. T. X., Huy, N. T., Kariu, T., Tajima, K., & Kamei, K. (2004). One-step concentration of malarial parasite-infected red blood cells and removal of contaminating white blood cells. *Malaria journal*, 3, 7. <https://doi.org/10.1186/1475-2875-3-7>
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3--new capabilities and interfaces. *Nucleic acids research*, 40(15), e115. <https://doi.org/10.1093/nar/gks596>
- Vaidya, A. B., & Mather, M. W. (2009). Mitochondrial Evolution and Functions in Malaria Parasites. *Annual Review of Microbiology*, 63(1), 249–267. <https://doi.org/10.1146/annurev.micro.091208.073424>
- Valkiūnas, G. (2004a). Methods of Species Identification and Recommendations for Their Description. In *Avian Malaria Parasites and other Haemosporidia* (S. 225–229). Boca Raton, Florida, USA: CRC Press.
- Valkiūnas, G. (2004b). A Brief Outline of the Ultrastructure. In *Avian Malaria Parasites and other Haemosporidia* (S. 47–65). Boca Raton, Florida, USA: CRC Press.
- Valkiūnas, G. (2004c). Life Cycle and Morphology. In *Avian Malaria Parasites and other Haemosporidia* (S. 17–45). Boca Raton, Florida, USA: CRC Press.
- Valkiūnas, G., Iezhova, T. A., Križanauskienė, A., Palinauskas, V., & Bensch, S. (2008). In Vitro Hybridization of Haemoproteus Spp.: An Experimental Approach for Direct Investigation of Reproductive Isolation of Parasites. *Journal of Parasitology*, 94(6), 1385–1394. <https://doi.org/10.1645/GE-1569.1>
- Valkiūnas, G., Palinauskas, V., Križanauskienė, A., Bernotienė, R., Kazlauskienė, R., & Iezhova, T. A. (2013). Further Observations on In Vitro Hybridization of Hemosporidian Parasites: Patterns of Ookinete Development in Haemoproteus Spp. *Journal of Parasitology*, 99(1), 124–136. <https://doi.org/10.1645/GE-3226.1>
- Valkiūnas, G., & Iezhova, T. A. (2017). Exo-erythrocytic development of avian malaria and related haemosporidian parasites. *Malaria journal*, 16(1), 101. <https://doi.org/10.1186/s12936-017-1746-7>
- van Dijk, M. R., Janse, C. J., Thompson, J., Waters, A. P., Braks, J. A. M., Dodemont, H. J., ... Eling, W. (2001). A Central Role for P48/45 in Malaria Parasite Male Gamete Fertility. *Cell*, 104(1), 153–164. [https://doi.org/10.1016/S0092-8674\(01\)00199-4](https://doi.org/10.1016/S0092-8674(01)00199-4)
- van Schaik, J., Janssen, R., Bosch, T., Haarsma, A.-J., Dekker, J. J. A., & Kranstauber, B. (2015). Bats Swarm Where They Hibernate: Compositional Similarity between Autumn Swarming and Winter Hibernation Assemblages at Five Underground Sites. *PLOS ONE*, 10(7), e0130850. <https://doi.org/10.1371/journal.pone.0130850>
- Vaughan, N. (1997). The diets of British bats (Chiroptera). *Mammal Review*, 27(2), 77–94. <https://doi.org/10.1111/j.1365-2907.1997.tb00373.x>
- Venkatesan, M., Amaratunga, C., Campino, S., Auburn, S., Koch, O., Lim, P., ... Plowe, C. V. (2012). Using CF11 cellulose columns to inexpensively and effectively remove human DNA from Plasmodium falciparum-infected whole blood samples. *Malaria Journal*, 11(1), 41. <https://doi.org/10.1186/1475-2875-11-41>
- Videvall, E. (2018). Plasmodium parasites of birds have the most AT-rich genes of eukaryotes. *Microbial Genomics*, 4(2). <https://doi.org/10.1099/mgen.0.000150>
- Videvall, E., Cornwallis, C. K., Palinauskas, V., Valkiūnas, G., & Hellgren, O. (2015). The Avian Transcriptome Response to Malaria Infection. *Molecular Biology and Evolution*, 32(5), 1255–1267. <https://doi.org/10.1093/molbev/msv016>
- Vulliez-Le Normand, B., Tonkin, M. L., Lamarque, M. H., Langer, S., Hoos, S., Roques, M., ... Lebrun, M. (2012). Structural and Functional Insights into the Malaria Parasite Moving Junction Complex. *PLoS Pathogens*, 8(6), e1002755. <https://doi.org/10.1371/journal.ppat.1002755>
- Wasmuth, J., Daub, J., Peregrin-Alvarez, J. M., Finney, C. A. M., & Parkinson, J. (2009). The origins of apicomplexan sequence innovation. *Genome Research*, 19(7), 1202–1213. <https://doi.org/10.1101/gr.083386.108>
- Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., ... Zdobnov, E. M. (2018). BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. *Molecular Biology and Evolution*, 35(3), 543–548. <https://doi.org/10.1093/molbev/msx319>
- Waters, A. P., Higgins, D. G., & McCutchan, T. F. (1991). Plasmodium falciparum appears to have arisen as a result of lateral transfer between avian and human hosts. *Proceedings of the National Academy of Sciences*, 88(8), 3140–3144. <https://doi.org/10.1073/pnas.88.8.3140>
- Whittaker, C. A., & Hynes, R. O. (2002). Distribution and Evolution of von Willebrand/Integrin A Domains: Widely Dispersed Domains with Roles in Cell Adhesion and Elsewhere. *Molecular Biology of the Cell*, 13(10), 3369–3387. <https://doi.org/10.1091/mbc.e02-05-0259>
- WHO. (2018). *World Malaria Report 2018*. Abgerufen von <https://www.who.int/malaria/publications/world-malaria-report-2018/report/en/>
- Wilkinson, D. A., Duron, O., Cordonin, C., Gomard, Y., Ramasindrazana, B., Mavingui, P., ... Tortosa, P. (2016). The Bacteriome of Bat Flies (Nycteribiidae) from the Malagasy Region: a Community Shaped by Host Ecology, Bacterial Transmission Mode, and Host-Vector Specificity. *Applied and Environmental Microbiology*, 82(6), 1778–1788. <https://doi.org/10.1128/AEM.03505-15>

- Wilkinson, L. C., Handel, C. M., Van Hemert, C., Loiseau, C., & Sehgal, R. N. M. (2016). Avian malaria in a boreal resident species: long-term temporal variability, and increased prevalence in birds with avian keratin disorder. *International Journal for Parasitology*, 46(4), 281–290. <https://doi.org/10.1016/j.ijpara.2015.12.008>
- Williamson, K. C., Fujioka, H., Aikawa, M., & Kaslow, D. C. (1996). Stage-specific processing of Pfs230, a *Plasmodium falciparum* transmission-blocking vaccine candidate. *Molecular and Biochemical Parasitology*, 78(1–2), 161–169. [https://doi.org/10.1016/S0166-6851\(96\)02621-7](https://doi.org/10.1016/S0166-6851(96)02621-7)
- Wilson, R. J. M., Gardner, M. J., Feagin, J. E., & Williamson, D. H. (1991). Have malaria parasites three genomes? *Parasitology Today*, 7(6), 134–136. [https://doi.org/10.1016/0169-4758\(91\)90276-T](https://doi.org/10.1016/0169-4758(91)90276-T)
- Wilson, R. J. M., Denny, P. W., Preiser, P. R., Rangachari, K., Roberts, K., Roy, A., ... Williamson, D. H. (1996). Complete Gene Map of the Plastid-like DNA of the Malaria Parasite *Plasmodium falciparum*. *Journal of Molecular Biology*, 261(2), 155–172. <https://doi.org/10.1006/jmbi.1996.0449>
- Wirth, C. C., Glushakova, S., Scheuermayer, M., Repnik, U., Garg, S., Schaack, D., ... Pradel, G. (2014). Perforin-like protein PPLP2 permeabilizes the red blood cell membrane during egress of *Plasmodium falciparum* gametocytes. *Cellular Microbiology*, 16(5), 709–733. <https://doi.org/10.1111/cmi.12288>
- Witsenburg, F., Salamin, N., & Christe, P. (2012). The evolutionary host switches of *Polychromophilus*: a multi-gene phylogeny of the bat malaria genus suggests a second invasion of mammals by a haemosporidian parasite. *Malaria Journal*, 11(1), 53. <https://doi.org/10.1186/1475-2875-11-53>
- Witsenburg, F., Schneider, F., & Christe, P. (2014). Epidemiological traits of the malaria-like parasite *Polychromophilus murinus* in the Daubenton's bat *Myotis daubentonii*. *Parasites and Vectors*, 7(1), 1–8. <https://doi.org/10.1186/s13071-014-0566-7>
- Witsenburg, F., Clément, L., López-Baucells, A., Palmeirim, J., Pavlinić, I., Scaravelli, D., ... Christe, P. (2015). How a haemosporidian parasite of bats gets around: the genetic structure of a parasite, vector and host compared. *Molecular Ecology*, 24(4), 926–940. <https://doi.org/10.1111/mec.13071>
- Witsenburg, F., Schneider, F., & Christe, P. (2015). Signs of a vector's adaptive choice: on the evasion of infectious hosts and parasite-induced mortality. *Oikos*, 124(5), 668–676. <https://doi.org/10.1111/oik.01785>
- Woo, Y. H., Ansari, H., Otto, T. D., Klinger, C. M., Kolisko, M., Michálek, J., ... Pain, A. (2015). Chromerid genomes reveal the evolutionary path from photosynthetic algae to obligate intracellular parasites. *eLife*, 4. <https://doi.org/10.7554/eLife.06974>
- Worthey, E. A., & Myler, P. J. (2005). Protozoan genomes: gene identification and annotation. *International Journal for Parasitology*, 35(5), 495–512. <https://doi.org/10.1016/j.ijpara.2005.02.008>
- Woyke, T., Tighe, D., Mavromatis, K., Clum, A., Copeland, A., Schackwitz, W., ... Cheng, J.-F. (2010). One Bacterial Cell, One Complete Genome. *PLoS ONE*, 5(4), e10314. <https://doi.org/10.1371/journal.pone.0010314>
- Yadav, M. K., & Swati, D. (2012). Comparative genome analysis of six malarial parasites using codon usage bias based tools. *Bioinformation*, 8(24), 1230–1239. <https://doi.org/10.6026/97320630081230>
- Yamauchi, L. M., Coppi, A., Snounou, G., & Sinnis, P. (2007). *Plasmodium* sporozoites trickle out of the injection site. *Cellular Microbiology*, 9(5), 1215–1222. <https://doi.org/10.1111/j.1462-5822.2006.00861.x>
- Yotoko, K. S. C., & Elisei, C. (2006). Malaria parasites (Apicomplexa, Haematozoa) and their relationships with their hosts: is there an evolutionary cost for the specialization? *Journal of Zoological Systematics and Evolutionary Research*, 44(4), 265–273. <https://doi.org/10.1111/j.1439-0469.2006.00377.x>
- Youssef, N. H., Couger, M. B., Struchtemeyer, C. G., Liggensstoffer, A. S., Prade, R. A., Najjar, F. Z., ... Elshahed, M. S. (2013). The Genome of the Anaerobic Fungus *Orpinomyces* sp. Strain C1A Reveals the Unique Evolutionary History of a Remarkable Plant Biomass Degrader. *Applied and Environmental Microbiology*, 79(15), 4620–4634. <https://doi.org/10.1128/AEM.00821-13>
- Yuda, M., Iwanaga, S., Shigenobu, S., Mair, G. R., Janse, C. J., Waters, A. P., ... Kaneko, I. (2009). Identification of a transcription factor in the mosquito-invasive stage of malaria parasites. *Molecular Microbiology*, 71(6), 1402–1414. <https://doi.org/10.1111/j.1365-2958.2009.06609.x>
- Yuda, M., Iwanaga, S., Shigenobu, S., Kato, T., & Kaneko, I. (2010). Transcription factor AP2-Sp and its target genes in malarial sporozoites. *Molecular Microbiology*, 75(4), 854–863. <https://doi.org/10.1111/j.1365-2958.2009.07005.x>
- Zerbino, D. R., & Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, 18(5), 821–829. <https://doi.org/10.1101/gr.074492.107>
- Zhang, G., Cowled, C., Shi, Z., Huang, Z., Bishop-Lilly, K. A., Fang, X., ... Wang, J. (2013). Comparative Analysis of Bat Genomes Provides Insight into the Evolution of Flight and Immunity. *Science*, 339(6118), 456–460. <https://doi.org/10.1126/science.1230835>
- Zuegge, J., Ralph, S., Schmuker, M., McFadden, G. I., & Schneider, G. (2001). Deciphering apicoplast targeting signals – feature extraction from nuclear-encoded precursors of *Plasmodium falciparum* apicoplast proteins. *Gene*, 280(1–2), 19–26. [https://doi.org/10.1016/S0378-1119\(01\)00776-4](https://doi.org/10.1016/S0378-1119(01)00776-4)
- Zwickl, D. J. (2006). *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion*. (The University of Texas at Austin). Abgerufen von <http://hdl.handle.net/2152/2666>

7 Anhang

7.1 Abkürzungsverzeichnis

+F	empirisch ermittelte Basenfrequenzen
+G	Gamma-Distribution der Substitutionsraten
+I	Anteil invariabler Positionen
ad	addiere/auffüllen auf
adh	ATP dependent helicase
Amp	Ampicillin
Api	Apikomplexa
asl	Adenylosuccinat lyase
aa	Aminosäure
AIC	Akaike Informationskriterium
AMA	apical membrane antigen
AP2	Apetala2
atp	ATPase subunit
ATP	Adenosintriphosphat
BI	Bayes'sche Interferenz
bidest.	bidestilliert
Blast	„Basic Local Alignment Search Tool“
BlastN	Nukleotide-Blast, Suchsequenz und Datenbank: Nukleotide
BlastX	Protein-Blast, Suchsequenz: Nukleotide, Datenbank: Proteine
b	Basen
bp	Basenpaare
BSA	Rinderserumalbumin
bzw.	Beziehungsweise
ca.	circa (ungefähr)
CellTOS	cell-traversal protein for <i>Plasmodium</i> ookinetes and sporozoites
clpc	caseinolytic protease
coI	cytochrome oxidase subunit I (siehe cox)
cox	cytochrome oxidase (COX1/2/3)
CLAG	cytoadherence linked asexual protein
CSB	conserved sequence block (konservierte Sequenz innerhalb der mtDNA)
CSP	circumsporozoite protein
CTRP	circumsporozoite protein and thrombospondin-related adhesive protein
cys	cystein
cytb	Cytochrom b-Gen
dest.	destilliert
D-Loop	displacement-Loop, Teil der Kontrollregion der mtDNA
DNA	desoxyribonucleic acid (Desoxyribonukleinsäure)

dNTP	Desoxynukleosidtriphosphat
dsDNA	doppelsträngige DNA
E. coli	Escherichia coli
EDTA	Ethylendiamintetraacetat
ef2a	Elongation Factor 2
EGF	epidermal growth factor-like
ERF	ethylen response factor
et al.	et alii, und andere
E-Value	Expect value (Erwartungswert)
ExoSAP	Exonuklease I – Shrimp Alkaline Phosphatase
EXP	exported protein
F/fwd	forward
GC-Gehalt	Anteil Basen Guanin und Cytosin
gDNA	genomische DNA
GEST	gamete egress and sporozoite traversal protein
GO	gene ontology
HPLC	high performance liquid chromatography (Hochleistungsflüssigkeitschromatographie)
HSP	heat shock protein
H-Strang	heavy strand, forward Strang der mtDNA
IPTG	Isopropyl- β -D-Thiogalactopyranosid
kbp	Kilobasenpaare
Krt	Kreatin
L-Strang	light strand, reverse Strang der mtDNA
LSU	large subunit ribosomal ribonucleic acid
M	molar
M. daubentonii/mda	Myotis daubentonii
Mg	Mikrogameten
ML	Maximum Likelihood
MP	Maximum Parsimony
mt	mitochondrial
nad	NADH Dehydrogenase Untereinheit
NCBI	national Center for Biotechnology Information
NCR	non Coding Region (Kontrollregion der mtDNA)
NGS	Next generation sequencing, Hochdurchsatz-Sequenzierungsverfahren
Nyc	Nycteribidae
OD	optische Dichte
OG	Orthologe Gruppen
ORF	open reading frame (offener Leserahmen)
OriL	origin of Replication (L-Strang)
<i>P. falciparum</i> /Pfa	<i>Plasmodium falciparum</i>
<i>P. murinus</i> /Pmu	<i>Polychromophilus murinus</i>

PCR	polymerase chain reaction (Polymerasekettenreaktion)
pdcd6	programmed cell death protein 6
PEXEL	<i>Plasmodium</i> Export Element
pH	lat. pondus Hydrogenii, Maß für Oxoniumionenkonzentration
plDNA	apikomplexe Plastide, Apikoplasmid
Pmu	Polychromophilus murinus
prlr	Prolactin Receptor
PTEx	<i>Plasmodium</i> translocon of exported proteins
R/rev	reverse
RON	rhoptry neck protein
rpl	ribosomal protein L
rpm	rounds per minute, Umdrehungen pro Minute
rps	ribosomal protein S
rRNA/rrn	ribosomale RNA
RT	Raumtemperatur
sf3b1	splicing factor 3B subunit 1
spec.	Species, nicht genau definierte Spezies
SPECT	sporozoite microneme protein essential for cell transversal
ssDNA	einzelsträngige DNA
SSU	small subunit ribosomal ribonucleic acid
SWGA	selective whole genome Amplification (selektive WGA)
TBE	Tris-Borat-EDTA
TRAP	thrombospondin-related anonymous protein
Tris	Tris-(hydroxymethyl)-aminomethane
tRNA/trn	transfer RNA
TSR	thrombospondin typ 1 repeat
TVN	tubovesikulären Netzwerks
UV	Ultraviolett
v/v	Volumen/Volumen-Verhältnis
vgl.	vergleiche
vs.	versus
VWA	von Willebrand Faktor A
w/v	Gewicht/Volumen-Verhältnis
WF	Wasserfledermaus
WGA	whole genome amplification
X-Gal	5-Bromo-4-chloro-3-indoyl- β -D-galactosid
z. B.	zum Beispiel

Tabellarische Übersichten über die in dieser Arbeit verwendeten Abkürzungskodes für Nukleinsäuren und Aminosäuren finden sich in Tab. 7.1 und Tab.7.2.

7.2 Ergänzende Tabellen

Tab. 7.1: Iupac-Abkürzungskode für Nukleotide.

Symbol	Bedeutung	Beschreibung
A		Adenin
C		Cytosin
G		Guanin
T		Thymin
U		Uracil
R	A oder G	Purine
Y	C oder T	Pyrimidine
W	A oder T	weak (schwache) Wasserstoffbrückenbindungen
S	G oder C	strong (starke) Wasserstoffbrückenbindungen
M	A oder C	Aminogruppe
K	G oder T	Ketogruppe
H	A, C, oder T	nicht G, (H folgt im Alphabet auf G)
B	G, C, oder T	nicht A, (B folgt im Alphabet auf A)
V	G, A, oder C	nicht T (U), (V folgt im Alphabet auf U)
D	G, A, oder T	nicht C, (D folgt im Alphabet auf C)
N	G, A, C oder T	any (alle Nukleotide möglich)

Tab. 7.2: Iupac-Abkürzungskodes für Aminosäuren.

Aminosäure	Dreibuchstabencode	Einbuchstabencode
Alanin	Ala	A
Arginin	Arg	R
Asparagin	Asn	N
Asparaginsäure	Asp	D
Cystein	Cys	C
Glutamin	Gln	Q
Glutaminsäure	Glu	E
Glycin	Gly	G
Histidin	His	H
Isoleucin	Ile	I
Leucin	Leu	L
Lysin	Lys	K
Methionin	Met	M
Phenylalanin	Phe	F
Prolin	Pro	P
Serin	Ser	S
Threonin	Thr	T
Tryptophan	Trp	W
Tyrosin	Tyr	Y
Valin	Val	V

Tab. 7.3: Übersicht der gefangenen Fledermäuse.

No.	Fledermausart	Geschlecht	Unterarm- länge	Gewicht	Ektoparasiten
Fransenfledermaus					
1	<i>Myotis natterii</i>	w	39,6 mm	7,8 g	Spin sp. (1)
2	<i>M. natterii</i>	m	39,5 mm	7,1 g	-
3	<i>M. natterii</i>	m	39,1 mm	7,9 g	-
4	<i>M. natterii</i>	m	39,0 mm	7,3 g	Spin sp. (1)
5	<i>M. natterii</i>	m	40,0 mm	7,9 g	-
6	<i>M. natterii</i>	m	40,0 mm	7,7 g	-
7	<i>M. natterii</i>	w	41,3 mm	8,7 g	-
8	<i>M. natterii</i>	w	40,5 mm	7,8 g	Spin sp. (1)
9	<i>M. natterii</i>	m	39,9 mm	7,6 g	Spin sp. (1)
10	<i>M. natterii</i>	w	39,0 mm	7,2 g	Spin sp. (1)
11	<i>M. natterii</i>	m	41,8 mm	7,6 g	-
12	<i>M. natterii</i>	m	40,0 mm	7,9 g	-
13	<i>M. natterii</i>	w	40,7 mm	7,4 g	-
14	<i>M. natterii</i>	m	entkommen		
15	<i>M. natterii</i>	m	41,5 mm	8,0 g	Spin sp. (2), Ischnopsyllus sp. (1)
Bechsteinfledermaus					
16	<i>Myotis bechsteinii</i>	w	42,0 mm	8,2 g	-
17	<i>M. bechsteinii</i>	m	42,1 mm	10,1 g	-
Teichfledermaus					
18	<i>Myotis dasycneme</i>	m	46,7 mm	16,7 g	Argas sp. (3)
19	<i>M. dasycneme</i>	m	46,8 mm	15,4 g	Argas sp. (2)
20	<i>M. dasycneme</i>	m	46,4 mm	18,5 g	Spin sp. (1)
21	<i>M. dasycneme</i>	m	45,0 mm	13,9 g	Spin sp., Argas sp. (1)
22	<i>M. dasycneme</i>	m	45,2 mm	14,4 g	-
23	<i>M. dasycneme</i>	m	45,9 mm	14,1 g	-
24	<i>M. dasycneme</i>	w	46,6 mm	24,5 g	Spin sp. (2), Ischnopsyllus sp. (1)
25	<i>M. dasycneme</i>	m	44,6 mm	19,9 g	Ischnopsyllus sp. (1)
26	<i>M. dasycneme</i>	w	47,5 mm	26,5 g	Spin sp. (3), Ischnopsyllus sp. (1)
27	<i>M. dasycneme</i>	m	46,8 mm	24,3 g	Spin sp. (1), Ischnopsyllus sp. (3)
28	<i>M. dasycneme</i>	w	44,9 mm	14,7 g	-
29	<i>M. dasycneme</i>	w	44,7 mm	14,3 g	-
Wasserfledermaus					
30	<i>Myotis daubentonii</i>	w	38,6 mm	14,0 g	-
31	<i>M. daubentonii</i>	w	38,2 mm	7,3 g	-
32	<i>M. daubentonii</i>	m	37 mm	11,0 g	Spin sp. (1), Nycteribia sp. (4)
33	<i>M. daubentonii</i>	w	36,9 mm	-	Spin sp. (1), Nycteribia sp. (2)
34	<i>M. daubentonii</i>	m	36,9 mm	7,8 g	-
35	<i>M. daubentonii</i>	w	40,1 mm	13,0 g	-
36	<i>M. daubentonii</i>	m	37,9 mm	12,6 g	Spin sp. (2), Nycteribia sp. (2)
37	<i>M. daubentonii</i>	m	37,5 mm	13,4 g	Nycteribia sp. (5)
38	<i>M. daubentonii</i>	m	37,5 mm	12,4 g	Spin sp. (2), Nycteribia sp. (4)
39	<i>M. daubentonii</i>	m	38,4 mm	12,0 g	-
40	<i>M. daubentonii</i>	m	36,3 mm	10,4 g	Spin sp. (2), Nycteribia sp. (2)
41	<i>M. daubentonii</i>	m	37,6 mm	12,2 g	Nycteribia sp. (2)
42	<i>M. daubentonii</i>	w	38,7 mm	7,9 g	-
43	<i>M. daubentonii</i>	m	37,3 mm	6,7 g	Spin sp. (1)
44	<i>M. daubentonii</i>	m	36 mm	9,6 g	Nycteribia sp. (1)
45	<i>M. daubentonii</i>	w	37,2 mm	7,9 g	Nycteribia sp. (1)
46	<i>M. daubentonii</i>	m	36,9 mm	7,2 g	Spin sp., Nycteribia sp. (2)
47	<i>M. daubentonii</i>	m	36,6 mm	11,2 g	Nycteribia sp. (1)
48	<i>M. daubentonii</i>	m	35,4 mm	7,8 g	-
49	<i>M. daubentonii</i>	w	36,8 mm	12,9 g	Spin sp. (1)
50	<i>M. daubentonii</i>	m	38,3 mm	9,4 g	Nycteribia sp. (2)

No., Nummer; m, männlich; w, weiblich; Zahl in Klammern gibt die Anzahl der gefundenen Parasiten an.

Tab. 7.4: GC-Gehalt genomischer Loci von *Polychromophilus* spp.

Gen	GC-Gehalt	Accessionnr.
adenylosuccinate lyase gene (asl)	22,4 %	JN990725
aminophospholipid-transporting P-ATPase 2	27,2 %	LN483006
DNA polymerase delta catalytic subunit	27,5 %	LN482965
LCCL domain-containing protein (CCp1)	26,7 %	LN483032
nuclear elongation factor 2 (EF2)	34,2 %	KF159742
ORF for putative eukaryotic translation initiation factor 2 gamma subunit	30,9 %	LN482994
ORF for putative metabolite/drug transporter	28,6 %	LN483021
pre-mRNA-processing factor 6	28,7 %	LN482982
putative 26S proteasome regulatory subunit 7	31,8 %	LN482947
putative acetyl-CoA synthetase (AACS)	26,8 %	LN483026
putative DNA repair protein RAD54	23,0 %	LN482971
putative DNA replication licensing factor MCM5	33,2 %	LN482988
putative DNA-directed RNA polymerase II	22,7 %	LN482999
putative importin beta	27,1 %	LN482931
putative splicing factor 3B subunit 1	29,7 %	LN483012
putative TCP-1/cpn60 chaperonin family (TCP1)	29,4 %	LN483017
Sec24 subunit a (Sec24a)	26,3%	LN482951

Tab. 7.5: Verwendete Genomdaten zur Rekonstruktion der Phylogenie der Haemosporida.

Spezies	Verbreitung	Spezies	Verbreitung
<i>Babesia bigemina</i>	BOND	<i>Plasmodium knowlesi</i>	Malayan Pk1 (A+)
<i>Babesia bovis</i>	T2Bo	<i>Plasmodium malariae</i>	kA
<i>Babesia microti</i>	RI	<i>Plasmodium ovale</i>	wallikeri
<i>Babesia orientalis</i>	Wuhan	<i>Plasmodium reichenowi</i>	SY57
<i>Haemoproteus tartakowskyi</i>	cyt b lineage SISKIN1	<i>Plasmodium relictum</i>	SGS1
<i>Plasmodium berghei</i>	ANKA	<i>Plasmodium vinckei</i>	vinckei
<i>Plasmodium chabaudi</i>	Chabaudi CB	<i>Plasmodium vivax</i>	Salvador I
<i>Plasmodium coatneyi</i>	CDC	<i>Plasmodium yoelii</i>	17NXL
<i>Plasmodium cynomolgi</i>	M	<i>Polychromophilus murinus</i>	WF43 (Diese Arbeit)
<i>Plasmodium falciparum</i>	3D7	<i>Theileria annulata</i>	Ankara
<i>Plasmodium fragile</i>	nilgiri	<i>Theileria equi</i>	WA
<i>Plasmodium gaboni</i>	SY75	<i>Theileria orientalis</i>	Shintoku
<i>Plasmodium gallinaceum</i>	8A	<i>Theileria parva</i>	Muguga
<i>Plasmodium inui</i>	San Antonio 1		

Genomdaten stammen jeweils PlasmoDB bzw. PiroplasmoDB (Release 32), auf Genom von *H. tartakowskyi* (Bensch et al., 2016) wurde im Juli 2017 zugegriffen.

Tab. 7.6: Verwendete OrthoMCL IDs zur Rekonstruktion der Phylogenie der Haemosporida.

OG5_126539	OG5_127220	OG5_127889	OG5_128499	OG5_130425	OG5_142642	OG5_150558
OG5_126610	OG5_127250	OG5_127890	OG5_128504	OG5_130601	OG5_142685	OG5_150559
OG5_126659	OG5_127287	OG5_127913	OG5_128505	OG5_130790	OG5_142982	OG5_150561
OG5_126684	OG5_127294	OG5_127919	OG5_128515	OG5_130885	OG5_143214	OG5_150563
OG5_126687	OG5_127310	OG5_127921	OG5_128543	OG5_130982	OG5_143221	OG5_150571
OG5_126714	OG5_127324	OG5_127942	OG5_128583	OG5_131529	OG5_143224	OG5_150573
OG5_126724	OG5_127349	OG5_127944	OG5_128613	OG5_132039	OG5_143229	OG5_150581
OG5_126734	OG5_127353	OG5_127970	OG5_128636	OG5_132955	OG5_144050	OG5_150591
OG5_126735	OG5_127390	OG5_127984	OG5_128650	OG5_133351	OG5_144310	OG5_152655
OG5_126742	OG5_127411	OG5_127998	OG5_128652	OG5_134455	OG5_144710	OG5_152790
OG5_126764	OG5_127428	OG5_128010	OG5_128667	OG5_134579	OG5_145113	OG5_153013
OG5_126776	OG5_127434	OG5_128039	OG5_128694	OG5_135732	OG5_145124	OG5_153529
OG5_126784	OG5_127449	OG5_128040	OG5_128697	OG5_136225	OG5_146288	OG5_153537
OG5_126788	OG5_127451	OG5_128047	OG5_128814	OG5_137087	OG5_146323	OG5_153541
OG5_126839	OG5_127457	OG5_128050	OG5_128828	OG5_137211	OG5_146350	OG5_153563
OG5_126852	OG5_127480	OG5_128053	OG5_128834	OG5_137246	OG5_146494	OG5_153564
OG5_126865	OG5_127512	OG5_128073	OG5_128892	OG5_138354	OG5_146589	OG5_153565
OG5_126882	OG5_127513	OG5_128081	OG5_128974	OG5_138480	OG5_146727	OG5_153611
OG5_126889	OG5_127562	OG5_128083	OG5_128978	OG5_138696	OG5_146787	OG5_153617
OG5_126897	OG5_127572	OG5_128086	OG5_129018	OG5_138808	OG5_147127	OG5_153618
OG5_126901	OG5_127575	OG5_128093	OG5_129068	OG5_139311	OG5_147285	OG5_153619
OG5_126932	OG5_127598	OG5_128179	OG5_129083	OG5_139591	OG5_147402	OG5_153627
OG5_126952	OG5_127606	OG5_128181	OG5_129091	OG5_139602	OG5_147418	OG5_155210
OG5_126960	OG5_127622	OG5_128190	OG5_129131	OG5_139975	OG5_147429	OG5_155292
OG5_126971	OG5_127633	OG5_128200	OG5_129167	OG5_139980	OG5_147441	OG5_156239
OG5_126983	OG5_127643	OG5_128203	OG5_129173	OG5_140089	OG5_147446	OG5_156793
OG5_126993	OG5_127675	OG5_128210	OG5_129196	OG5_140270	OG5_147457	OG5_156826
OG5_126994	OG5_127685	OG5_128212	OG5_129228	OG5_140509	OG5_147458	OG5_156833
OG5_126998	OG5_127690	OG5_128215	OG5_129252	OG5_140536	OG5_147462	OG5_158126
OG5_127002	OG5_127702	OG5_128222	OG5_129362	OG5_140540	OG5_147467	OG5_158593
OG5_127016	OG5_127716	OG5_128229	OG5_129418	OG5_140542	OG5_147469	OG5_160404
OG5_127036	OG5_127732	OG5_128247	OG5_129462	OG5_141018	OG5_147470	OG5_160569
OG5_127038	OG5_127734	OG5_128259	OG5_129558	OG5_141138	OG5_147471	OG5_160573
OG5_127044	OG5_127749	OG5_128264	OG5_129568	OG5_141175	OG5_147638	OG5_160584
OG5_127058	OG5_127760	OG5_128302	OG5_129610	OG5_141200	OG5_149434	OG5_160602
OG5_127068	OG5_127761	OG5_128306	OG5_129701	OG5_141228	OG5_149560	OG5_160659
OG5_127077	OG5_127781	OG5_128323	OG5_129811	OG5_141231	OG5_149657	OG5_162077
OG5_127086	OG5_127790	OG5_128335	OG5_129825	OG5_141425	OG5_149676	OG5_162099
OG5_127103	OG5_127796	OG5_128338	OG5_129889	OG5_141599	OG5_150124	OG5_162107
OG5_127107	OG5_127804	OG5_128358	OG5_129939	OG5_141603	OG5_150521	OG5_163889
OG5_127125	OG5_127822	OG5_128366	OG5_130096	OG5_141747	OG5_150523	OG5_163933
OG5_127138	OG5_127838	OG5_128399	OG5_130105	OG5_141753	OG5_150527	OG5_165875
OG5_127160	OG5_127848	OG5_128435	OG5_130121	OG5_141755	OG5_150529	OG5_167172
OG5_127162	OG5_127855	OG5_128472	OG5_130212	OG5_141761	OG5_150532	OG5_167323
OG5_127188	OG5_127886	OG5_128492	OG5_130409	OG5_142437	OG5_150557	OG5_171643

Tab. 7.7: Annotation der mitochondrialen DNA von *Myotis daubentonii*.

Gen	Start-position	Stop-position	Länge (bp)	Antikodon	Startkodon	Stopkodon	Strang
tRNA ^{Phe}	1	68	68	GAA			+
12S-rRNA	69	1.037	989				+
tRNA ^{Val}	1.038	1.105	68	UAC			+
16S-rRNA	1.105	2.675	1.571				+
tRNA ^{Leu}	2.677	2.751	75	UAA			+
nad1	2.757	3.713	957		ATG	TAA	+
tRNA ^{Ile}	3.713	3.781	69	GAU			+
tRNA ^{Gln}	3.779	3.852	74	UUG			-
tRNA ^{Met}	3.853	3.922	70	CAU			+
nad2	3.923	4.966	1.044		ATA	TAG	+
tRNA ^{Trp}	4.965	5.032	68	UCA			+
tRNA ^{Ala}	5.038	5.106	69	UGC			-
tRNA ^{Asn}	5.108	5.180	73	U ¹ TT			-
OriL	5.181	5.213	33				+
tRNA ^{Cys}	5.213	5.278	66	UCA			-
tRNA ^{Tyr}	5.279	5.345	67	UTA			-
cox1	5.347	6.891	1.545		ATG	TAG	+
tRNA ^{Ser}	6.906	6.974	69	UGA			-
tRNA ^{Asp}	6.982	7.048	67	GUC			+
cox2	7.049	7.732	684		ATG	TAG	+
tRNA ^{Lys}	7.736	7.802	67	UUU			+
atp8	7.804	8.007	204		ATG	TAA	+
atp6	7.965	8.645	681		ATG	TAA	+
cox3	8.645	9.428	783		ATG	T*	+
tRNA ^{Gly}	9.429	9.498	70	UCC			+
nad3	9.499	9.845	345		ATA	TA*	+
tRNA ^{Arg}	9.846	9.914	69	UCG			+
nad4L	9.916	10.212	298		ATG	TAA	+
nad4	10.206	11.573	1.368		ATG	T*	+
tRNA ^{His}	11.584	11.651	68	GUG			+
tRNA ^{Ser}	11.652	11.710	59	GCU			+
tRNA ^{Leu}	11.711	11.780	70	UAG			+
nad5	11.781	13.601	1.821		ATA	TAA	+
nad6	13.585	14.109	525		ATA	TAA	-
tRNA ^{Glu}	14.113	14.181	69	UUC			-
cytb	14.188	15.327	1.140		ATG	AGA	+
tRNA ^{Thr}	15.328	15.397	70	UGU			+
tRNA ^{Pro}	15.397	15.462	66	UGG			-
D-loop	18.464	17.149	1.686				+

* unvollständiges Stopkodon

Tab. 7.8: MtDNA Sequenzdaten zur Rekonstruktion der Phylogenie der Gattung *Myotis*.

Spezies	Verbreitung	Accessionnr.	Spezies	Verbreitung	Accessionnr.
<i>Murina leucogaster</i>	paläarktisch	NC_025949	<i>Myotis leibii</i>	nearktisch	NC_036321
<i>Murina ussuriensis</i>	paläarktisch	NC_021119	<i>Myotis lucifugus</i>	nearktisch	NC_029849
<i>Myotis albescens</i>	neotropisch	NC_036327	<i>Myotis macrodactylus</i>	paläarktisch	NC_022694
<i>Myotis atacamensis</i>	neotropisch	NC_036324	<i>Myotis martiniquensis</i>	neotropisch	NC_036328
<i>Myotis auriculus</i>	nearktisch	NC_036320	<i>Myotis melanorhinus</i>	nearktisch	NC_036322
<i>Myotis bechsteinii</i>	paläarktisch	NC_034227	<i>Myotis muricola</i>	orientalisch	NC_029422
<i>Myotis bombinus</i>	paläarktisch	NC_029342	<i>Myotis myotis</i>	paläarktisch	NC_029346
<i>Myotis brandtii</i>	paläarktisch	NC_025308	<i>Myotis nigricans</i>	neotropisch	NC_036318
<i>Myotis davidii</i>	paläarktisch	NC_025568	<i>Myotis oxyotus</i>	neotropisch	NC_036316
<i>Myotis dominicensis</i>	neotropisch	NC_036312	<i>Myotis petax</i>	paläarktisch	NC_027973
<i>Myotis evotis</i>	nearktisch	NC_036313	<i>Myotis riparius</i>	neotropisch	NC_036317
<i>Myotis formosus</i>	orientalisch	NC_015828	<i>Myotis ruber</i>	neotropisch	NC_036315
<i>Myotis borsfieldii</i>	orientalisch	NC_036325	<i>Myotis thysanodes</i>	nearktisch	NC_036323
<i>Myotis ikonnikovi</i>	paläarktisch	NC_022698	<i>Myotis volans</i>	nearktisch	NC_036326
<i>Myotis keaysi</i>	neotropisch	NC_036314	<i>Myotis yumanensis</i>	nearktisch	NC_036319

Tab. 7.9: Anteil repetitiver DNA im Genom von *M. daubentonii*.

	Anzahl der Elemente	Gesamt Größe [bp]	Anteil der DNA [%]
SINEs	323.019	47.886.266	3,32
ALUs	48	4.634	0,00
MIRs	316.841	46.967.178	3,26
LINEs	651.408	236.812.474	16,43
LINE1	37.2001	16.4584.255	11,42
LINE2	245.542	61.883.208	4,29
L3/CR1	23.286	6.767.305	0,47
LTR elements	245.240	78.622.592	5,46
ERV1	65.880	23.145.958	1,61
ERV1-MaLRs	111.987	3.4019.755	2,36
ERV_classI	38.935	12.405.391	0,86
ERV_classII	2.235	675.433	0,05
DNA elements	300.860	56.462.343	3,92
hAT-Charlie	164.319	28.014.064	1,94
TcMar-Tigger	62.669	14.416.438	1,00
Unclassified	6.497	1.349.610	0,09
Total interspersed repeats		421.133.285	29,22
Small RNA	133.301	8.229.890	0,57
Satellites	2.06	22.706	0,00
Simple repeats	439.384	18.620.299	1,29
Low complexity	114.908	5.387.052	0,37

Tab. 7.10: Vergleich der transponierbaren Elemente im Säugetiergenom.

	<i>P. vampyrus</i>	<i>R. sinicus</i>	<i>H. armiger</i>	<i>M. davidii</i>	<i>M. daubentonii</i>	<i>H. sapiens</i>	<i>M. musculus</i>
DNA	3,45	3,64	3,31	3,57	3,92	3,3	2,3
LINE	16,43	16,6	14,95	14,66	16,43	17,5	18,2
LRT	4,69	4,83	4,38	4,83	5,46	8,3	10,4
SINE	3,15	3,36	3,15	3,00	3,32	11,3	6,1
Andere	0,05	0,05	0,05	0,08	0,09	1,1	1,9
Gesamt	27,77	28,49	25,84	26,14	29,22	41,5	38,9

(nach Dong et al., 2017)

Tab. 7.11: Kodonverwendung in der mtDNA von *P. murinus*.

TTT	F	0,79	7,31(81)	TCT	S	0,33	2,89(32)	TAT	Y	0,87	4,97(55)	TGT	C	0,80	1,45(16)
TTC	F	0,21	1,99(22)	TCC	S	0,02	0,18(2)	TAC	Y	0,13	0,72(8)	TGC	C	0,20	0,36(4)
TTA	L	0,86	12,57(139)	TCA	S	0,35	3,07(34)	TAA	*	1,00	0,27(3)	TGA	*	0,00	0,0(0)
TTG	L	0,02	0,27(3)	TCG	S	0,00	0,0(0)	TAG	*	0,00	0,0(0)	TGG	W	1,00	1,9 (22)
CTT	L	0,05	0,72(8)	CCT	P	0,43	1,63(18)	CAT	H	0,94	2,89(32)	CGT	R	0,23	0,45(5)
CTC	L	0,01	0,09(1)	CCC	P	0,02	0,09(1)	CAC	H	0,06	0,18(2)	CGC	R	0,00	0,0(0)
CTA	L	0,07	5,34(11)	CCA	P	0,55	2,08 (23)	CAA	Q	1,00	1,81(20)	CGA	R	0,00	0,0(0)
CTG	L	0,00	2,08 (0)	CCG	P	0,00	0,0(0)	CAG	Q	0,00	0,0(0)	CGG	R	0,00	0,0(0)
ATT	I	0,55	6,51(72)	ACT	T	0,35	2,35(26)	AAT	N	0,93	3,35(37)	AGT	S	0,29	2,53(28)
ATC	I	0,01	0,09(1)	ACC	T	0,01	0,09(1)	AAC	N	0,07	0,27(3)	AGC	S	0,02	0,18(2)
ATA	I	0,45	5,34 (59)	ACA	T	0,64	4,25(47)	AAA	K	1,00	1,45(16)	AGA	R	0,77	1,54(17)
ATG	M	1,00	2,06(23)	ACG	T	0,00	0,0(0)	AAG	K	0,00	0,0(0)	AGG	R	0,00	0,0 (0)
GTT	V	0,40	2,26(25)	GCT	A	0,54	2,62(29)	GAT	D	0,89	1,54(17)	GGT	G	0,63	4,07(45)
GTC	V	0,00	0,0 (0)	GCC	A	0,00	0,0 (0)	GAC	D	0,11	0,18(2)	GGC	G	0,00	0,0(0)
GTA	V	0,60	3,44(38)	GCA	A	0,46	2,26(25)	GAA	E	0,96	2,17(24)	GGA	G	0,37	2,35(26)
GTG	V	0,00	0,0(0)	GCG	A	0,00	0,0(0)	GAG	E	0,04	0,09(1)	GGG	G	0,00	0,0(0)

Gezeigt sind jeweils Kodon, entsprechende Aminosäure, relative Häufigkeit der Verwendung von synonymen Kodons, sowie Häufigkeit der Kodons gemessen an den Genen cytb, cox1 und cox3.

Tab. 7.12: Eigenschaften der Primersets zur selektiven Amplifizierung von *P. murinus*.

SWGA-Set	Score	Primer- anzahl	Gini- Index	Mittlere Distanz (bp) <i>Polychromophilus</i>	Mittlere Distanz (bp) <i>Myotis</i>
Set1	0,00063	2	0,63	527	520.118
Set2	0,00063	2	0,63	527	520.118
Set3	0,00064	2	0,63	528	519.718
Set4	0,00064	2	0,63	528	519.718
Set5	0,001296	3	0,62	518	251.215
Set6	0,00196	4	0,62	494	157.185
Set7	0,00185	4	0,62	507	170.015
Set8	0,00170	4	0,62	512	187.957

Score: Die Einstufung des Primersets erfolgt nach dem Verhältnis zwischen Bindungshäufigkeit der Primer im Ziel und Hintergrundgenom; Gini-Index ist ein Maß für die gleichmäßige Verteilung der Primerbindungsstellen auf der Zielsequenz, wobei 1 extrem ungleichmäßig ist und 0 eine einheitliche Verteilung repräsentiert.

Tab. 7.13: MtDNA Sequenzdaten zur Rekonstruktion der Phylogenie der Apikomplexa.

Spezies	Wirt	mtDNA	Accessionnr.		
			Cox3	Cox1	cytb
<i>Haemoproteus balmoralis</i>	<i>Luscinia luscinia</i>	KY653800	ATD12984	ATD12985	ATD12986
<i>Haemoproteus belopolskyi</i>	<i>Hippolais icterina</i>	KY653790	ATD12954	ATD12955	ATD12956
<i>Haemoproteus coatneyi</i>	<i>Arremon brunneinucha</i>	KT698210	AMS24000	AMS24001	AMS24002
<i>Haemoproteus columbae</i>	<i>Columba livia</i>	KY653761	ATD12867"	ATD12868"	ATD12869
<i>Haemoproteus erythrogravidus</i>	<i>Zonotrichia capensis</i>	KT698209	AMS23997	AMS23998	AMS23999
<i>Haemoproteus lanii</i>	<i>Lanius collurio</i>	KY653787	ATD12945	ATD12946	ATD12947
<i>Haemoproteus pastoris</i>	<i>Sturnus vulgaris</i>	KY653793	ATD12963	ATD12964	ATD12965
<i>Haemoproteus sacharovi</i>	<i>Zenaida macroura</i>	KY653811	ATD13017	ATD13018	ATD13019
<i>Haemoproteus tartakovskyi</i>	<i>Spinus spinus</i>	KY653810	ATD13014	ATD13015	ATD13016
<i>Haemoproteus witii</i>	<i>Eriocnemis derbyi</i>	KY653794	ATD12966	ATD12967	ATD12968
<i>Hepatocystis</i> sp.	<i>Macaca fascicularis</i>	KY653783	ATD12933	ATD12934	ATD12935
<i>Hepatocystis</i> sp.	<i>Macaca fascicularis</i>	KY653782	ATD12930	ATD12931	ATD12932
<i>Hepatocystis</i> sp.	<i>Pteropus hypomelanus</i>	FJ168565	ACL80167	ACL80168	ACL80169
<i>Leucocytozoon danilevskyi</i>	<i>Asio otus</i>	KY653781	ATD12927	ATD12928	ATD12929
<i>Leucocytozoon dubrenili</i>	<i>Turdus merula</i>	KY653795	ATD12969	ATD12970	ATD12971
<i>Leucocytozoon fringillinarum</i>	<i>Turdus fuscater</i>	KT162004	ANB66447	ANB66448	ANB66449
<i>Leucocytozoon pterotennisi</i>	<i>Grallaria ruficapilla</i>	KM610046	AJA33414	AJA33415	AJA33416
<i>Nycteria gabonensis</i>	<i>Rhinolophus alycone</i>	KX090647	AOF41412	AOF41413	AOF41414
<i>Nycteria heischei</i>	<i>Megaderma spasma</i>	KX090648	AOF41415	AOF41416	AOF41417
<i>Nycteria medusiformis</i>	<i>Nycterus grandis</i>	KX090645	AOF41408	AOF41407	AOF41406
<i>Nycteria</i> sp.	<i>Nycterus</i> sp.	KX090646	AOF41411	AOF41410	AOF41409
<i>Plasmodium bubalis</i>	<i>Bubalus bubalis</i>	LC090213	BAU68591	BAU68592	BAU68593
<i>Plasmodium chiricabuae</i>	<i>Sceloporus jarrovi</i>	KY653779	ATD12921	ATD12922	ATD12923
<i>Plasmodium cynomolgi</i>	<i>Macaca nemestrina</i>	AB444129	BAI82327	BAI82328	BAI82329
<i>Plasmodium elongatum</i>	<i>Acrocephalus scirpaceus</i>	KY653801	ATD12987	ATD12988	ATD12989
<i>Plasmodium falciparum</i>	<i>Homo sapiens</i>	AY282930	AAP57986	AAP57984	AAP57985
<i>Plasmodium homocircumflexum</i>	<i>Lanius collurio</i>	KY653784	ATD12936	ATD12937	ATD12938
<i>Plasmodium inui</i>	<i>Macaca sinica</i>	GQ355483	ADD00623	ADD00624	ADD00625
<i>Plasmodium juxtanucleare</i>	<i>Turdus subalaris</i>	MG598398	AWQ13383	AWQ13384	AWQ13385
<i>Plasmodium kentropyxi</i>	<i>Cnemidophorus gramivagus</i>	KY653753	ATD12843	ATD12844	ATD12845
<i>Plasmodium knowlesi</i>	<i>Macaca nemestrina</i>	NC_007232	YP_272028	YP_272029	YP_272030
<i>Plasmodium lutzi</i>	<i>Turdus fuscater</i>	KC138226	AGA63740	AGA63741	AGA63742
<i>Plasmodium malariae</i>	<i>Homo sapiens</i>	AB354570	BAG68530	BAG68531	BAG68532
<i>Plasmodium ovale</i>	<i>Homo sapiens</i>	HQ712052	AFC62059	AFC62060	AFC62061
<i>Plasmodium simiovale</i>	<i>Macaca sinica</i>	AB434920	BAG68557	BAG68558	BAG68559
<i>Plasmodium simium</i>	<i>Simia belzebub</i>	KY923297	AVP52753	AVP52754	AVP52755
<i>Plasmodium</i> sp.	<i>Capra aegagrus hircus</i>	LC090215	BBA78539	BBA78540	BBA78541
<i>Plasmodium</i> sp.	<i>Bubalus bubalis</i>	LC090214	BAU68594	BAU68595	BAU68596
<i>Plasmodium</i> sp.	<i>Plica cf. plica</i>	KY653796	ATD12972	ATD12973	ATD12974
<i>Plasmodium unalis</i>	<i>Turdus fuscater</i>	KY653813	ATD13023	ATD13024	ATD13025
<i>Plasmodium vaughani</i>	<i>Turdus merula</i>	KY653792	ATD12960	ATD12961	ATD12962
<i>Plasmodium vivax</i>	<i>Homo sapiens</i>	AY598140	AAV26891	AAV26892	AAV26893

Tab. 7.14: Annotation der Apikoplast-DNA von *Polychromohilus murinus*.

Gen	Start-position	Stop-position	Länge (bp)	Anti- kodon	Start- kodon	Stop- kodon	Strang
rps4	1	630	630		ATG	TGA	+
tRNA ^{His}	640	713	74	GUG			+
tRNA ^{Cys}	730	803	74	GCA			+
*tRNA ^{Lys}	812	1083	75	AAT			+
tRNA ^{Met}	1.052	1.140	89	CAU			+
tRNA ^{Tyr}	1153	1.240	88	GUA			+
tRNA ^{Ser}	1.254	1.348	95	GCU			+
tRNA ^{Asp}	1.351	1.426	76	GUC			+
tRNA ^{Lys}	1.444	1.517	74	UUU			+
tRNA ^{Glu}	1.523	1.593	71	UUC			+
tRNA ^{Pro}	1.602	1.675	74	UGG			+
rpl4	1.701	2.303	603		ATG	TAA	+
rpl3	2.312	2.530	219		ATG	TAA	+
rpl2	2.538	3.269	732		ATG	TAA	+
rps19	3.286	3.561	276		ATG	TAA	+
rsp7	3.563	4.207	645		ATG	TAA	+
rpl16	4.224	4.619	396		ATG	TAA	+
rsp17	4.633	4.854	222		ATG	TGA	+
rpl14	4.851	5.213	363		ATG	TGA	+
rps8	5.221	5.607	387		ATG	TGA	+
Rps6	5.628	6.134	507		ATG	TGA	
rps5	6.131	6.823	693		ATG	TAA	+
ORF91	6.851	7.117	267		ATG	TAA	+
rpl36	7.083	7.193	111		ATG	TGA	+
rps11	7.190	7.597	408		ATG	TAA	+
rps12	7.599	7.967	369		ATG	TAA	+
rps7	8.004	8.408	405		ATG	TAA	+
tufa	8.413	9.642	1.230		ATG	TAA	+
ORF78	9.650	9.901	252		ATG	TGA	+
tRNA ^{Phe}	9.878	9949	72	CTT			-
tRNA ^{Gln}	9.971	10.042	72	UUG			+
tRNA ^{Gly}	10.046	10.119	74	ACC			+
tRNA ^{Trp}	10.128	10.200	73	CCA			+
ORF129	10.223	10.594	372		ATG	TGA	+
clpc	10.591	12.894	2.304		ATG	TAA	+
tRNA ^{Gly}	12.908	12.979	72	UGG			+
ORF79	13.007	13.243	237		ATG	TAA	+
tRNA ^{Ser}	13.247	13.334	88	UGA			+
ORF105	13.358	13.666	309		ATG	TAA	+
rps2	13.653	14.339	687		ATG	TAA	-
rpoc2b	14.353	15.621	1.269		ATG	TAA	-
rpoc2a	15.653	17.233	1.581		ATG	TGA	-
rpoc1	17.249	18.961	1.713		ATG	TAA	-
ropb	18.965	22.048	3.084		ATG	TGA	-
ORF101	22.051	22.374	324		ATG	TAA	-
ORF51	22.388	22.549	162		ATG	TAA	-
sufb	22.565	23.974	1.410		ATG	TAA	-
tRNA ^{Thr}	24.002	24.074	73	UGU			-
23S-rRNA	24.135; 26.541	26.770; 29.176	2.636				-; +
tRNA ^{Met}	26.808; 26.430	26.881; 26.503	74	CAU			-; +
tRNA ^{Arg}	26.888; 26.350	26.961; 26.423	74	ACU			-; +
tRNA ^{Val}	26.978; 26.260	27.051; 26.333	74	UAC			+; -
tRNA ^{Arg}	27.058; 26.182	27.129; 26.253	74	UCU			+; -
tRNA ^{Leu}	27.144; 26.088	27.232; 26.167	80	UAG			+; -
tRNA ^{Asp}	27.233; 26.007	27.304; 26.078	72	GUU			-; +
tRNA ^{Ala}	27.337; 25.902	27.409; 25.974	73	UGC			-; +
16S-rRNA	27.447; 24.397	28.914; 25.864	1.468				-; +
tRNA ^{Ile}	29.054; 24.187	29.129; 24.257	71	GAU			-; +
tRNA ^{Thr}	29.237	29.309	73	UGU			+

* enthält 141 bp Intron; grau alternative Positionen der zweiten Isoform

Tab. 7.15: Kodonverwendung im Apikoplast-Genom von *P. murinus*.

TTT	F	0,98	6,48(487)	TCT	S	0,43	2,00(150)	TAT	Y	0,99	9,41(707)	TGT	C	0,99	1,65(80)
TTC	F	0,02	0,13(10)	TCC	S	0,01	0,03(2)	TAC	Y	0,01	0,13(10)	TGC	C	0,01	0,01(1)
TTA	L	0,94	11,81(887)	TCA	S	0,25	1,17(88)	TAA	*	0,75	0,32(24)	TGA	*	0,25	0,11(8)
TTG	L	0,02	0,03(23)	TCG	S	0,01	0,04(3)	TAG	*	0,00	0,00(0)	TGG	W	1,00	0,39(29)
CTT	L	0,02	0,28(21)	CCT	P	0,77	1,01(82)	CAT	H	0,96	0,87(65)	CGT	R	0,12	0,16(12)
CTC	L	0,00	0,00(0)	CCC	P	0,00	0,00(0)	CAC	H	0,04	0,04(3)	CGC	R	0,00	0,0(0)
CTA	L	0,01	0,11(8)	CCA	P	0,22	0,32(24)	CAA	Q	0,97	1,70(128)	CGA	R	0,00	0,0(0)
CTG	L	0,00	0,00(0)	CCG	P	0,01	0,01(1)	CAG	Q	0,03	0,05(4)	CGG	R	0,00	0,0(0)
ATT	I	0,41	6,86(515)	ACT	T	0,46	1,32(99)	AAT	N	0,99	1,77(133)	AGT	S	0,30	1,39(105)
ATC	I	0,01	0,08(6)	ACC	T	0,00	0,0(0)	AAC	N	0,01	2,29(172)	AGC	S	0,01	0,03(2)
ATA	I	0,58	0,97(727)	ACA	T	0,53	1,32(99)	AAA	K	0,96	13,67(1027)	AGA	R	0,88	1,19(90)
ATG	M	1,00	1,16(87)	ACG	T	0,01	0,01(1)	AAG	K	0,04	0,51(38)	AGG	R	0,00	0,0(0)
GTT	V	0,43	0,92(69)	GCT	A	0,65	0,28(42)	GAT	D	1,00	1,71(133)	GGT	G	0,43	1,3(96)
GTC	V	0,00	0,0(0)	GCC	A	0,00	0,0(0)	GAC	D	0,0	0,0	GGC	G	0,00	0,0(0)
GTA	V	0,56	1,19(89)	GCA	A	0,32	2,26(21)	GAA	E	0,98	2,29(172)	GGA	G	0,55	1,62(122)
GTG	V	0,01	0,03(2)	GCG	A	0,03	0,03(2)	GAG	E	0,02	0,05(4)	GGG	G	0,01	0,04(3)

Gezeigt sind jeweils Kodon, entsprechende Aminosäure, relative Häufigkeit der Verwendung von synonymen Kodons, sowie Häufigkeit der Kodons gemessen an allen annotierten Genen der plDNA.

Tab. 7.16: PIDNA Sequenzdaten zur Rekonstruktion der Phylogenie der Apikomplexa.

Spezies	Stamm	Größe (bp)	GC-Gehalt	Accessionnr.
<i>Babesia bovis</i>	T2Bo	35.107	22 %	AAXT01000007.1
<i>Babesia microti</i>	RI	28.657	14 %	LK028575.1
<i>Babesia orientalis</i>	Wuhan	33.200	21 %	KT428643.1
<i>Hepatozoon sp.</i>	Ex <i>P. tephrosceles</i> isolate RC106	~29.276	13 %	PDMG01016242.1 PDMG01025106.1 PDMG01043485.1
<i>Leucocytozoon canlleryi</i>	Niigata	34.779	15 %	AP013071.1
<i>Plasmodium berghei</i>	ANKA	29.464	14 %	AB649421.1
<i>Plasmodium brasilianum</i>	Bolivian I	28.948	13 %	CM007351.1
<i>Plasmodium chabaudi</i>	Chabaudi CB	29.623	14 %	HF563595.1
<i>Plasmodium coatneyi</i>	CDC	29.055	13 %	AB649420.1
<i>Plasmodium cynomolgi</i>	M	34.521	14 %	LT841394
<i>Plasmodium falciparum</i>	3D7	34.250	14 %	LN999985.1
<i>Plasmodium fragile</i>	nilgiri	35.616	13 %	KQ001809
<i>Plasmodium gaboni</i>	SY75	29.387	13 %	LVLB01000001.1
<i>Plasmodium gallinaceum</i>	8A	29.456	13 %	LN835293.1
<i>Plasmodium gonderi</i>	ATCC 30045	34.515	14 %	BDQF01000742.1
<i>Plasmodium knowlesi</i>	Malayan Pk1 (A+)	32.097	14 %	MF373420.1
<i>Plasmodium malariae</i>	kA	34.324	14 %	LT594636.1
<i>Plasmodium ovale</i>	wallikeri	29.270	13 %	KX611805.1
<i>Plasmodium reichenowi</i>	SY57	24.235	12 %	CM003883.1
<i>Plasmodium relictum</i>	SGS1	29.365	13 %	LN835310.1
<i>Plasmodium vinckei</i>	vinckei	29.543	14 %	LR215071.1
<i>Plasmodium vivax</i>	Salvador I	29.093	13 %	AB649419.1
<i>Plasmodium yoelii</i>	17NXL	29.227	14 %	AB649422.1

Tab. 7.17: Kodonverwendung in der genomischen DNA von *P. murinus*.

TTT	F	0,90	4,60	TCT	S	0,49	2,39	TAT	Y	0,92	5,46	TGT	C	0,85	1,42
TTC	F	0,10	0,49	TCC	S	0,12	0,59	TAC	Y	0,08	0,50	TAC	C	0,15	0,25
TTA	L	0,71	6,16	TCA	S	0,85	1,64	TAA	*	0,75	0,15	TGA	*	0,14	0,03
TTG	L	0,11	0,96	TCG	S	0,15	0,29	TAG	*	0,10	0,02	TGG	W	1,00	0,51
CIT	L	0,09	0,77	CCT	P	0,48	1,03	CAT	H	0,91	2,00	CGT	R	0,09	0,20
CTC	L	0,01	0,08	CCC	P	0,09	0,18	CAC	H	0,09	0,20	CGC	R	0,01	0,02
CTA	L	0,07	0,59	CCA	P	0,41	0,87	CAA	Q	0,91	2,79	CGA	R	0,18	0,30
CTG	L	0,01	0,05	CCG	P	0,03	0,07	CAG	Q	0,09	0,28	CGG	R	0,01	0,03
ATT	I	0,46	5,00	ACT	T	0,41	1,71	AAT	N	0,89	11,86	AGT	S	0,85	1,66
ATC	I	0,06	0,64	ACC	T	0,10	0,40	AAC	N	0,11	1,53	AGC	S	0,15	0,29
ATA	I	0,48	5,22	ACA	T	0,44	1,85	AAA	K	0,86	9,88	AGA	R	0,64	1,41
ATG	M	1,00	1,86	ACG	T	0,06	0,24	AAG	K	0,14	1,57	AGG	R	0,12	0,26
GTT	V	0,46	1,57	GCT	A	0,43	0,80	GAT	D	0,91	5,17	GGT	G	0,30	0,73
GTC	V	0,04	0,14	GCC	A	0,06	0,11	GAC	D	0,09	0,50	GGC	G	0,03	0,07
GTA	V	0,43	1,46	GCA	A	0,47	0,87	GAA	E	0,89	5,57	GGA	G	0,56	1,38
GTG	V	0,07	0,24	GCG	A	0,04	0,07	GAG	E	0,11	0,72	GGG	G	0,12	0,28

Gezeigt sind jeweils Kodon, entsprechende Aminosäure, relative Häufigkeit der Verwendung von synonymen Kodons, sowie Häufigkeit der Kodons der kodierenden Genen.

Tab. 7.18: Exportierte Proteine von *Polychromophilus murinus*.

<i>P. murinus</i>	PEXEL	<i>P. falciparum</i>	PEXEL	Protein
Konservierte Proteine der Aconoidasida				
Pmur00073	RLLSE	PF3D7_1311800	RLLKS	M1-family alanyl aminopeptidase
Pmur00442		PF3D7_1366400		rhoptry protein RHOP148
Pmur00884		PF3D7_1452000		rhoptry neck protein 2
Pmur01208	REINE	PF3D7_1133300	RFLYS	LEM3/CDC50 family protein
Pmur01209		PF3D7_1133400		apical membrane antigen 1
Pmur01480		PF3D7_1430200	RILSD	plasmepsin IX
Pmur01751	KSIKE	PF3D7_0506900		rhomboid protease ROM4
Pmur02179	RLLID	PF3D7_0817700		rhoptry neck protein 5
Pmur02358	KNISE	PF3D7_1018300	RKLKE	conserved <i>Plasmodium</i> protein
Pmur02450	RILSE	PF3D7_1033000	RILSE	conserved <i>Plasmodium</i> protein
Pmur02499	RKLLQ	PF3D7_0718100		exported serine/threonine protein kinase
Pmur02874		PF3D7_0318100		stomatin-like protein
Pmur02930	K'TILE	PF3D7_0321200	KILLE	UDP-N-acetylglucosamine--dolichyl-phosphate N-acetylglucosaminephosphotransferase
Pmur03079		PF3D7_0825700	RNYSE	conserved <i>Plasmodium</i> protein
Konservierte Proteine der Haemosporida				
augustus-Contig_0007099-processed-gene-0.2		PF3D7_0404500		6-cysteine protein
Pmur00081	KDIIE	PF3D7_1313300	RDLID	peptidyl-prolyl cis-trans isomerase
Pmur00247	KNIPQ	PF3D7_1339900	RAFLD	ABC transporter B family member 5
Pmur00396	RILSA	PF3D7_1360000		conserved <i>Plasmodium</i> protein
Pmur00524		PF3D7_1210700	RILPS	conserved <i>Plasmodium</i> protein
Pmur00535		PF3D7_1212100	RILKE	peripheral plastid protein 1
Pmur00569		PF3D7_1216600		cell traversal protein for ookinetes and sporozoites
Pmur00570	KAIHE	PF3D7_1216700	RNLEQ	perforin-like protein 2
Pmur00580	KLILD	PF3D7_1218000		thrombospondin-related apical membrane protein
Pmur00603		PF3D7_1222100		conserved <i>Plasmodium</i> protein

<i>P. murinus</i>	PEXEL	<i>P. falciparum</i>	PEXEL	Protein
Pmur00681	KSLYD	PF3D7_1233800	RCLCA	conserved <i>Plasmodium</i> protein
Pmur00932	KMLND	PF3D7_1458000	RVLLE	cysteine proteinase falcipain 1
Pmur00973	RDLGE	PF3D7_1463900	RILFE	EF-hand calcium-binding domain-containing protein transcript product
Pmur00981	RNLSE	PF3D7_1464600	RILDE	serine/threonine protein phosphatase UIS2
Pmur01097	KDLCQ	PF3D7_1114100		rhomboid protease ROM1
Pmur01211	KELTQ	PF3D7_1133700		FHA domain-containing protein
Pmur01235	RSLSE	PF3D7_1136900		subtilisin-like protease 2
Pmur01431		PF3D7_1418100		liver specific protein 1
Pmur01755		PF3D7_0507500		subtilisin-like protease 1
Pmur02103	KKIRE	PF3D7_0805700		serine/threonine protein kinase, FIKK family
Pmur02146	KDISE	PF3D7_0811600	RSIDE	conserved <i>Plasmodium</i> protein
Pmur02172		PF3D7_0815500		conserved <i>Plasmodium</i> protein
Pmur02237		PF3D7_0828800		GPI-anchored micronemal antigen
Pmur02253	KKLHE	PF3D7_0404400		6-cysteine protein
Pmur02254	RIIYD	PF3D7_0404700		dipeptidyl aminopeptidase 3
Pmur02262	[KSINN E]	PF3D7_0405900		apical sushi protein
Pmur02307	KILQQ	PF3D7_0414900		armadillo-domain containing rhoptry protein
Pmur02345	KNIQD	PF3D7_0419700		apical merozoite protein
Pmur02572		PF3D7_0207300		serine repeat antigen 8
Pmur02587		PF3D7_0210600		conserved <i>Plasmodium</i> protein
Pmur02772	RLIID	PF3D7_0310400	RLLVD	parasite-infected erythrocyte surface protein
Pmur02859	KHIID	PF3D7_0315200		circumsporozoite- and TRAP-related protein
Pmur02861		PF3D7_0316000		microneme associated antigen
Pmur02950	KCLRE	PF3D7_0323400		Rh5 interacting protein
Pmur03045		PF3D7_0104200	RILKE	StAR-related lipid transfer protein
Pmur03100	RILYD	PF3D7_1316200	RILYE	ADP-ribosylation factor
Konservierte Proteine der Plasmodiiae				
Pmur00294		PF3D7_1346800		6-cysteine protein
Pmur01458		PF3D7_1425900	RNLIE	conserved <i>Plasmodium</i> protein
Pmur02692	KEIAE	PF3D7_1012200		rhoptry associated adhesin
Pmur02706	KNLPE	PF3D7_1013700	RILTE	conserved <i>Plasmodium</i> protein
Pmur03089	RSLNF	PF3D7_1007000	RILSV	conserved membrane protein
maker- Contig_0006937-		PF3D7_0523900	RNLDH	conserved <i>Plasmodium</i> membrane protein
Pmur00293	[KLLPK E]	PF3D7_1346700		6-cysteine protein
Pmur01979		PF3D7_0612700		6-cysteine protein
Pmur02255		PF3D7_0404900		6-cysteine protein
Pmur02576		PF3D7_0208900		6-cysteine protein
Pmur03120	RILNE	PF3D7_0104000	RLJNE	thrombospondin-related sporozoite protein

In Grau: Proteine, die in den Gattungen *Polychromophilus* und *Plasmodium* konserviert sind jedoch nicht in *Hepaticocystis*.

Tab. 7.19: Konservierte Proteine der Plasmodiidae, die nicht in fledermausinfizierenden Gattungen enthalten sind

Pfa ID	Protein	Pfa ID	Protein
PF3D7_0104000	thrombospondin-related sporozoite	PF3D7_1030300	conserved <i>Plasmodium</i> protein
PF3D7_0106200	EF hand domain-containing protein	PF3D7_1031300	conserved <i>Plasmodium</i> protein
PF3D7_0112100	conserved <i>Plasmodium</i> protein	PF3D7_1105300	conserved <i>Plasmodium</i> protein
PF3D7_0204300	conserved <i>Plasmodium</i> protein	PF3D7_1114600	conserved <i>Plasmodium</i> protein
PF3D7_0209000	6-cysteine protein P230	PF3D7_1117900	conserved <i>Plasmodium</i> protein
PF3D7_0215600	conserved <i>Plasmodium</i> protein	PF3D7_1120200	conserved <i>Plasmodium</i> protein
PF3D7_0216100	conserved <i>Plasmodium</i> protein	PF3D7_1121600	exported protein 1
PF3D7_0220800	cytoadherence linked asexual protein 2	PF3D7_1126700	autophagy-related protein 23
PF3D7_0302200	cytoadherence linked asexual protein 3	PF3D7_1127700	conserved <i>Plasmodium</i> protein
PF3D7_0302500	cytoadherence linked asexual protein 3	PF3D7_1129100	parasitophorous vacuolar protein 1 sporozoite surface protein essential for liver stage development
PF3D7_0309900	conserved <i>Plasmodium</i> protein	PF3D7_1137800	conserved <i>Plasmodium</i> protein
PF3D7_0313000	conserved <i>Plasmodium</i> protein	PF3D7_1145000	conserved <i>Plasmodium</i> protein
PF3D7_0315900	conserved <i>Plasmodium</i> protein	PF3D7_1147600	conserved <i>Plasmodium</i> protein
PF3D7_0321300	conserved <i>Plasmodium</i> protein	PF3D7_1203100	conserved <i>Plasmodium</i> protein
PF3D7_0404800	conserved <i>Plasmodium</i> protein	PF3D7_1204100	conserved <i>Plasmodium</i> protein
PF3D7_0404900	6-cysteine protein P41	PF3D7_1214500	conserved <i>Plasmodium</i> protein
PF3D7_0405500	conserved <i>Plasmodium</i> protein	PF3D7_1221200	conserved <i>Plasmodium</i> protein
PF3D7_0406500	NYN domain-containing protein	PF3D7_1224200	BRO1 domain-containing protein, putative
PF3D7_0407400	conserved <i>Plasmodium</i> protein	PF3D7_1226100	haloacid dehalogenase-like hydrolase
PF3D7_0407800	conserved <i>Plasmodium</i> protein	PF3D7_1229300	conserved <i>Plasmodium</i> protein
PF3D7_0419100	conserved <i>Plasmodium</i> protein	PF3D7_1232600	conserved <i>Plasmodium</i> protein
PF3D7_0419400	conserved <i>Plasmodium</i> protein	PF3D7_1233700	homocysteine S-methyltransferase
PF3D7_0502300	conserved <i>Plasmodium</i> protein	PF3D7_1241500	conserved <i>Plasmodium</i> protein
PF3D7_0507900	conserved <i>Plasmodium</i> protein	PF3D7_1243300	conserved <i>Plasmodium</i> protein
PF3D7_0509300	conserved <i>Plasmodium</i> protein	PF3D7_1245000	5-formyltetrahydrofolate cyclo-ligase
PF3D7_0510700	conserved <i>Plasmodium</i> protein	PF3D7_1247700	conserved <i>Plasmodium</i> protein
PF3D7_0511600	apical rhoptry neck protein	PF3D7_1248400	conserved <i>Plasmodium</i> protein
PF3D7_0512400	conserved <i>Plasmodium</i> protein	PF3D7_1249200	conserved <i>Plasmodium</i> protein
PF3D7_0526100	conserved <i>Plasmodium</i> membrane protein	PF3D7_1318300	conserved <i>Plasmodium</i> protein
PF3D7_0527600	conserved <i>Plasmodium</i> protein	PF3D7_1319000	conserved <i>Plasmodium</i> protein
PF3D7_0606600	conserved <i>Plasmodium</i> protein	PF3D7_1321100	conserved <i>Plasmodium</i> protein
PF3D7_0609300	conserved <i>Plasmodium</i> protein	PF3D7_1322700	conserved <i>Plasmodium</i> protein
PF3D7_0625200	conserved <i>Plasmodium</i> protein	PF3D7_1329200	conserved <i>Plasmodium</i> protein
PF3D7_0625900	conserved <i>Plasmodium</i> protein bifunctional methylenetetrahydrofolate dehydrogenase/cyclohydrolase,	PF3D7_1346300	DNA/RNA-binding protein Alba 2
PF3D7_0630700	conserved <i>Plasmodium</i> protein	PF3D7_1346700	6-cysteine protein
PF3D7_0710700	conserved <i>Plasmodium</i> protein	PF3D7_1351800	conserved <i>Plasmodium</i> protein
PF3D7_0717800	conserved <i>Plasmodium</i> protein	PF3D7_1354100	conserved <i>Plasmodium</i> protein
PF3D7_0727900	conserved <i>Plasmodium</i> protein	PF3D7_1354700	conserved <i>Plasmodium</i> protein
PF3D7_0806600	kinesin-like protein, putative	PF3D7_1359500	conserved <i>Plasmodium</i> protein
PF3D7_0814500	conserved protein	PF3D7_1364400	conserved <i>Plasmodium</i> protein
PF3D7_0814700	conserved <i>Plasmodium</i> protein	PF3D7_1366700	conserved <i>Plasmodium</i> protein
PF3D7_0816900	adenylate kinase 2	PF3D7_1337800	calcium-dependent protein kinase 5
PF3D7_0831600	cytoadherence linked asexual protein 8	PF3D7_1404200	conserved <i>Plasmodium</i> protein

Pfa ID	Protein	Pfa ID	Protein
PF3D7_0914100	conserved <i>Plasmodium</i> protein	PF3D7_1407400	conserved <i>Plasmodium</i> protein
PF3D7_0915600	conserved <i>Plasmodium</i> protein	PF3D7_1407600	conserved <i>Plasmodium</i> protein
PF3D7_0920300	conserved <i>Plasmodium</i> protein	PF3D7_1413700	conserved <i>Plasmodium</i> protein
PF3D7_0925000	conserved <i>Plasmodium</i> protein	PF3D7_1422300	DnaJ protein, putative
PF3D7_0925100	conserved <i>Plasmodium</i> protein	PF3D7_1430800	PhIL1 interacting protein PIP3
PF3D7_0926500	tetratricopeptide repeat protein	PF3D7_1431100	PhIL1 interacting protein PIP2
PF3D7_0933700	conserved <i>Plasmodium</i> protein	PF3D7_1436400	conserved <i>Plasmodium</i> protein
PF3D7_1002300	conserved <i>Plasmodium</i> protein	PF3D7_1447100	conserved <i>Plasmodium</i> protein conserved oligomeric Golgi complex subunit 6
PF3D7_1005400	conserved <i>Plasmodium</i> protein	PF3D7_1465100	conserved protein
PF3D7_1016100	conserved <i>Plasmodium</i> protein	PF3D7_1465400	conserved protein
PF3D7_1023700	conserved <i>Plasmodium</i> protein	PF3D7_1470300	conserved <i>Plasmodium</i> protein
PF3D7_1026000	conserved <i>Plasmodium</i> protein	PF3D7_1473000	conserved <i>Plasmodium</i> protein

Pfa ID, *P. falciparum* PlasmoDB ID

7.3 Ergänzende Abbildungen

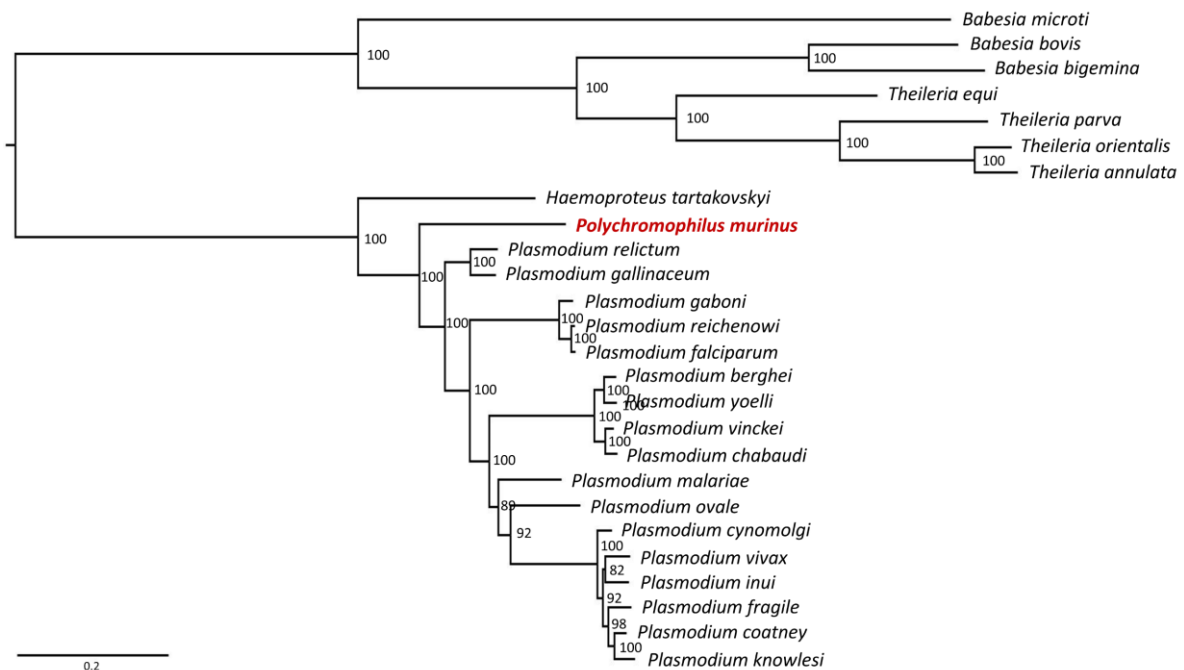


Abb. 7.1: ML-Phylogenie der Haemosporida basierend auf genomischen Sequenzen.

Dargestellt ist der Konsensusbaum nach 1000 Replikaten, Bootstrap Unterstützungen finden sich an den Verzweigungen. Der Baum wurde mit den Sequenzen der Piroplasmiden (*Theileria* und *Babesia*) gewurzelt.

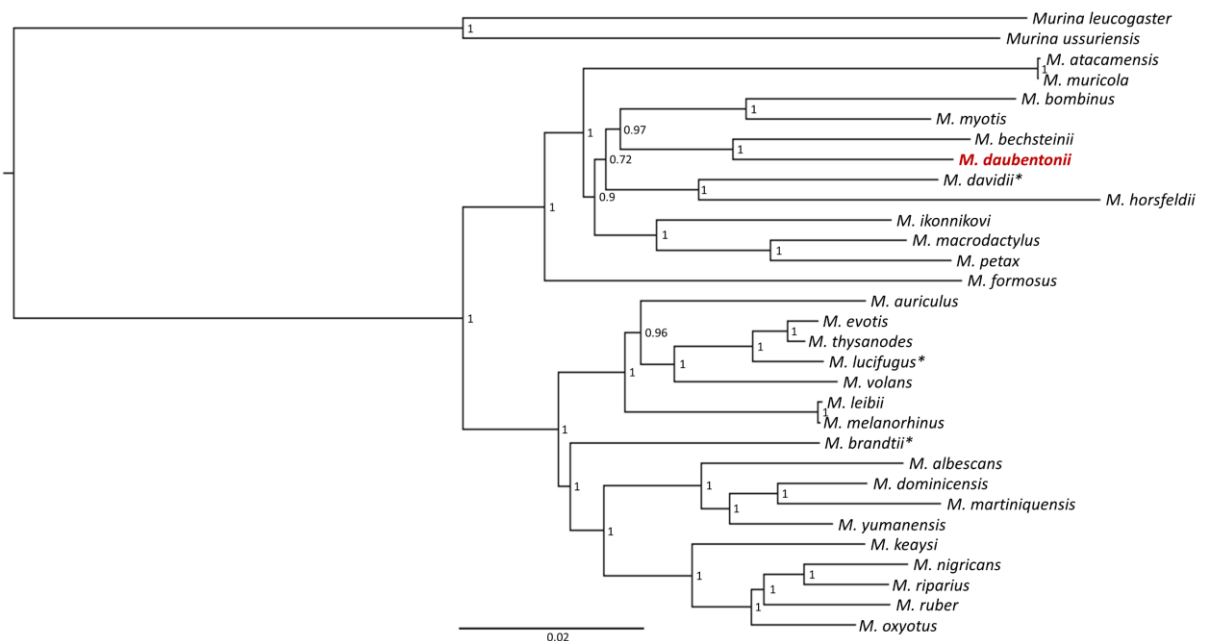


Abb. 7.2: Phylogenie der Gattung *Myotis* basierend auf BI-Analysen.

A priori Wahrscheinlichkeiten sind jeweils an den Verzweigungen dargestellt. Der Baum wurde mit den Sequenzen von zwei Fledermäusen der Gattung *Murina* gewurzelt. Sternchen: Fledermäuse, deren Genom bislang sequenziert wurde; rot hervorgehoben: *M. daubentonii*.

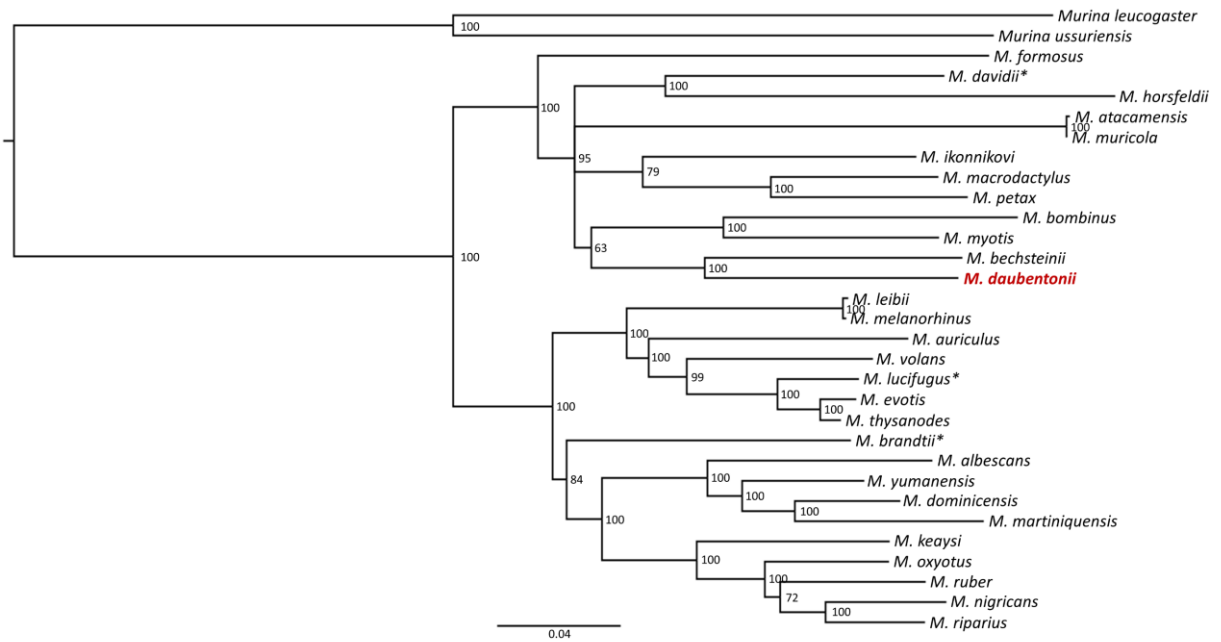


Abb. 7.3: Phylogenie der Gattung *Myotis* basierend auf ML-Analysen.

Dargestellt ist der Konsensusbaum nach 1000 Replikaten, Bootstrap Unterstützungen finden sich an den Verzweigungen. Der Baum wurde mit den Sequenzen von zwei Fledermäusen der Gattung *Murina* gewurzelt. Sternchen: Fledermäuse, deren Genom bislang sequenziert wurde; rot hervorgehoben: *M. daubentonii*.

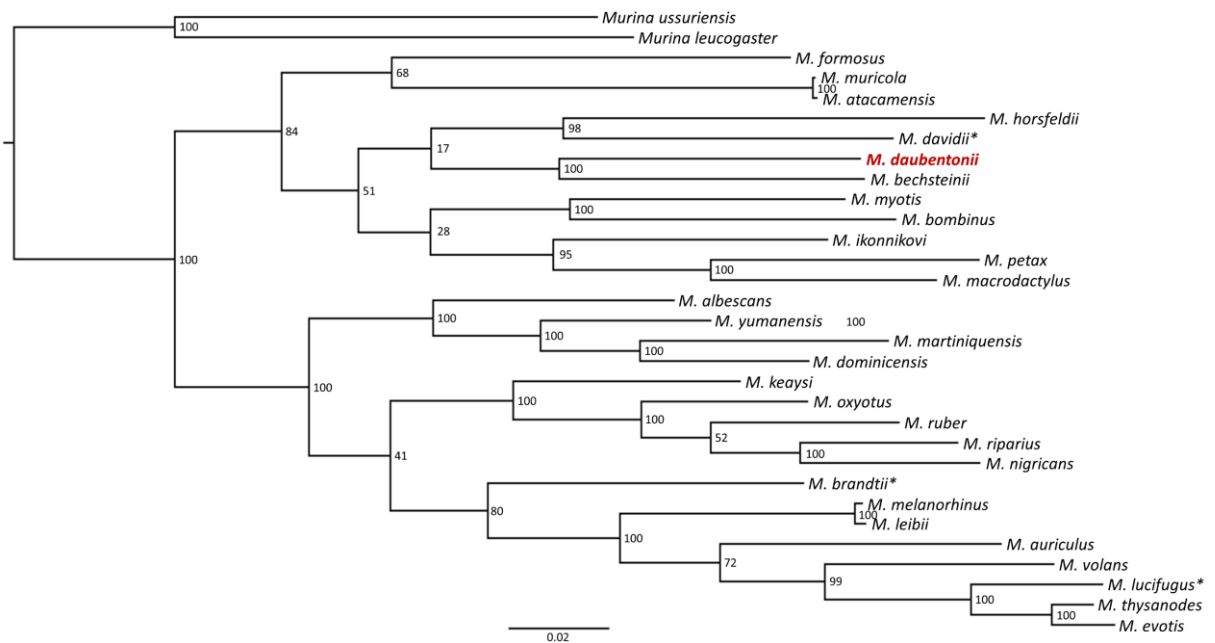


Abb. 7.4: Phylogenie der Gattung *Myotis* basierend auf MP-Analysen.

Dargestellt ist der Konsensusbaum nach 1000 Replikaten, Bootstrap Unterstützungen finden sich an den Verzweigungen. Der Baum wurde mit den Sequenzen von zwei Fledermäusen der Gattung *Murina* gewurzelt. Sternchen: Fledermäuse, deren Genom bislang sequenziert wurde; rot hervorgehoben: *M. daubentonii*.

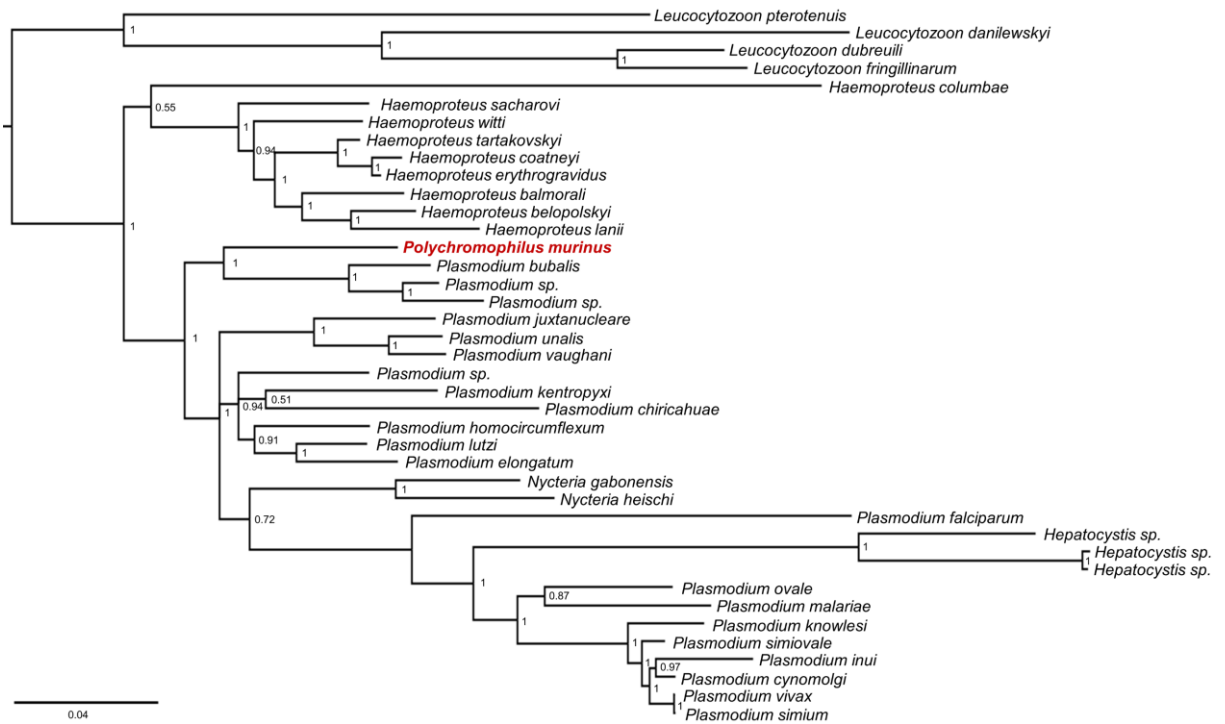


Abb. 7.5: Phylogenie der Haemosporida basierend auf BI-Analysen der mtDNA.

A priori Wahrscheinlichkeiten sind jeweils an den Verzweigungen dargestellt. Der Baum wurde mit *Leucocytozoon* gewurzelt.

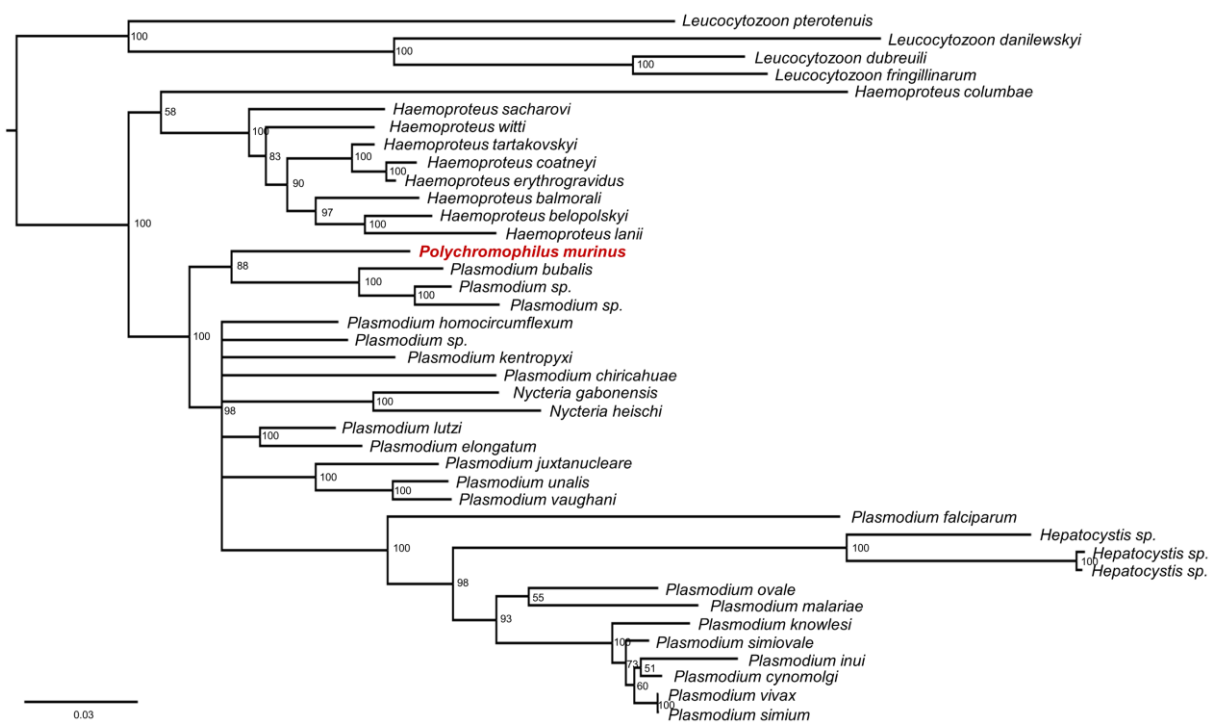


Abb. 7.6: Phylogenie der Haemosporida basierend auf ML-Analysen der mtDNA.

Dargestellt ist der Konsensusbaum nach 1000 Replikaten, Bootstrap Unterstützungen finden sich an den Verzweigungen. Der Baum wurde mit *Leucocytozoon* gewurzelt.

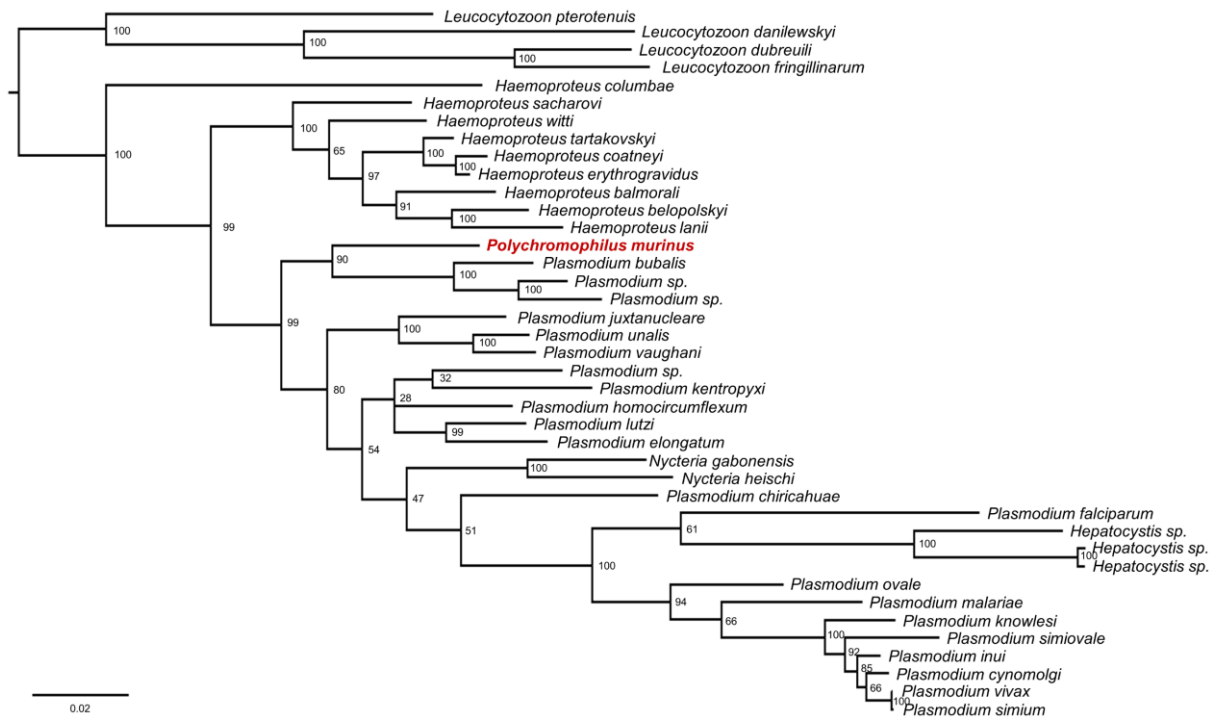


Abb. 7.7: Phylogenie der Haemosporida basierend auf MP-Analysen der mtDNA.

Dargestellt ist der Konsensusbaum nach 1000 Replikaten, Bootstrap Unterstützungen finden sich an den Verzweigungen. Der Baum wurde mit *Leucocytozoon* gewurzelt.

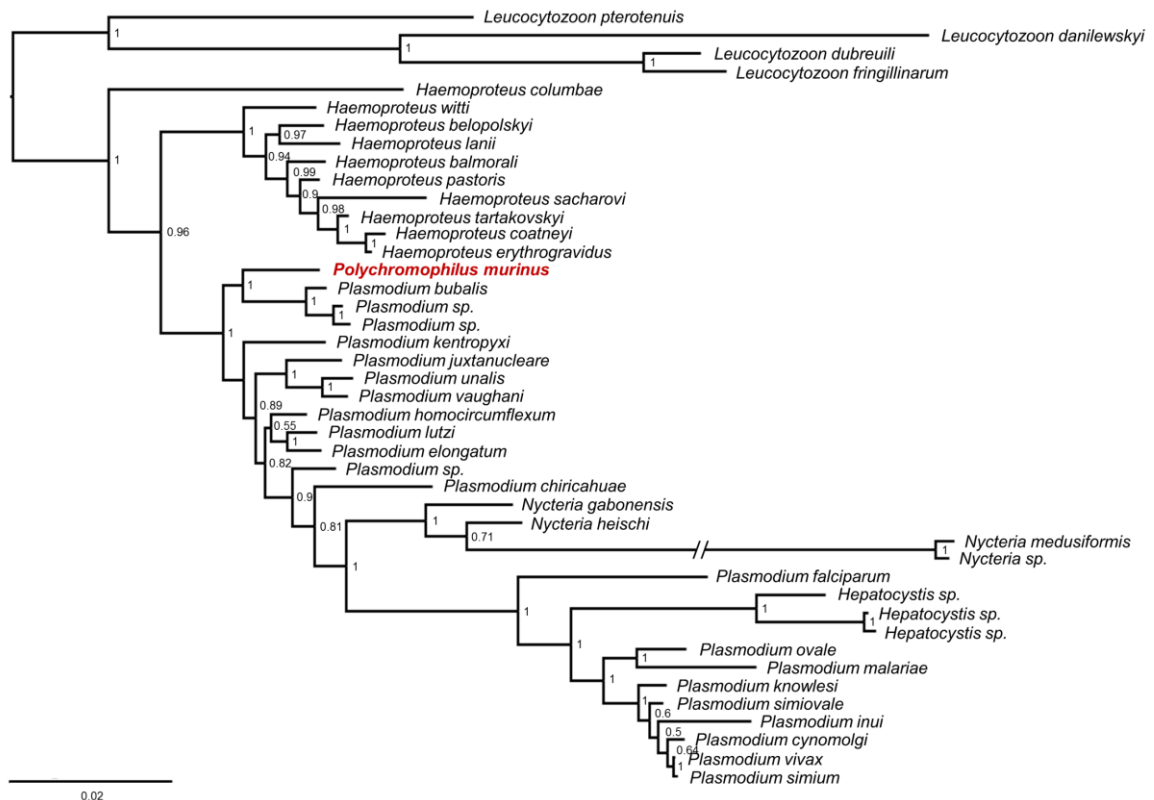


Abb. 7.8: BI-Analyse mitochondrialer Proteine der Haemosporida.

A priori Wahrscheinlichkeiten sind jeweils an den Verzweigungen dargestellt. Der Baum wurde mit *Leucocytozoon* gewurzelt.

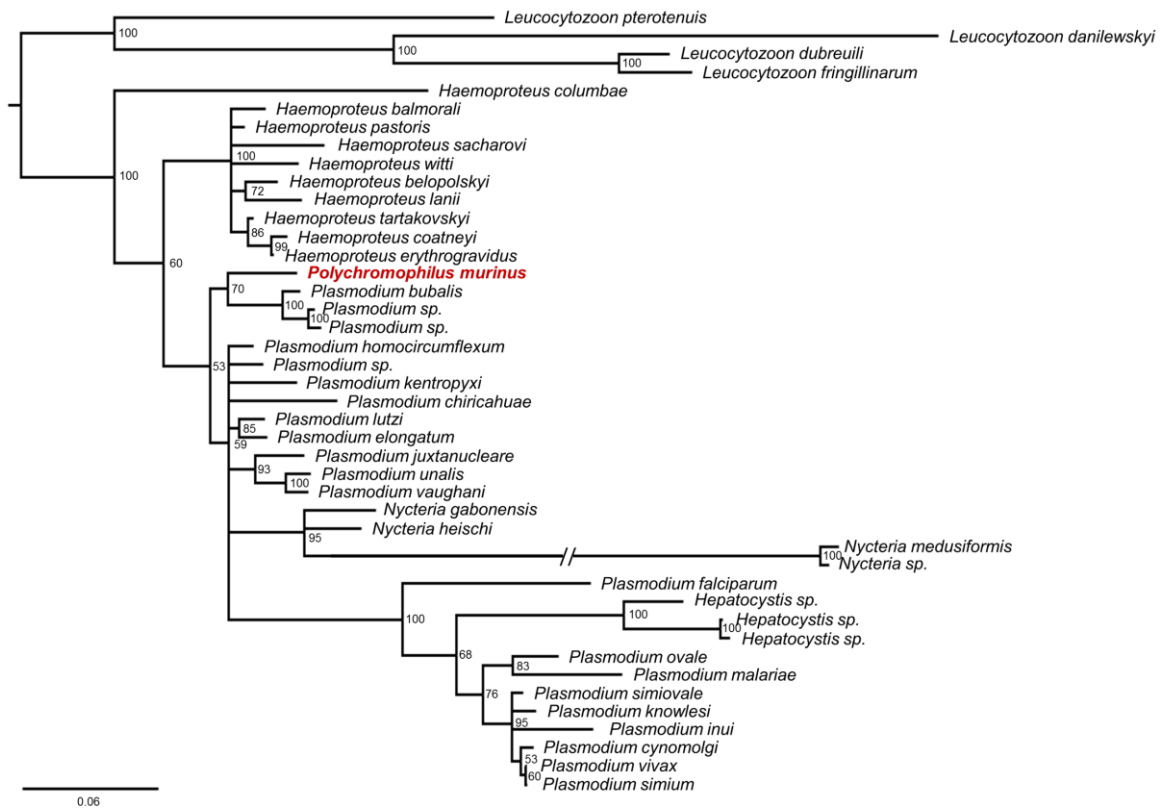
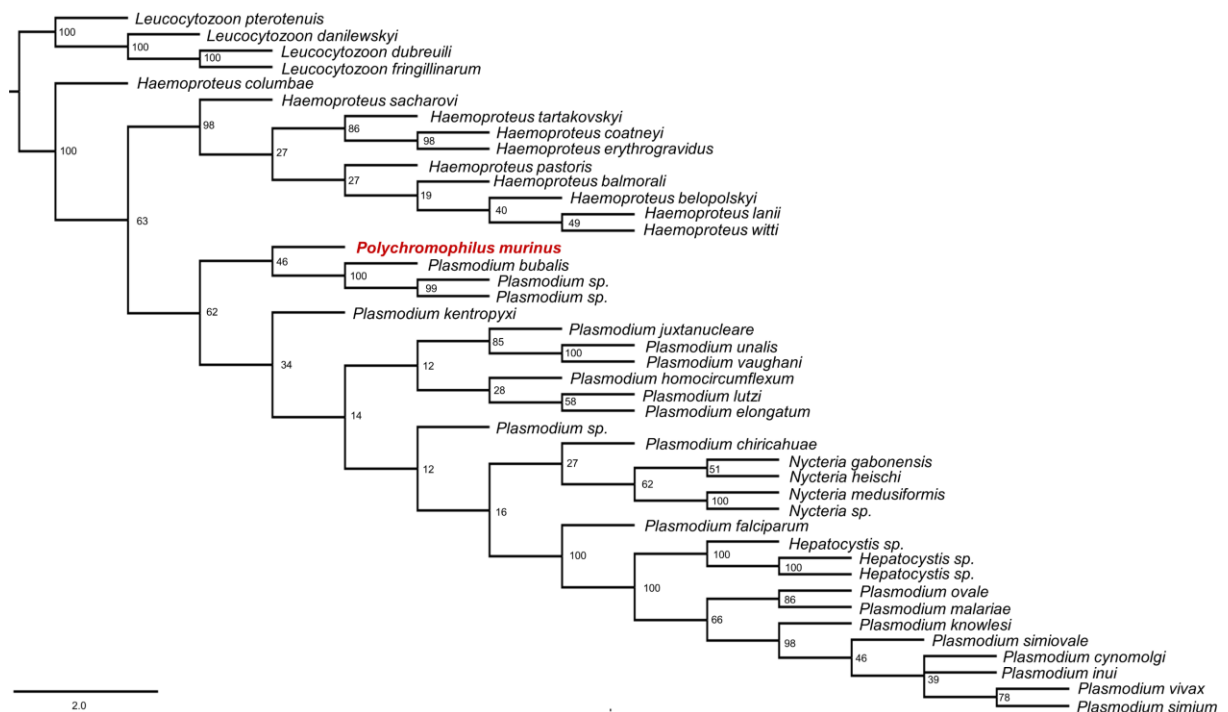


Abb. 7.9: MP-Analyse mitochondrialer Proteine der Haemosporida.

Dargestellt ist der Konsensusbaum nach 1000 Replikaten, Bootstrap Unterstützungen finden sich an den Verzweigungen. Der Baum wurde mit *Leucocytozoon* gewurzelt.

Abb. 7.11: Phylogenetische Stellung von *P. murinus* basierend auf Daten der mtDNA.

Dargestellt ist der Konsensusbaum nach 1000 Replikaten, Bootstrap Unterstützungen finden sich an den Verzweigungen. Der Baum wurde mit *Leucocytozoon* gewurzelt.



Abb. 7.12: Vergleich der repetitiven Region der Apikoplasten von *P. murinus* und *P. chabaudi*.

Abkürzungen: lsu, large subunit ribosomal ribonucleic acid; rps4, ribosomal protein S4; sufB, sulphur assimilation protein B, trnT, tRNA T.

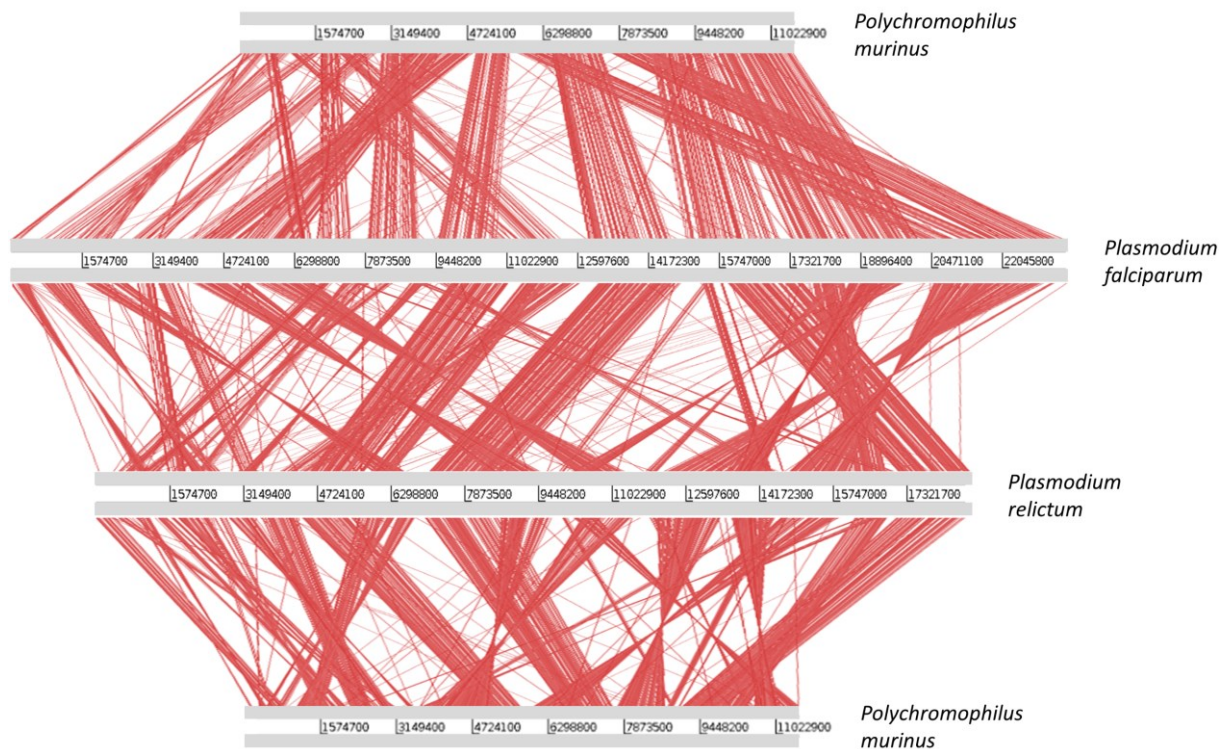


Abb. 7.13: Konservierte Syntanie-Blöcke von *P. murinus* im Vergleich.

Dargestellt sind die pseudochromosom-artigen Sequenzen von *P. murinus* und die Chromosomen von *P. falciparum* und *P. relictum* jeweils als Konkater, beginnend von der kleinsten (links) bis hin zur größten Sequenz (rechts). Abb. wurde mit dem Artemis Comparison Tool erzeugt.

Eidesstattliche Versicherung

Hiermit erkläre ich an Eides statt, dass die vorliegende Dissertationsschrift von mir selbstständig verfasst wurde und ich keine anderen als die angegebenen Hilfsmittel –insbesondere keine im Quellenverzeichnis nicht benannten Quellen – benutzt habe und die Arbeit von mir vorher nicht einem anderen Prüfungsverfahren eingereicht wurde. Die eingereichte schriftliche Fassung entspricht der auf dem Speichermedium.

Hamburg, März 2020

.....

Elsa Aul

Danke

Allen, die mich im Laufe der Zeit, in der diese Arbeit entstanden ist, begleitet haben und ohne deren Hilfe die Anfertigung dieser Dissertation niemals möglich gewesen wäre, möchte ich an dieser Stelle Danke sagen.

Insbesondere möchte ich mich bei Prof. Dr. Thorsten Burmester bedanken — für die Heranführung an ein wahnsinnig spannendes Thema, das mir nicht erst seit Beginn der Doktorarbeit ans Herz gewachsen ist. Danke für die Aufnahme in die Arbeitsgruppe, die Anregungen während meiner Arbeit und auch den gewährten wissenschaftlichen Gestaltungsfreiraum. Ich wünsche dir weiterhin alles Gute!

Bei Frau Prof. Dr. Iris Bruchhaus möchte ich mich ganz herzlich für die Übernahme des Erstgutachtens und die Unterstützung in der Endphase meiner Promotion bedanken.

Vielen, vielen Dank auch an Prof. Dr. Susanne Dobler für die spontane Zusage, als Erstgutachter für diese Arbeit zu fungieren.

Ein großes Dankeschön auch an Dr. Janus Borner und Dr. Andrej Fabrizius. Ohne Janus hätte ich wahrscheinlich noch wesentlich länger für diese Arbeit gebraucht. Ich bin dir wahnsinnig dankbar für jedes Skript und jede Erläuterung, die im Endeffekt so viel meiner Zeit gespart haben. Danke, Andrej, für die Bereitschaft meine Arbeit zu lesen und zu korrigieren; mir ist bewusst wie viel Arbeit das für dich bedeutet hat. Danke aber auch für die stets geöffnete Tür, die Diskussionsbereitschaft und die wegweisenden Gespräche — nicht nur fachlicher Natur.

Darüber hinaus bedanke ich mich bei Florian Gloza-Rausch, ohne den ich an kein Probenmaterial gekommen wäre. Danke, dass du dir die Zeit für unserer Fledermausfangaktion genommen hast.

Ganz herzlich möchte ich mich auch bei allen Mitgliedern der Arbeitsgruppe Burmester für die Hilfsbereitschaft, Unterstützung und vor allem für angenehme Arbeitsatmosphäre bedanken. Danke an Dr. Cornelia Geßner, Maria Machola, Michelle Wisniewsky, Katja Reimann und Marilou Kloock für jede auch noch so kleine Hilfe im Labor. Dir, Maria, danke ich außerdem für jede einzelne PCR, und jedes einzelne Gelbild in dem *Haemoproteus*-Projekt, das in dieser Arbeit leider keinen Platz gefunden hat.

Danke an meine Mitstreiter Alena Krüger, Julia Orth, Kathrin Helfenrath, Paressa Papadopoulou-Wörner und Sven Flachsbarth, die — obwohl selbst, oder gerade weil promovierend — immer ein offenes Ohr für Diskussionen hatten. Der Austausch von Ideen und Gedanken, die motivierenden Gespräche aber auch jede Feier haben wesentlich zum Gelingen dieser Arbeit beigetragen.

Meiner Familie danke ich ganz besonders dafür, dass sie mir diese Ausbildung ermöglicht haben und stets mit tatkräftiger Hilfe an meiner Seite stehen. Insbesondere bin ich meinem Mann Vitali dankbar. Danke, dass du auch noch an diese Arbeit — an mich — geglaubt hast, als ich selbst gezweifelt habe. Danke dafür, dass du mir den Raum gegeben hast, den ich gebraucht habe, dass du mich an allen Ecken und Enden unterstützt hast und mit mir diese — teilweise sehr nervenraubende — Zeit durchgestanden hast.