

Aus dem Institut für Medizinische Mikrobiologie und Hygiene,
Medizinische Fakultät der Universität des Saarlandes, Homburg/Saar

Direktor: Prof. Dr. Dr. Sören Becker

Anwendung der MALDI-TOF Massenspektrometrie zur Identifikation der Trematoden
Fasciola und Schistosoma

Dissertation zur Erlangung des Grades eines Doktors der Humanmedizin

der Medizinischen Fakultät

der UNIVERSITÄT DES SAARLANDES

2023

vorgelegt von:

Jurena Christiane Ebersbach

geb. am: 17.01.1997 in Innsbruck, Österreich

Tag der Promotion: 25.04.2024

Dekan: Prof. Dr. M.D. Menger

1. Berichterstatter: Prof. Dr. Dr. S. Becker

2. Berichterstatter: Prof. Dr. Markus Meyer

Für Enno

Inhaltsverzeichnis

I. ABKÜRZUNGSVERZEICHNIS	V
II. ABBILDUNGSVERZEICHNIS	VII
III. TABELLENVERZEICHNIS	IX
1. ZUSAMMENFASSUNG	1
2. ABSTRACT	3
3. EINLEITUNG	5
3.1. VERNACHLÄSSIGTE TROPENKRANKHEITEN	5
3.2. FASCIOLA.....	7
3.2.1. ÜBERTRAGUNG UND VERBREITUNG	7
3.2.2. MORPHOLOGIE	7
3.2.3. LEBENSZYKLUS	8
3.2.4. ROLLE IN DER VETERINÄRMEDIZIN	10
3.2.5. KLINISCHE CHARAKTERISTIKA DER HUMANEN FASZIOLOSE.....	10
3.2.6. DIAGNOSTIK.....	11
3.2.7. THERAPIE DER FASZIOLOSE	13

3.3. SCHISTOSOMA.....	15
3.3.1. ÜBERTRAGUNG UND VERBREITUNG	15
3.3.2. MORPHOLOGIE	15
3.3.3. LEBENSZYKLUS	16
3.3.4. ROLLE IN DER VETERINÄRMEDIZIN	17
3.3.5. KLINISCHE CHARAKTERISTIKA DER HUMANEN SCHISTOSOMIASIS	18
3.3.6. DIAGNOSTIK.....	19
3.3.7 THERAPIE DER SCHISTOSOMIASIS.....	21
3.4. MALDI-TOF MS IN DER MIKROBIOLOGISCH-INFEKTIONS-DIAGNOSTISCHEN ANWENDUNG	22
4. ZIEL DER ARBEIT / FRAGESTELLUNG	25
5. MATERIAL UND METHODEN	26
5.1. VERSUCHSAUFBAU	26
5.2. PROBENVORBEREITUNG.....	28
5.2.1. PROBENHERKUNFT	28
5.2.2. LAGERUNGSMEDIEN	29
5.2.3. PROTEINEXTRAKTION.....	30
5.3. MASSENSPEKTROMETRISCHE MESSUNG MITTELS MALDI-TOF MS	33
5.3.1. MALDI TARGET PLATE VORBEREITUNG UND MESSUNG	33
5.3.2. MALDI-TOF MS PARAMETER	34

5.4. DATENANALYSE MIT MALDI-TOF BIOTYPER	35
5.4.1. SPEKTRENAUFBEREITUNG UND ERSTELLEN EINER HAUSINTERNEN DATENBANK	35
5.4.2. AUSWERTUNG DER MESSERGESBNISSE.....	37
5.4.3. VALIDIERUNG DER DATENBANK	37
5.5. KLASSIFIZIERUNGS- UND VERGLEICHSANALYSE MITTELS „CLOVER“- SOFTWARE	38
5.5.1. PARAMETER ZUR VORVERARBEITUNG.....	38
5.5.2. KLASSIFIZIERUNG UND ALGORITHMEN DES MASCHINELLEN LERNENS	38
6. GERÄTE, REAGENZIEN, SOFTWARE	44
7. ERGEBNISSE	45
7.1. FASCIOLA GIGANTICA UND FASCIOLA HEPATICA	45
7.1.1. VALIDIERUNG DER MALDI-TOF MS BASIERTEN DIAGNOSTIK VON F. HEPATICA UND F. GIGANTICA EIERN.....	45
7.1.2. INTERSPEZIES-KLASSIFIKATION: SPEZIESSPEZIFISCHE DIFFERENZIERUNG DURCH MALDI-TOF MS	50
7.1.3. INTRASPEZIES-KLASSIFIKATION: EINFLUSS DES ENTWICKLUNGSSTADIUMS AUF DIE MALDI-TOF BASIERTE DIAGNOSTIK	52
7.1.4. INTRASPEZIES-KLASSIFIKATION: EINFLUSS DES VERWENDETEN KÖRPERANTEILS.....	56

7.2. SCHISTOSOMA MANSONI UND SCHISTOSOMA JAPONICUM.....	57
7.2.1. VALIDIERUNG DER MALDI-TOF MS BASIERTEN DIAGNOSTIK VON S. JAPONICUM UND S. MANSONI.....	57
7.2.2. INTERSPEZIES-KLASSIFIKATION: SPEZIESSPEZIFISCHE DIFFERENZIERUNG DURCH MALDI-TOF MS	60
7.2.3. INTRASPEZIES-KLASSIFIKATION: EINFLUSS DES LAGERUNGSMEDIUMS AUF MALDI-TOF MS BASIERTE DIAGNOSTIK.....	62
8. DISKUSSION	65
8.1. KRITISCHE BETRACHTUNG DER EIGENEN UNTERSUCHUNG	65
8.2. PRAKTISCHE ANWENDUNG DER ERGEBNISSE	69
8.3. VERGLEICH DER EIGENEN ERGEBNISSE MIT DENEN AUS DER LITERATUR.....	71
8.4. SCHLUSSFOLGERUNG	76
9. LITERATURVERZEICHNIS	77
10. PUBLIKATION	90
11. DANKSAGUNG	91
12. LEBENSLAUF	92
13. ANHANG	93
14. ERKLÄRUNG AN EIDES STATT.....	988

I. Abkürzungsverzeichnis

Abb.	Abbildung
AUPR-Kurve	Precision-Recall-Kurve
AUROC-Kurve	Receiver-Operating-Characteristic-Kurve
BTS	Bacterial Test Standard
bzw.	Beziehungsweise
°C	Grad Celsius
Da	Dalton
EPG	Eiern pro Gramm Stuhl
FDA	Food and Drug Administration
<i>F. gigantica</i>	<i>Fasciola gigantica</i>
<i>F. hepatica</i>	<i>Fasciola hepatica</i>
PCA	Hauptkomponentenanalyse
KNN	K-nearest neighbors
LA	Lernalgorithmen
LOO	Leave One Out
LSV	log-scores value
m/z	Masse – Ladungsverhältnis
MALDI TOF	Matrix-assisted laser desorption/ionization time-of-flight
min	Minuten
ML	Maschinelles Lernen
µl	Microliter
mg/kgKG	Milligramm pro Kilogramm Körpergewicht
mm	Millimeter
MS	Mass spectrometry
MSP	Main Spectrum Profil (Referenzspektrum)

NaCl	Natriumchloridlösung
nm	Nanometer
ns	Nanosekunden
NPV	negativer Vorhersagewert
TIC	Gesamtionenstrom
TOF	Time of flight
PBS	Phosphatgepufferte Salzlösung (englisch „phosphate-buffered saline“)
PLS-DA	Partial Least Squares-Discriminant Analysis
PCR	Polymerase chain reaction
ppm	Parts per million
PPV	positiver Vorhersagewert
RF	Random Forest
ROC	Receiver-Operating-Characteristic-Kurve
rpm	Rounds per minute
sek	Sekunden
<i>S. mansoni</i>	<i>Schistosoma mansoni</i>
<i>S. japonicum</i>	<i>Schistosoma japonicum</i>
<i>S. haematobium</i>	<i>Schistosoma haematobium</i>
SVM	Lineare Supervektormaschine
u.a.	unter anderem
WHO	World Health Organization

II. *Abbildungsverzeichnis*

Abbildung 1: Neglected tropical diseases	6
Abbildung 2: verschiedene Stadien und Spezies der <i>Fasciola</i>	8
Abbildung 3: Lebenszyklus <i>F. gigantica</i> und <i>F. hepatica</i>	9
Abbildung 4: Lebenszyklus <i>S. mansoni</i> , <i>S. japonicum</i> und <i>S. haematobium</i>	17
Abbildung 5: MALDI-TOF MS Funktionsweise	23
Abbildung 6: Workflow zur Arbeit mit <i>Fasciola</i> und <i>Schistosoma</i>	27
Abbildung 7: Vergleich: Anzahl der Proteinidentifizierungen vier Konservierungstechniken	29
Abbildung 8: Protokoll Proteinextraktion	30
Abbildung 9: <i>F. gigantica</i> . Anatomie	31
Abbildung 10: <i>F. gigantica</i> Eier	32
Abbildung 11: <i>MSP 96 Polished Steel MALDI Target Plate:</i>	34
Abbildung 12: Spektrenaufbereitung	35
Abbildung 13: Probe von <i>F. gigantica</i> -Eiern - Proteinspektrum der Probe.....	36
Abbildung 14: Unterschied überwachtes vs. unüberwachtes Lernen	40
Abbildung 15: Proteinspektren von <i>F. gigantica</i> Eiern	46
Abbildung 16: Proteinspektren von <i>F. hepatica</i> -Eiern und adulten <i>F. hepatica</i> -Proben.....	47
Abbildung 17: Vergleich der Proteinspektren von <i>F. gigantica</i> und <i>F. hepatica</i>	47
Abbildung 18: Unüberwachter LA (PCA): Darstellung aller <i>Fasciola</i> -Proben.....	51
Abbildung 19: Unüberwachter LA (PCA): Darstellung adulte <i>F. hepatica</i> (blau) und adulte <i>F. gigantica</i> -Proben	51
Abbildung 20: Überwachte LA: Darstellung <i>F. gigantica</i> -Proben und <i>F. hepatica</i> -Proben ..	52
Abbildung 21: Unüberwachter LA (PCA): Darstellung adulte <i>F. gigantica</i> -Proben und Eier	54
Abbildung 22: Überwachte LA: Darstellung adulte <i>F. gigantica</i> -Proben und <i>F. gigantica</i> -Eier	54
Abbildung 23: Unüberwachter LA: Darstellung adulte <i>F. hepatica</i> -Proben und <i>F. hepatica</i> -Eier	55

Abbildung 24: überwachter LA: Darstellung adulte *F.hepatica*-Proben und *F.hepatica*-Eier 55

Abbildung 25: Proteinspektren von *S. mansoni*-Proben und *S. japonicum*-Proben 57

Abbildung 26 unüberwachter LA (PCA): Darstellung *S. mansoni* und *S. japonicum* 61

Abbildung 27: überwachter LA: Darstellung *S. mansoni* und *S. japonicum* 61

Abbildung 28: unüberwachter LA (PCA): Darstellung *S. mansoni* in 70% (v/v) Ethanol und RNAlater® 62

Abbildung 29: überwachter LA: Darstellung *S. mansoni* in 70% (v/v) Ethanol und RNAlater® 63

Abbildung 30: unüberwachter LA (PCA): Darstellung *S. japonicum* in 70% (v/v) Ethanol und RNAlater® 63

Abbildung 31: überwachter LA: Darstellung von *S. japonicum* in 70% (v/v) Ethanol und RNAlater® 64

Anhang:

Abbildung 32: AUROC und ARCP – *Fasciola* spp..... 97

Abbildung 33: AUROC und ARCP – *Schistosoma* spp.: A - PLS-DA, B – SVM, C – KNN, D - RF..... 98

III. Tabellenverzeichnis

Tabelle 1: <i>Fasciola</i> spp. vs. BDAL-Datenbank von Bruker Daltronic	45
Tabelle 2: Interne und Externe Validierung: Identifizierung von <i>Fasciola</i> -Proben, in verschiedenen Entwicklungsstadien (adulte Proben, Eier)	49
Tabelle 3: <i>Schistosoma</i> spp. vs, BDAL- Datenbank von Bruker Daltronic.....	57
Tabelle 4: Interne und Externe Validierung: Identifizierung von <i>Schistosoma</i> -Proben in verschiedenen Lagerungsmedien (70% (v/v) Ethanol und RNAlater®).....	59
Anhang:	
Tabelle 5: 10-fache Kreuzvalidierung <i>Fasciola</i> spp.	94
Tabelle 6: 10-fache Kreuzvalidierung <i>Schistosoma</i> spp.	96

1. Zusammenfassung

Einleitung: Die Schistosomiasis und die Fasziole sind die humanmedizinisch bedeutsamsten Trematoden-Infektionen (Saugwurm-Infektionen) und können sowohl zu akuten Erkrankungen als auch zu chronischen Komplikationen führen. Beide Erkrankungen gehören zu den sogenannten vernachlässigten Tropenkrankheiten (neglected tropical diseases, NTD), deren höchste Krankheitslast in armen Ländern des globalen Südens liegt. Die heutige Routinediagnostik der humanen Schistosomiasis und Fasziole wird vorrangig mittels spezieller mikroskopischer Untersuchungen durchgeführt. Aufgrund der Seltenheit dieser Infektionen in Europa sinkt die Expertise des Laborpersonals in der korrekten mikroskopischen Speziesidentifizierung. Darüber hinaus ist aufgrund der ähnlichen Morphologie der Wurmeier eine speziesspezifische Diagnose nicht immer möglich. Andere diagnostische Methoden wie molekulare PCR-Untersuchungen oder serologische Tests sind zeitaufwendig, teuer und in vielen Ländern des globalen Südens sowie auch in Europa außerhalb von Referenzlaboren nicht flächendeckend verfügbar. In der schnellen und genauen Diagnostik von Bakterien und Pilzen ist die sogenannte matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) seit Jahren eine Standard-Identifikationsmethode. Die Vorteile dieser massenspektrometrischen Untersuchungstechnik liegen besonders darin, dass Ergebnisse innerhalb von Minuten, hochspezifisch und zu einem geringen Preis pro untersuchtem Analyten vorliegen. Seit wenigen Jahren wird auch die Anwendung von MALDI-TOF MS für die Diagnostik verschiedener Parasitosen erforscht, und zeigte vielversprechende Ergebnisse. Ziel dieser Arbeit war daher, die erstmalige Anwendung der MALDI-TOF MS zur Detektion von *Schistosoma* spp. (*Schistosoma mansoni* und *Schistosoma japonicum*) sowie *Fasciola* spp. (*Fasciola gigantica* und *Fasciola hepatica*) zu untersuchen und verschiedene Einflussfaktoren auf die Identifizierungsraten zu evaluieren.

Material und Methoden: Die MALDI-TOF-Analytik basiert auf der „Wiedererkennung“ erregerspezifischer Proteinspektrenprofile, wobei im Massenspektrometer ein Abgleich der gemessenen Spektren mit pathogenspezifischen „main spectra profiles“ (MSPs) erfolgt. Die hierfür erforderliche Datenbank mit MSPs muss nach der Erstellung internen und externen Validationsschritten unterzogen werden. Für die vorliegende Arbeit wurden als Referenzspektren 13 molekulargenetisch identifizierte Proben von *Fasciola* spp. aus Nigeria und der Schweiz und 4

Schistosoma spp. Proben aus Japan verwendet. Zur externen Validierung dieser generierten MSP-Datenbank wurden 111 unbekannte *Fasciola* spp. Proben und 58 unbekannte *Schistosoma* spp. Isolate untersucht worden. Die Identifikation wurde auf Spezies- bzw. Gattungsebene nach ihren log-scores value (LSV) angegeben, wobei ein LSV von $\geq 1,7$ eine korrekte Identifikation der Gattung und ein LSV von $\geq 2,0$ eine sichere Identifikation der Spezies anzeigt. Untersucht wurden dabei verschiedene Einflussfaktoren auf die Zuverlässigkeit der Identifizierung. Zusätzlich wurden die Ergebnisse in Anwendung unüberwachter (Hauptkomponentenanalyse (PCA)) und überwachter (Support Vector Machine (SVM), k-nearest neighbor (KNN), random forest (RF) und partial least square-discriminant analysis (PLS-DA)) Lernalgorithmen (LA) des maschinellen Lernens (ML) klassifiziert, um eine Verbesserung der Identifikation durch bestimmte LA zu untersuchen. Als Gradmesser zur Einschätzung einer Modellleistung wurde der F1-Score eingesetzt, welcher Aufschluss über eine die Sensitivität und Präzision des LA gibt.

Ergebnisse: Von 100% (n=169 der Proben konnten Proteinspektren mittels MALDI-TOF MS generiert werden. In der externen Validierung war es möglich, mit der neu erstellten in-house-Database eine korrekte Speziesidentifizierung mit einem LSV $> 2,0$ für 85% (94/111) der *Fasciola* spp. Proben zu erzielen. Ähnliche Ergebnisse ergaben sich bei *Schistosoma* spp. mit einer speziesspezifischen Identifizierung von 81% (47/58) Proben. Es zeigte sich eine Abhängigkeit der Identifikationsraten vom Entwicklungsstadium: So konnten die Eier von *Fasciola* spp. nur in 68% (15/22) korrekt identifiziert werden, wenn lediglich von adulten Würmern gewonnene MSPs als Referenz genutzt wurden. Auch gab es Unterschiede zwischen verschiedenen Lagerungsmedien von *Schistosoma* spp.: Wurde RNAlater® eingesetzt, war eine korrekte Identifizierung von 96% (29/30) der Proben möglich, wohingegen nur 36% (10/28) nach Lagerung in Ethanol korrekt identifiziert wurden. Diese Beobachtungen ließen sich auch in verschiedenen LA bestätigen. Dabei stellte der LA SVM bei den meisten Proben den geeignetsten LA dar, jedoch konnten alle LA mit einem F1 Score von über 90% eine gute Modellleistung aufweisen.

Schlussfolgerung: Eine erfolgreiche Identifikation von *Fasciola*- und *Schistosoma*-Proben mit MALDI-TOF MS ist möglich, wobei das gewählte Lagerungsmedium und auch das Entwicklungsstadium des Wurms einen Einfluss auf die Identifikationsraten haben. Weitere Untersuchungen von Würmern aus unterschiedlichen Weltregionen sind erforderlich, um die Methodik erfolgreich zu validieren und ggf. vermehrt in der klinischen Routinediagnostik einzusetzen.

2. Abstract

Background: Schistosomiasis and fascioliasis are part of the neglected tropical diseases, defined by the WHO as the infections with the highest disease burden in poor countries of the global south, causing devastating social, health and economic consequences to more than one billion people. Today the routine diagnostic is performed by microscopic examinations. Due to the rarity of these diseases in Europe, the missing of microscopic expertise is waning outside endemic areas, and the difficulty to distinguish the egg morphology of different species, this technique is not sufficient. Other diagnostic methods such as molecular PCR investigations or serological tests are time-consuming, expensive and not widely available outside reference laboratories in many countries of the global South and also in Europe. This indicates that new methods of diagnosis are needed. For bacteria and fungi, the MALDI-TOF MS has become an important diagnostic tool, because it provides results within minutes while being highly specific and at a low cost per analyte. During the last years the application of MALDI-TOF MS for the diagnosis of different parasitosis has been investigated and showed promising results. The aim of this work was to investigate the first application of MALDI-TOF MS for the detection of *Schistosoma* spp. (*S.mansoni* and *S.japonicum*) and *Fasciola* spp. (*F.gigantica* and *F.hepatica*) and to evaluate different factors influencing the identification rates.

Methods: For this work, 13 molecularly identified samples of *Fasciola* spp. and 4 *Schistosoma* spp. samples were used as reference spectra. These reference spectra are used in MALDI analysis for the recognition of specific protein spectra profiles. For external validation, 111 unknown *Fasciola* spp. samples and 58 unknown *Schistosoma* spp. isolates have been analyzed. Identification was reported at the species or genus level according to their log-scores value (LSV) (LSV > 1.7 correct identification of genus, LSV > 2.0 correctly identified species). In addition, various influencing factors, such as storage media and developmental stages on the reliability of identification were investigated. Finally, the results were validated using unsupervised (PCA) and supervised (SVM, KNN, RF, PLS-DA) learning algorithms. The F1 score, which provides information about the sensitivity and precision of the LA, was used as a measure to assess model performance.

Results: Protein spectra could be generated from 100% (169/169) of the samples by MALDI-TOF MS. For external validation it was possible to assign 85% (94/111) of the *Fasciola* spp.

with correct species identification (LSV > 2.0)) using the newly created in-house-database. Similar results were obtained for *Schistosoma* spp. with a species-specific identification of 81% (47/58) samples. The strongest influence on the protein spectra had the developmental stage of *Fasciola* spp. samples. Differences in the spectra of the *Schistosoma* spp. samples in relation to the storage medium were also found. These results could be confirmed in the examination of the different LA. In this regard, the LA SVM represented the most accurate LA for most samples, however, all LA were able to show good model performance with an F1-Score above 90%.

Conclusion: MALDI-TOF MS can reliably distinguish between the species of *Schistosoma* spp. and *Fasciola* spp. samples. An influence of different developmental stages or storage media on the generated protein spectra could be observed. This is a further step towards the application of MALDI-TOF MS in clinical routine.

3. Einleitung

3.1. Vernachlässigte Tropenkrankheiten

Neglected tropical diseases (NTD), zu Deutsch „Vernachlässigte Tropenkrankheiten“, umfassen laut WHO-Liste eine Gruppe von rund 20 viralen, bakteriellen und parasitären Erkrankungen, von welchen etwa als 1,7 Milliarden Menschen weltweit betroffen und schätzungsweise weitere 2 Milliarden Menschen bedroht sind. Die Auswirkungen der NTDs lassen sich in sozialen, ökonomischen und gesundheitlichen Bereichen beobachten. Vernachlässigt wurden diese Erkrankungen lange sowohl vom Weltgesundheitssystem als auch von globalen Förderorganisationen sowie hinsichtlich der Forschung und Medikamentenentwicklung. Dies steht auch in starkem Gegensatz zu Erkrankungen wie Malaria, Tuberkulose und HIV/AIDS, welche in den letzten Jahrzehnten Gegenstand weitreichender Forschung und neuer Präventions- und Behandlungsansätze waren. Ein Grund für diesen Gegensatz mag die Verbreitung in entlegenen Gebieten und die oft von Armut betroffene Patient*innengruppe sein. Hauptsächlich treten diese Krankheiten in tropischen Gebieten mit unzureichender Infrastruktur und begrenzten Bildungsmöglichkeiten auf. Dazu zählen besonders vom Klimawandel stark betroffene Regionen, wodurch zusammengenommen wiederum ein Kreislauf der Stigmatisierung und sozialen Ausgrenzung dieser dort ansässigen Bevölkerungsgruppen aufrechterhalten wird. So ist beispielsweise die Förderung von Landwirtschaft und Fischerei wichtig zur Bekämpfung von Armut und Hunger, birgt aber die Gefahr der Infektion durch kontaminierte Bewässerungssysteme und Übertragung von Erkrankung von Tieren auf Menschen [30]. Es handelt sich bei den NTDs meist um vektorassoziierte Erkrankungen, oft durch Tiere wie Weidevieh oder in verschmutzten Gewässern ansässige Zwischenwirte. Verbesserte Aufklärung, Hygiene und Infrastruktur sind an dieser Stelle somit von großer Bedeutung in der Prävention. Mit dem Ziel der Eliminierung dieser Gruppe von Erkrankung wurde in den letzten Jahren auf politischer Ebene ein Umdenken angestoßen, finanzielle Mittel bereitgestellt und die Forschung vorangetrieben [104]. Dazu hat die WHO 2020 eine „Road Map“ unter dem Titel "Ending the neglect to attain the Sustainable Development Goals: a road map for neglected tropical diseases 2021-2030" vorgestellt und das Ziel der Reduzierung aller behandlungsbedürftigen NTDs um 90% bis 2030 festgeschrieben. Hierzu sollen vorrangig folgende Strategien im Zentrum stehen: Vorbeugende Chemotherapien gegen NTDs (= „Massenbehandlungen“ in Endemiegebieten), veterinärmedizinische und humanmedizinische Zusammenarbeit an der Schnittstelle Mensch-Tier, Kontrolle

von Vektoren und Zwischenwirte, Bereitstellung von sauberer Trinkwasserversorgung, Abwasserentsorgung und Hygiene. Eine verbesserte Diagnostik wird ebenfalls als wichtiger Baustein zur Eliminierung der NTDs beschrieben. Hierzu wird die Entwicklung innovativer diagnostischer Tests angestrebt, die auch in ressourcenarmen Gebieten zugänglich sein müssen, also kostengünstig, durch weitgehend ungelernetes Personal durchführbar, empfindlich und hochspezifisch sind, einen hohen Durchsatz ermöglichen, hitzestabil sind und wenig und/oder einfache Ausrüstung benötigen und deren Qualität durch Qualitätskontrollen gewährleistet ist [109].

Die in der vorliegenden Arbeit behandelten parasitären Erkrankungen, Schistosomiasis und Fasziole, gehören zu der 2017 von der WHO veröffentlichten Liste der NTDs. Die Schistosomiasis zählt dabei sogar zu den 5 wichtigsten NTDs, welche bis Ende 2030 laut WHO „Road Map“ eliminiert sein sollen [116]. Für Schistosomiasis und die Fasziole (letztere gehört zu den über Lebensmittel übertragenen Trematoden-Infektionen) hat die WHO konkrete Ziele zur Weiterentwicklung von Diagnostikmethoden aufgestellt (siehe Abbildung (Abb.) 1)

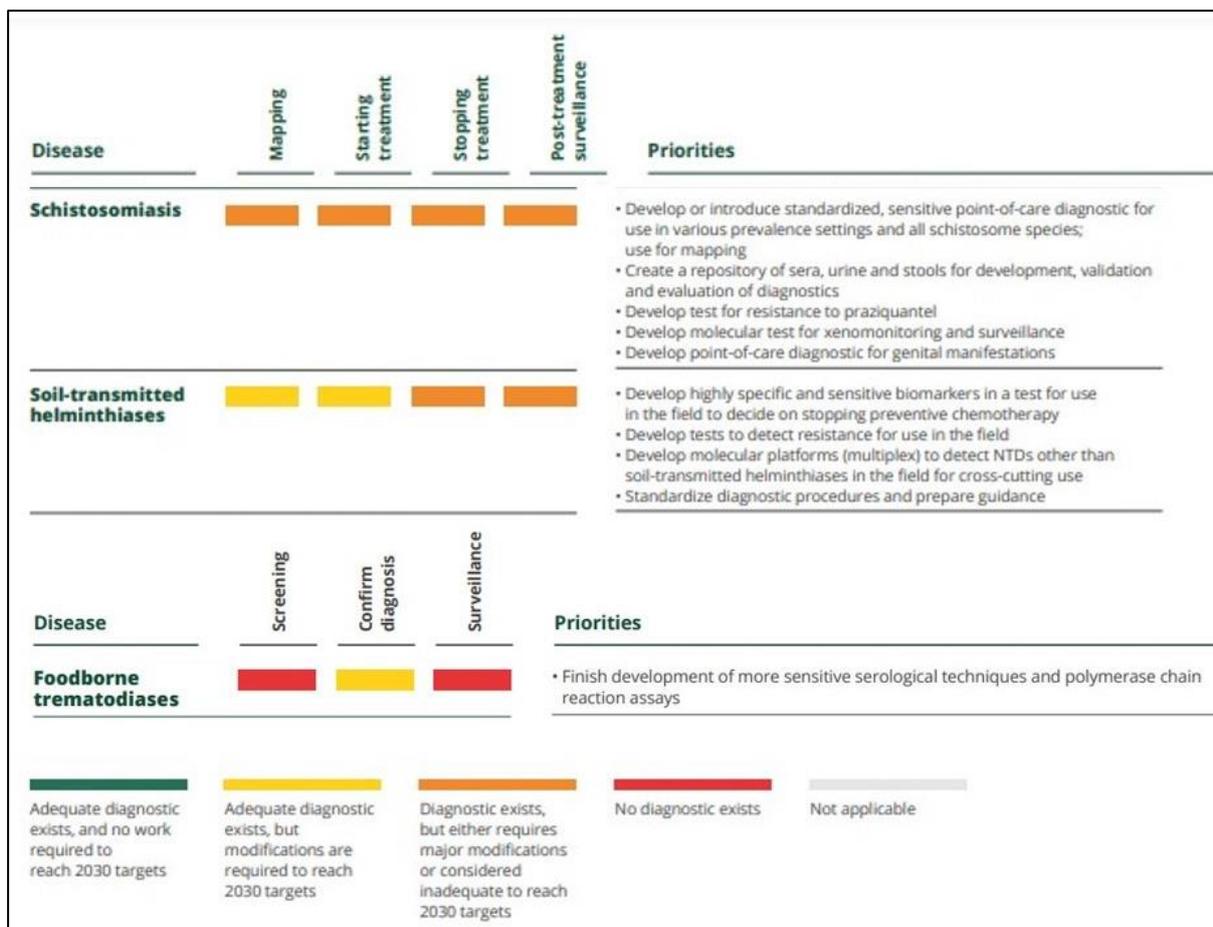


Abbildung 1: Neglected tropical diseases, Resistenztestungen, point-of-care-Diagnostik und Screening-Tests sind Ziele der WHO bis 2030 [109]

3.2. *Fasciola*

3.2.1. *Übertragung und Verbreitung*

Die Fasziose ist eine weit verbreitete parasitäre Erkrankung bei Menschen und Tieren, welche über Kontamination von Lebensmitteln, Trinkwasser oder ungewaschener Brunnenkresse, durch Kot von Schafen, Rindern, aber auch Eseln, Schweinen und Büffeln übertragen wird [55][68][69]. Die Fasziose zählt zu den NTDs, wobei die Vernachlässigung dieser Erkrankung vermutlich ihren Ursprung darin hatte, dass der Focus lange Zeit auf der tierischen Fasziose lag und so hauptsächlich in der Veterinärmedizin erforscht wurde. Erst Anfang 1990 erlangte die humane Fasziose mehr Beachtung in der öffentlichen Gesundheitsversorgung [69]. Heute sind schätzungsweise etwa 2,4 bis 17 Millionen Menschen weltweit durch *Fasciola* spp. infiziert, [84] wobei die Dunkelziffer vermutlich höher liegt, da aus einigen Ländern mit unzureichender Infrastruktur und Diagnostikmethoden ein realistischer Überblick fehlt. Bei Frauen und Kinder im Alter von 5-15 Jahren sind besonders hohe Infektionsraten zu beobachten [69]. Zu den humanpathogenen Spezies gehören *Fasciola gigantica* (*F. gigantica*) und *Fasciola hepatica* (*F. hepatica*) [56], wobei *F. gigantica* fast ausschließlich in Gebieten in Asien und Afrika vorkommt [121][15] und *F. hepatica* in Afrika, Asien, Ozeanien und Amerika weit verbreitet ist [100][76]. Auf Grund der evolutionär wenig differenzierten Spezies, sind sowohl hybride Formen, als auch ein Vorkommen beider Spezies in derselben Region und teilweise sogar eine zeitgleichen Infektion des selben Wirtes möglich [69].

3.2.2. *Morphologie*

Fascioliadea sind Hermaphroditen, eine Unterscheidung von Geschlechtern gibt es, wie bei den meisten Trematoden, somit nicht. Morphologisch unterscheiden sich *F. hepatica* und *F. gigantica* hauptsächlich durch ihre Länge (siehe Abb. 2). *F. gigantica* weisen einen recht lange, als auch schmale Körperform auf und erreichen eine Körperlänge von 15 mm (Millimeter) bis 75 mm. *F. hepatica* sind etwas breiter und erreichen etwa eine Größe von 15mm bis 30 mm [121]. Damit sind sie im Schnitt deutlich kleiner als *F. gigantica*. Ebenso unterscheiden sich die Eier beider Spezies anhand der Größe, wobei auch die *F. gigantica* Eier größer als die der *F. hepatica* Art sind. Eine Unterscheidung auf Grundlage der Verzweigung der Geschlechtsorgane ist ebenfalls möglich, aber weniger aussagekräftig [69].

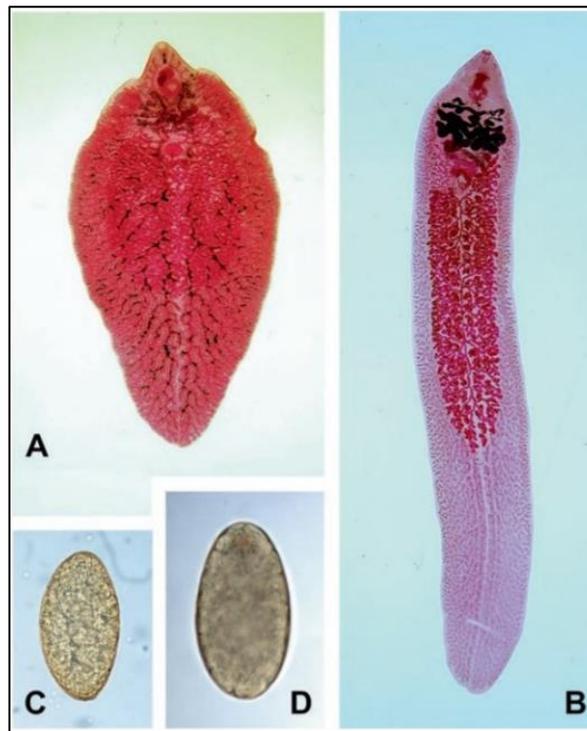


Abbildung 2: verschiedene Stadien und Spezies der *Fasciola*: (A) *F. hepatica* im erwachsenen Stadium (B) *F. gigantica* im erwachsenen Stadium (C) *F. hepatica* im Ei Stadium (D) *F. gigantica* im Ei Stadium [69]

3.2.3. *Lebenszyklus*

Der Lebenszyklus von beiden *Fasciola*-Arten ist sehr ähnlich und erfordert einen Zwischenwirt zur Entwicklung und einen Endwirt, wobei es sich dabei ausschließlich um Säugetiere, wie Rinder und Schafe, und Menschen handelt. Als Zwischenwirt fungieren für *F. gigantica* und *F. hepatica* spezielle Frischwasserschnecken der Familie der Lymnaeidea.

In den Gallengängen der Endwirte befinden sich die erwachsenen Würmer, wo sie Eier ohne Embryonen freisetzen, welche mit dem Kot ausgeschieden und erst bei Kontakt mit Wasser Embryonen entwickeln (Siehe Abb 3, Schritt1). Die frisch geschlüpften Mircadien werden von ihrem jeweiligen Zwischenwirt aufgenommen und gelangen als Sporozysten in die Verdauungsdrüsen der Schnecken. Die Aufnahme der Mircadien muss innerhalb von 24 Stunden erfolgen, da sie in diesem Stadium nicht länger überleben können. Im Zwischenwirt vermehren sich die Sporozysten klonal und werden dann als Zerkarien in Gewässern und an angrenzender Vegetation abgelegt (Siehe Abb 3, Schritt 5). Durch Enzystierung entstehen nun Metazerkarien, welche sehr widerstandsfähig sind und auch bei Temperaturen unter 10 Grad Celsius (°C) bis zu einem Jahr überleben [16]. *F. hepatica* Metazerkarien haben dabei eine deutlich geringere Optimaltemperatur (9 °C bis 26 °C), als dieses Stadium der *F. gigantica* Spezies, welche eine Mindesttemperatur von 16 °C benötigen [71]. Durch den Klimawandel hervorgerufene Temperaturveränderung wirken sich somit in unterschiedlicher Weise auf die verschiedenen Spezies

aus. Metazerkarien setzten sich vor allem an Pflanzen fest, durch deren Verzehr sie in den Magen-Darmtrakt der Endwirte gelangen. Durch die Magensäfte werden die Metazerkarien exzystiert und gelangen durch Penetration der Dünndarmwand in die Leber [55]. Dort entwickeln sich im Leberparenchym die *Fasciola*-Larven. Unreife *F. hepatica*-Würmer wandern nun bis zu 5-6 Wochen und *F. gigantica*-Würmer bis zu 11 Wochen durch die Leber und legen nach abgeschlossener Reifung ihre Eier in die Gallengänge ihrer Wirte, von wo aus sie nun wieder über den Kot ausgeschieden werden [72] [16].

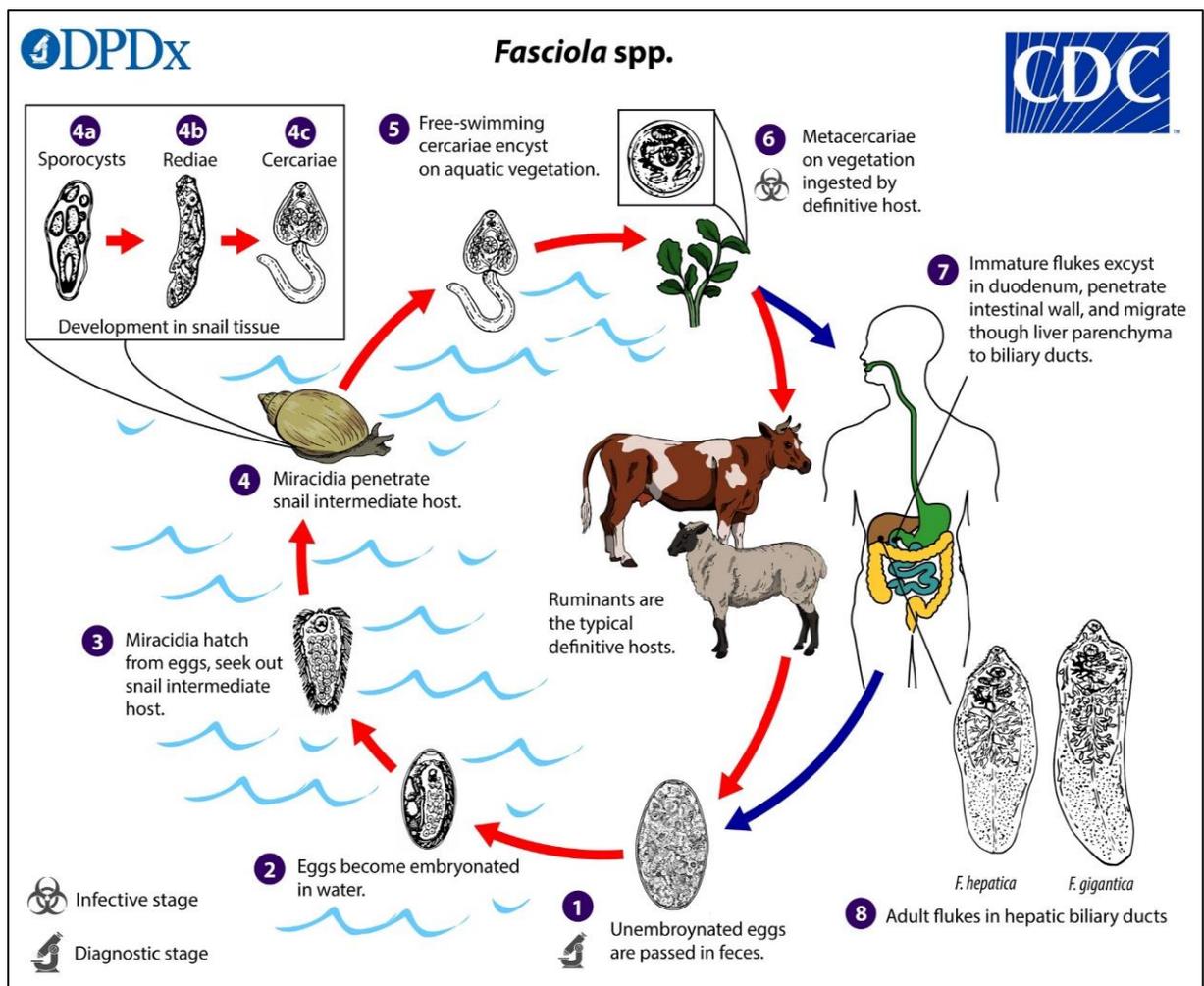


Abbildung 3: Lebenszyklus *F. gigantica* und *F. hepatica* [121]

3.2.4. *Rolle in der Veterinärmedizin*

Von der tierischen Fasziole sind hauptsächlich Hauswiederkäuer, vor allem Schafe, Ziegen und Rinder, betroffen. Aber auch Büffel, Pferde, Schweine, Wildwiederkäuer, Kaninchen, Hasen, Nutrias und andere Herbivoren können infiziert werden. Wie der Mensch fungieren die Tiere als Endwirt und scheiden die Eier über ihren Kot aus. Oft bleiben die Tiere asymptomatisch, es sind aber auch Symptome wie Anämie, Unfruchtbarkeit, submandibuläre Ödeme und eine verringerte Milchproduktion möglich. Die adulten Würmer können über mehrere Monate bis Jahre in der Leber der Tiere überleben. Schafen kommt dabei eine besondere Bedeutung zu, da in ihnen die Würmer über mehrere Jahre überleben können und eine hohe Eiproduktion vorweisen. Die Eier werden über einen langen Zeitraum ausgeschieden und können somit große Mengen Trinkwasser kontaminieren und für eine schnelle und weite Ausbreitung der Parasiten sorgen.

Eine Infektion ist sowohl im Kot, mittels Mikroskopie der Eier, möglich oder auch bei Milchkühen durch Nachweis von spezifischen Antikörpern in der Milch [33]. Eine 4-mal jährliche Behandlung von Weidevieh mit einem Anthelminthikum, meist wird Triclabendazol verwendet, ist eine wirksame Prophylaxe, jedoch besonders in Ländern des globalen Südens oft auf Grund von hohen Kosten und mangelnder Verfügbarkeit des Wirkstoffs nicht flächendeckend durchführbar [86]. Die Auswirkung auf die Landwirtschaft und die damit einhergehende Nahrungsmittelversorgung, besonders in ländlichen Gebieten, ist durch die weitverbreitete Infektion von Weidevieh und die Möglichkeit der Übertragung auf den Menschen relevant und verschärft Hunger und Armut in diesen Gebieten.

3.2.5. *Klinische Charakteristika der humanen Fasziole*

Die menschliche Fasziole ist zumeist in der Leber und den Gallengängen lokalisiert, kann aber auch andere Körperregionen befallen. Im akuten Stadium der Infektion reichen die Symptome von Fieber, Dyspepsie, Anorexie, Aszites und Urtikaria bis hin zu Hepato- und Splenomegalie und Ikterus [90]. Bei Befall der Gallengänge, der chronischen Phase, kann es dann weiterhin zu Gallenkoliken, Intoleranz gegenüber fetthaltigen Nahrungsmitteln, Cholangitis und akuter Pankreatitis kommen [121][15][72]. Bei Kindern scheint besonders der Bauchschmerz und ein Ikterus im Mittelpunkt der Symptome zu stehen [63]. *F. hepatica* entwickelt sich um etwa 1-2 Wochen verzögert im Vergleich zu *F. gigantica*. Grund dafür ist vermutlich, dass adulte *F. gigantica*-Würmer im Schnitt größer sind, die genetischen Unterschiede sind allerdings zu gering, um die erhöhte Pathogenität zu begründen [2].

Im Labor findet sich bei den meisten Patienten eine ausgeprägte Eosinophilie und durch Befall der Leber und der Gallengänge auch ein Anstieg der alkalische Phosphatase und der Globulinwerte (Hypergammaglobulinämie) [61].

3.2.6. Diagnostik

Aktuell kann eine Fasziole durch den mikroskopischen Nachweis von *Fasciola*-Eiern in einer Stuhl- oder Urinprobe detektiert werden (jedoch schließt ein negatives Untersuchungsergebnis die Erkrankung nicht aus, da häufig keine Eier nachweisbar sind), auf Grundlage eines positiven serologischen Tests, oder mittels Polymerase-Kettenreaktion (Polymerase chain reaction, PCR) nachgewiesen werden. PCR ist ein weit verbreitetes *in vitro* Verfahren, das durch Amplifikation von Gen-Sequenzen eine sehr sensitive, aber auch aufwendige und teure Diagnostik erlaubt. Radiologische Befunde insbesondere im Bereich der Leber können zusätzliche Hinweise auf eine Fasziole geben [91][51]. Die verschiedenen Methoden unterscheiden sich sowohl in den Kosten der Durchführung, dem Aufwand und der nötigen Ausbildung des Laborpersonals, der Sensitivität und Spezifität. Von besonderer Relevanz ist der Zeitpunkt ab wann eine Infektion erkannt werden kann. So ist mit einem ELISA-Test serologisch eine Infektion bereits nach 6 Wochen nachweisbar, eine PCR wird etwa nach 7 Wochen positiv und durch Mikroskopie einer Stuhlprobe sind erst nach 9 Wochen Eier der *Fasciola* zu finden [14].

Mikroskopischer Einachweis

Ein Nachweis der Eier ist erst ab einem fortgeschrittenerem Stadium der Infektion möglich, da sich diese bereits im Stuhl befinden und zudem auch in ausreichender Menge vorhanden sein müssen, um mikroskopisch detektierbar zu sein. In humanen Endemiegebieten sind bei akuten Infektionen häufig Konzentrationen von 400 bis 5000 Eiern pro Gramm Stuhl (eggs per gram of stool, EPG) zu finden [69]. Aufgrund der nicht eindeutig erkennbaren Unterschiede der Eier von *F. hepatica* und *F. gigantica* ist eine Unterscheidung der Spezies mittels Mikroskopie in der Regel nicht möglich. Teilweise gleichen sich die Eier unterschiedlicher Trematoden, was zusätzlich zu Verwechslungen verschiedener Gattungen führen kann [52].

Für den mikroskopischen Nachweis der Eier im Stuhl gibt es verschiedene Aufbereitungsmöglichkeiten. Für *F. hepatica* hat sich die verhältnismäßig neue Methode FLOTAC als am geeignetsten erwiesen [114]. FLOTAC ist eine Methode zur Aufbereitung von frischen und konservierten Stuhlproben von Menschen und Tieren. Eine kopromikroskopische Untersuchung der Probe ist sowohl quantitativ als auch qualitativ möglich. Die Aufbereitung der Proben benötigt dabei nur zwischen 12 und 15 Minuten (min). FLOTAC weist eine höhere Empfindlichkeit und

Genauigkeit auf als die sonst meistgenutzte Kato-Katz-Technik, welche sowohl in der Diagnostik von *F. hepatica*, als auch *F. gigantica* angewandt wird [23].

Die Kato-Katz Technik wird überwiegend bei Infektion mit *Schistosoma* angewandt, ist aber auch zur Detektion von Eiern anderer Helminthenspezies möglich. Dafür wird eine geringe Menge der Stuhlprobe auf einen Objektträger gestrichen und nach einer Wartezeit von 40 bis 60 Minuten unter einem Lichtmikroskop nach Eiern durchsucht [11].

Serologie

Auch schon bevor Eier der Parasiten im Stuhl zu finden sind, kann über serologische Tests, wie dem ELISA-Test eine Infektion erkannt werden. Sowohl akute als auch chronische Infektionen können erkannt werden. Dazu wurden verschiedene Antigene beschrieben, welche spezifisch für *Fasciola*-Infektionen sind, wie beispielsweise Kathepsine und 27-kDa-Proteine. Eine Untersuchung des Blutes zur Erfolgskontrolle einer Behandlung der Infektion ist allerdings nicht sinnvoll, weil es oft bis zu 4-5 Monate dauert, bis die Ergebnisse auch nach der Therapie wieder negativ sind. Zudem sind länger zurückliegende Infektionen oft nicht von akuten zu unterscheiden, da Antikörper über lange Zeiträume persistieren können [67][106]. In Endemiegebieten spielt die Serodiagnostik daher eine nur untergeordnete Rolle. Sie wird vor allem zu Untersuchung von Reiserückkehrer*innen aus Endemiegebieten oder Emigrant*innen genutzt.

Molekulare Diagnostik

PCR ist die wohl aufwendigste, aber auch sensitivste und spezifischste Methode zum Nachweis einer Infektion durch *Fasciola* und einer Differenzierung der jeweiligen Spezies. Sie dient zur Amplifikation von Gen-Sequenzen innerhalb einer vorliegenden DNA. Es handelt sich dabei um ein weit verbreitetes in-vitro Verfahren in der Diagnostik und Forschung.

Die DNA Proben können aus dem Stuhl extrahiert werden, jedoch muss dieser vor der Untersuchung durch verschiedene Aufreinigungsschritte bearbeitet werden [64]. Anschließend werden einzelne Nukleotidsequenzen der DNA - Proben mit Hilfe von spezifischen Primern amplifiziert und sequenziert. Die daraus resultierenden Amplifikate können mit bereits bekannten Sequenzen verglichen werden, oder anhand ihrer Größe durch eine Agarose Gelelektrophorese getrennt und identifiziert werden. Eine schnellere, aber auch teurere Variante der herkömmlichen PCR ist, die real time PCR. Hierbei wird durch Beigabe von Fluoreszenzstoffen eine quantitative Bestimmung der DNA erreicht und es ist möglich den Schritt zur Trennung der Amplifikate durch Gelelektrophorese einzusparen. So wird der Arbeitsaufwand verringert, allerdings sind sowohl die Reagenzien, als auch die Geräte teurer [4]. Eine Routinediagnostik mittels PCR

ist vielerorts nicht möglich, da besonders in Ländern mit unzureichender Infrastruktur oftmals Labore mit der nötigen Ausstattung und dem ausgebildeten Personal rar sind [64]. Zudem ist die Durchführung verhältnismäßig teuer und zeitintensiv [55]. Da es für die *Fasciola*-Infektion keine kommerziell verfügbaren PCR-Testkits gibt, ist auch in der westlichen Welt eine PCR-basierte Diagnostik nur in wenigen Referenzlaboren routinemäßig möglich.

Bildgebende Verfahren

Die Diagnose kann mittels bildgebender Verfahren, wie Ultraschall, Kontrastmittel-Ultraschall und Röntgen, gestellt oder gestützt werden. Dazu wird nach unspezifischen Zeichen wie Hepatosplenomegalie und Lymphadenopathie gesucht. Spezifischer für eine *Fasciola*-Infektion sind aber beispielsweise hypodense abszedierende Areale, welche im zeitlichen Verlauf die Leber durchwandern und sowohl im Ultraschall als auch im Röntgen zu erkennen sind. Dabei ist aber zu beachten, dass dieses Zeichen oft inkonstant auftritt. Ein fehlender Nachweis dieser Zeichen ist also kein Ausschlusskriterium für die Erkrankung.

In manchen Fällen ist es auch möglich adulte *Fasciola*-Würmer mit einer Länge von 1,5-3 cm im Ultraschall zu erkennen. Im weiteren Verlauf der Erkrankung, kann es zu intermittierenden Gallenkoliken und weiter zu Cholezystitis, Cholangitis, Ulzerationen und Cholezystolithiasis kommen, welche dann wiederum durch Wandverdickungen oder Lamellierung der Gallenblase ebenfalls im Ultraschall zu erkennen sind [28].

3.2.7. Therapie der Fasziole

Als einziges hochwirksames Medikament steht zur Behandlung der Fasziole Triclabendazol, ein Benzimidazol-Derivat, zur Verfügung [3][39]. Dieses wird einmalig mit einer Dosierung von 10 mg/kgKG (Kilogramm Körpergewicht) mit einer fetthaltigen Mahlzeit verabreicht. Die Heilungsrate liegt dabei zwischen 75-100% [106]. Triclabendazol wird seit 1986 zur Behandlung humaner Fasziole angewendet und ist für Kinder ab 6 Jahren durch die FDA (Food and Drug Administration, US-Behörde für Lebens- und Arzneimittel) zugelassen [65]. Bei ausbleibendem Therapieerfolg, kann die Gabe wiederholt werden [28]. Resistenzen gegen Triclabendazol werden immer häufiger zu einem Problem, da das Medikament in der Vergangenheit sehr oft in der Veterinärmedizin verwendet wurde, um Viehbestände nicht zu gefährden [106]. Seit 1977 zählt Triclabendazol zu der Liste der unentbehrlichen Medikamente („essential drugs“) der WHO. Hierzu zählen Medikamente, welche für die dringlichste medizinische Versorgung

einer Bevölkerung benötigt werden und daher als Richtlinie für die Versorgungsstandards einzelner Staaten dient [111]. Dennoch ist die Beschaffung von Triclabendazol, auch in Europa weiterhin oft schwierig.

Die Gabe von Triclabendazol ist sehr nebenwirkungsarm, auch bei mehrmaliger Gabe sind unerwünschte Nebenwirkungen selten. In wenigen Fällen wurde von Bauchschmerzen, Kopfschmerzen, Übelkeit, Schwäche oder Pruritus berichtet [101][105].

3.3. *Schistosoma*

3.3.1. *Übertragung und Verbreitung*

Die Schistosomiasis gehört zu den am weitesten verbreiteten parasitären Erkrankungen der Welt. Weltweit betrifft sie über 200 Millionen Menschen und ca. 779 Millionen Menschen leben im Risiko, sich zu infizieren [115]. Damit ist die Schistosomiasis, unter den parasitären Erkrankungen, hinsichtlich des Infektionsrisikos an zweiter Stelle hinter Malaria [98]. In der vorliegenden Arbeit habe ich mich mit den beiden Arten *S. japonicum* und *S. mansoni* beschäftigt, welche neben *Schistosoma haematobium* (*S. haematobium*) zu den 3 häufigsten humanpathogenen Spezies der *Schistosoma* zählen [48].

Schistosoma gehören zum Stamm der Plattwürmer, der Klasse der Trematoden und Familie der *Schistosomatidea*. *S. mansoni* kommt vor allem in Subsahara-Afrika und Südamerika vor und *S. japonicum* hauptsächlich in China, den Philippinen und auf der indonesischen Insel Sulawesi. Für die Verbreitung spielen Haustiere und Wildtiere als die Hauptüberträger der *Schistosoma* eine bedeutende Rolle, die Übertragungen der humanen Schistosomiasis findet aber über infiziertes Süßwasser statt. 2014 wurden etwa 61.6 Millionen Menschen auf Grund von diagnostizierter Schistosomiasis behandelt und 258 Millionen Menschen erhielten eine präventive Behandlung aufgrund des Verdachtes auf eine Infektion durch *Schistosoma* [48].

3.3.2. *Morphologie*

Im Gegensatz zu vielen anderen Trematoden sind adulte *Schistosoma* getrenntgeschlechtlich. Morphologisch unterscheiden sich *S. mansoni* und *S. japonicum* hauptsächlich durch ihre Größe, welche ebenfalls je nach Alter und Geschlecht der Würmer variiert. *S. mansoni* sind in der Regel deutlich kürzer als *S. japonicum*, ebenso sind die männlichen Würmer meist deutlich kleiner als die schlanken, langen weiblichen Exemplare. Männliche *S. mansoni* Würmer erreichen eine Körperlänge von 9 mm bis 12 mm, mit einer Breite von 1 mm bis 1.2 mm und weibliche Würmer sind etwa 12 mm bis 16 mm lang und 0.15 mm bis 0.2 mm breit. *S. japonicum* Männchen messen etwa 12 mm bis 20 mm in der Länge und 0.5 mm in der Breite und die Weibchen sind 18 mm bis 25 mm lang und 0.2 mm bis 0.3 mm breit. Im Gegensatz zu Körperlänge sind die Eier der *S. mansoni* Spezies mit 140 µm in der Länge und 60 µm in der Breite größer als die der *S. japonicum* Spezies mit 81 µm in der Länge und 63 µm in der Breite.

Farblich fallen männliche *Schistosoma* makroskopisch deutlich heller und durchscheinender auf als die weiblichen, eher schwarzen *Schistosoma*. Im altgriechischen wird „*schistoma*“ mit „getrennter Körper“ übersetzt, was sich auf die Morphologie der adulten Männchen bezieht, welche am seitlichen Körperwand eine Furche besitzen, den gynecophoralen Kanal, in welchem das Weibchen nach der Kopulation verbleibt.

3.3.3. *Lebenszyklus*

Schistosoma-Eier werden über den Stuhl oder Urin ausgeschieden, im Falle von *S. mansoni* und *S. japonicum* hauptsächlich mit dem Stuhl (Siehe Abb 4, Schritt 1). Unter geeigneten Bedingungen schlüpfen die Eier und setzen Miracidien (engl. „miracidia“) frei, welche schwimmen können und sich daher besonders in Gewässern schnell verbreiten können. Die Miracidien dringen in Schnecken als Zwischenwirte ein. *S. mansoni* nutzen Schnecken der Gattung *Biomphalaria* als Zwischenwirt, *S. japonicum* Schnecken der Gattung *Oncomelania* [62]. In der Schnecke entwickeln sich die Miracidien zu Sporozysten und beginnen mit der Produktion von Zerkarien (Siehe Abb 4, Schritt 3-5). Diese können nach Freisetzung aus der Schnecke über die Haut in menschlichen Wirt eindringen, von wo sie zu *Schistosomulae* werden. Die *Schistosomulae* wandern über den venösen Kreislauf in die Lunge, dann in das Herz, und in der Leber, welche sie nach abgeschlossenem Reifungsprozess über das Pfortadersystem verlassen. Männliche und weibliche erwachsene Würmer kopulieren, wobei weibliche Würmer diese Paarung zur physischen und reproduktiven Entwicklung benötigen, wohingegen männliche Würmer sich auch ungepaart entwickeln können. Die adulten Würmer halten sich in den Mesenterialvenen auf. *S. japonicum* wird häufiger in den oberen Mesenterialvenen gefunden, die den Dünndarm drainieren, und *S. mansoni* kommt häufiger in den unteren Mesenterialvenen vor, die den Dickdarm drainieren. Beide Arten können jedoch beide Orte besetzen und sind in der Lage, zwischen den Orten zu wechseln. Ohne antiparasitäre Behandlung können adulte Würmer über mehrere Jahre im menschlichen Verdauungstrakt überleben. In mehreren epidemiologischen Studien wurde von Überlebenszeiträumen zwischen 5 bis 30 Jahren berichtet [62]. Die Weibchen, legen ihre Eier in den kleinen Venolen des Portal- und Perivesikalsystems ab, von wo diese allmählich in Richtung des Darmlumens bewegt werden und mit dem Kot ausgeschieden werden. *S. mansoni* Weibchen können bis zu 300 befruchtete Eier pro Tag ablegen. Im Ei-Stadium können sie mittels Mikroskopie identifiziert werden. Dieser Transmissionszyklus schließt sich, sobald der menschliche Wirt über seine Exkremente, wie Stuhl oder Urin, welche die Eier der Parasiten enthalten, Frischwasser kontaminiert und die Parasiten schlüpfen können [115].

Der komplexe Lebenszyklus der *Schistosoma* ist in mehreren Entwicklungsstadien vom Klimawandel abhängig. Sowohl Wasser- und Lufttemperatur als auch physiochemische Einflüsse auf Wasserqualität werden durch den Klimawandel verändert. Zum aktuellen Zeitpunkt ist aber noch nicht vorauszusagen ob sich dies eher als verstärkte Verbreitung oder Reduzierung der *Schistosoma* Transmissionen äußern wird [10][70].

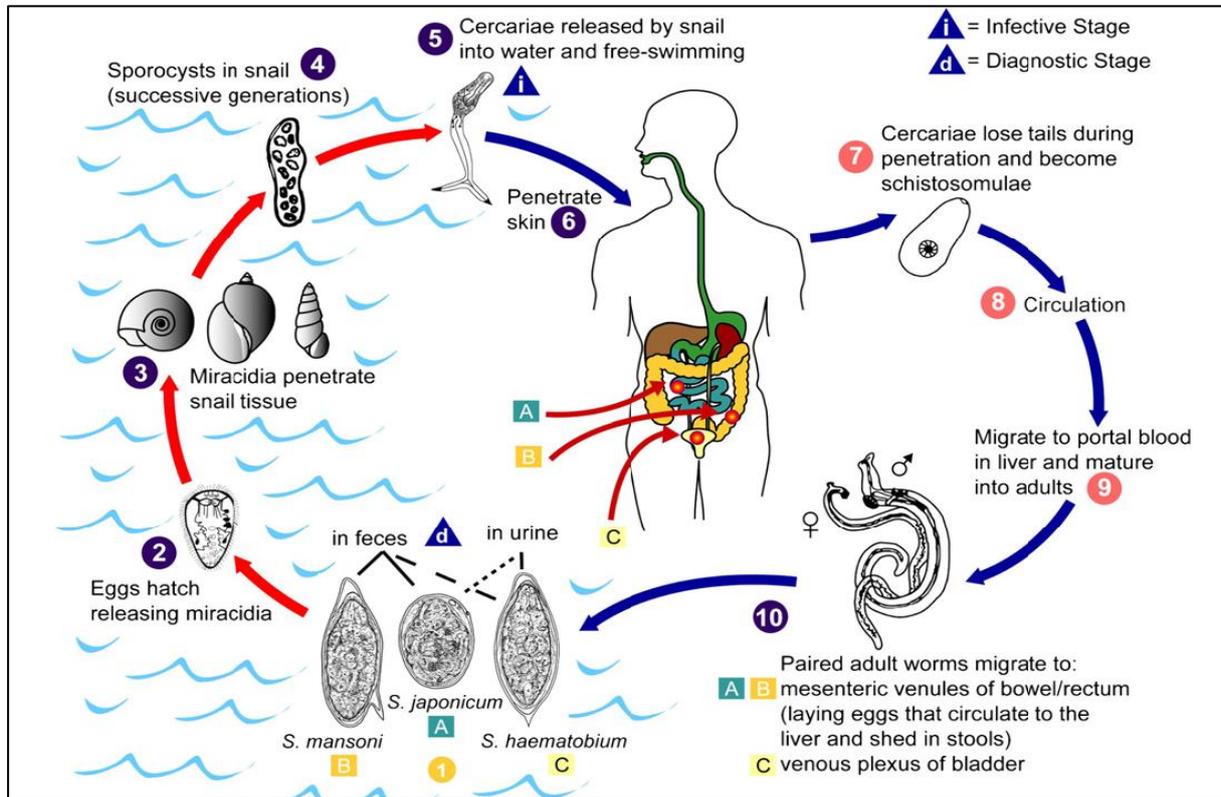


Abbildung 4: Lebenszyklus der *S. mansoni*, *S. japonicum* und *S. haematobium* [117]

3.3.4. Rolle in der Veterinärmedizin

In der Veterinärmedizin wird die Schistosomiasis hauptsächlich bei Rindern und Wasserbüffeln beobachtet, aber auch bei Hausschweinen und Hunden ist eine Infektion möglich. Besonders zur Regenzeit verbreiten sich die Infektionen durch *Schistosoma* unter Weidevieh sehr schnell, da die ausgeschiedenen Eier ihre Zwischenwirte in Gewässern finden und der Kot bei starken Regenfällen schnell verbreitet wird [74]. Ein Nachweis der Infektion erfolgt über eine mikroskopische Untersuchung des Stuhls auf *Schistosoma*-Eier. Auf den Menschen können die Parasiten durch kontaminierte Lebensmittel übertragen werden, ein enges Zusammenleben zwischen Mensch und Tier stellt somit ein stark erhöhtes Risiko dar [45]. Die Behandlung mit Praziquantel ist auch bei Tieren möglich, wird aber aktuell nicht zur Prävention eingesetzt.

3.3.5. *Klinische Charakteristika der humanen Schistosomiasis*

Die Schistosomiasis, auch bekannt als Bilharziose (nach Theodor Bilharz, dem Erstbeschreiber der *Schistosoma* im Menschen, 1851), kann sehr unterschiedliche Verlaufsformen aufweisen. Viele Infektionen verlaufen asymptomatisch. Die Symptome der Schistosomiasis werden nicht durch die Würmer selbst verursacht, sondern durch die Reaktion des Körpers auf deren Eier. Das Eindringen der Zerkarien in die Haut kann mitunter zu vorübergehendem Juckreiz führen, aber auch unbemerkt bleiben, da die Zerkarien sehr klein sind und mit dem bloßen Auge nicht zu erkennen sind [62]. Die akute Schistosomiasis ist eine systemische Überempfindlichkeitsreaktion, welche meist wenige Wochen nach der Erstinfektion auftreten kann und mit Fieber (sogenanntes Katayama-Fieber), Diarrhö und Bauchschmerzen einhergehen kann [34][89]. Eine über Jahre bestehende Schistosomiasis kann zu schweren Langzeitkomplikationen führen, da es zu einem Umbau des Gewebes in befallenen Organen kommt, die meist durch die immunologische Reaktion auf die dortige Ablage von *Schistosoma*-Eiern erklärt werden kann. Klinisch manifestiert sich die chronische Schistosomiasis häufig durch Aszites, Ösophagusvarizen und Leberzirrhosen. Auch eine Lungenbeteiligung ist in 20% der Fälle nachzuweisen [47]. Besonders für schwangere Frauen kann eine durch die Infektion ausgelöste Anämie erhöhte Blutungsneigung unter der Geburt lebensbedrohlich sein. Kinder im Vorschulalter und Frauen im gebärfähigen Alter sind oft von schwerwiegenden Infektionen bedroht. Entwicklungsverzögerungen sowohl in der Schwangerschaft als auch bis hin zum frühen Schulalter sowie verstärkte Blutungen unter der Geburt gefährden diese Patient*innengruppe besonders.

In seltenen Fällen kann es auch zu Neuroschistosomiasis mit intrazerebralen Granulomen (häufig durch *S.japonicum*) und spinalen Granulomen (häufig *S.mansoni*), welche ihre Symptomatik hauptsächlich durch die raumfordernde Wirkung hervorrufen.

In Gebieten, in denen sowohl HIV/AIDS, als auch Schistosomiasis endemisch sind, scheint die Schistosomiasis ein Kofaktor für die Ausbreitung und das Fortschreiten von HIV/AIDS zu sein, vermutlich am ehesten durch eine erleichterte Übertragung [94].

3.3.6. Diagnostik

Da *Schistosoma*-Eier meist über den Stuhl und den Urin ausgeschieden werden, erfolgt der Nachweis einer Darmschistosomiasis in der Regel aus Stuhlproben, ist aber in manchen Fällen auch über das Sputum, Biopsiematerial oder als Serologie oder PCR Nachweis aus dem Blut möglich [47] [110]. Ebenso wie bei der Fasziole ist ein rasches Handeln wichtig zur erfolgreichen Therapie der Infektion. Eine präzise und frühzeitige Diagnostik ist daher von zentraler Bedeutung. Dabei erfüllen die aktuell vorhandenen Diagnostikmethoden diese Anforderungen nicht in gleichem Maße. Wie in Kapitel 3.2.6.) bereits beschrieben, ist der PCR-Test der zeitintensivste und gleichzeitig sensitivste verfügbare Test. Auf Grund der hohen Kosten und des anspruchsvollen Testaufbaus wird er aber selten außerhalb von Studien durchgeführt, in Endemiegebieten stellt die Mikroskopie den wichtigsten diagnostischen Test dar.

Mikroskopischer Einachweis

Zur mikroskopischen Untersuchung von Stuhlproben, die weiterhin als Goldstandard in der Diagnostik der intestinalen Schistosomiasis gilt, wird meist die Kato-Katz Technik verwendet, da sie eine Quantifizierung der Infektionen anhand der Eizahl in EPG zulässt (siehe Kapitel 3.2.6.). Die Eier sind unter dem Mikroskop leicht durch ihre Größe, Form und den typischen Seiten- oder Endstachel zu erkennen und zu identifizieren. Allerdings können sie bei leichten Infektionen schnell übersehen werden, da es selbst durch verschiedene Konzentrationstechniken, wie Sedimentation in einer Glycerinlösung oder Zentrifugation oft schwierig ist, Eier nachzuweisen [35]. Mit einer Sensitivität von 30 EPG sind meist serielle Tests nötig, um Infektionen nicht zu übersehen [72]. Zum Nachweis von Eiern im Urin wird ein Urinsediment, welches um die Mittagszeit oder nach körperlicher Anstrengung gewonnen wurde, mikroskopiert, allerdings lassen sich in der Regel selbst *S. haematobium* auch im Stuhl nachweisen.

Serologie

Für den serologischen Nachweis von Antikörpern in der Diagnostik der Schistosomiasis werden meist IgG, IgM oder IgE durch indirekte Hämagglutination oder Immunfluoreszenz nachgewiesen. Antikörperbasierte Testungen sind zwar sehr empfindlich, allerdings fällt die Unterscheidung zwischen einer früheren Exposition und einer aktiven Infektion oft schwer. Die Serokonversion erfolgt in der Regel innerhalb von 4-8 Wochen nach der Infektion, das Intervall kann aber auch das 4- bis 6-fache betragen [50]. Viele Tests liefern noch bis zu 2 Jahre oder länger positive Ergebnisse. Zudem kann es zu Kreuzreaktionen mit anderen Helminthen kommen, sodass eine Differenzierung nicht immer möglich ist.

Bei bis zu 90% aller Patienten mit Schistosomiasis konnten Antigene wie das zirkulierende anodische Antigen (circulating anodic antigen) und zirkulierendes kathodisches Antigen (circulating cathodic antigen) nachgewiesen werden [120]. Diese können mittels markierter monoklonaler Antikörper im Serum oder Urin nachgewiesen und quantifiziert werden. Allerdings ist auch diese Methode, vor allem bei leichten Infektionen, nicht ausreichend sensitiv. Es ist auch ein point of care Test verfügbar (POC-CCA point-of-care circulating cathodic antigen), welcher auch außerhalb von Laboren durchgeführt werden kann und auf der Detektion des zirkulierenden kathodischen Antigens im Urin beruht [17][9]. Jedoch wird dieser Test aktuell noch aufgrund geringer Spezifität eher im Einzelfall als eine nützliche, zusätzliche direkte Nachweismethode für Reiserückkehrer*innen aus Endemiegebieten oder Emigrant*innen verwendet [118].

Molekulare Diagnostik

Die molekulargenetische Untersuchung mittels PCR ist in der Routinediagnostik der Schistosomiasis nicht sehr weit verbreitet, obwohl sie eine höhere Sensitivität hat als der mikroskopische Nachweis der Eier im Stuhl oder im Urin [44]. Besonders für Patient*innen mit einer geringen Erregerlast ist die PCR eine wichtige Diagnostik, da auch bei leichten Infektionen ein Nachweis möglich ist. Der Nachweis ist sowohl im Stuhl, im Urin, als auch im Blut möglich. Wie auch in der Diagnostik der Fasziole (siehe Kapitel 3.2.6.3.), ist die PCR sowohl die zeitaufwendigste, als auch teuerste Methode zum Nachweis einer Schistosomiasis [96].

Bildgebende Verfahren

In der bildgebenden Diagnostik werden die Zystoskopie und die Endoskopie eingesetzt, um Läsionen in der Blase oder Ösophagusvarizen nachzuweisen. Gastro-ösophageale Varizen und Pfortadererweiterungen, welche für eine hepatische Schistosomiasis sprechen, könne zudem auch mit Hilfe kontrastmittelgestützter Radiografie nachgewiesen werden. Zum Nachweis von periportal Fibrosierungen und granulomatösen Entzündungen in der Leber stehen Verfahren wie die Laparoskopie und gegebenenfalls eine histologische Untersuchung von Biopsien zur Verfügung [42].

Zur Diagnose der Neuroschistosomiasis dient auch der Nachweis von Läsionen im CT, Myelographie und MRT.

In der Felddiagnostik werden zunehmend auch tragbare Ultraschall-Geräte verwendet, um eine bildgebende Diagnostik auch außerhalb großer medizinischer Einrichtungen zu gewährleisten [46].

3.3.7. *Therapie der Schistosomiasis*

Mitte der 1970er Jahre wurde von den Firmen Bayer AG und MERCK KGaA in Deutschland das Medikament Praziquantel entwickelt und konnte durch zahlreiche Studien in seiner Wirksamkeit gegen die Schistosomiasis bestätigt werden. Praziquantel wirkt über einen vermehrten Calciumeinstrom durch die Wurmoberfläche, welche zu Muskelkontraktionen und spastischen Paralysen mit anschließendem Versterben des Wurmes führt. Die WHO empfiehlt eine einmalige Gabe von 40 mg/kgKG, was die Compliance im Gegensatz zu mehrmaligen Gaben oder Infusionstherapien erhöht. Zudem ist Praziquantel kostengünstig mit einem Preis von unter 0.10 US-Dollar pro Einmalgabe. Die Heilungsrate ist nach einmaliger Gabe hoch und wird zwischen 50% bis 100% angegeben. Eine Re-infektion ist aber auch nach erfolgreicher Behandlung möglich und birgt damit das Risiko, bei häufiger Behandlung mit Praziquantel Resistenzen zu provozieren. An Nebenwirkungen können Bauchschmerzen, Übelkeit, Erbrechen, Diarrhö und in seltenen Fällen Krampfanfälle oder allergische Reaktionen auftreten.

Im Rahmen der 2020 von der WHO veröffentlichten Road Map zur Eliminierung von NTDs "Ending the neglect to attain the Sustainable Development Goals: a road map for neglected tropical diseases 2021-2030" [109] ist eine pädiatrische Dosierung von Praziquantel für Kinder im Vorschulalter in groß angelegte Behandlungskampagnen, auch als präventive Chemotherapie vorgesehen. In einer groß angelegten Studie konnte nachgewiesen werden, dass die regelmäßige Behandlung von Kindern im Schulalter mit Praziquantel die Prävalenz der Bilharziose in Afrika südlich der Sahara in den letzten 20 Jahren um fast 60 % verringert hat [57].

3.4. *MALDI-TOF MS in der mikrobiologisch-infektionsdiagnostischen Anwendung*

Im heutigen klinischen und wissenschaftlichen Alltag wird die MALDI-TOF MS vorrangig zur Diagnostik von Bakterien, Viren und Pilze verwendet [80]. Hierbei wird eine Probe durch Laserbeschuss in ihre Moleküle aufgespalten, ionisiert und durch Beschleunigung in einem elektrischen Feld nach ihrem Masse–Ladungsverhältnis aufgetrennt. Eine Detektormembran analysiert die hierbei zu beobachtenden Proteinspektren und vergleicht sie mit einer internen Datenbank, um anhand der Ähnlichkeit der Spektren zu bereits in der Datenbank hinterlegten Spektren eine Speziesdiagnose zu ermöglichen.

Zur Vorbereitung der Proben wird ein standardisiertes Protokoll verwendet und dient der Extraktion einzelner Proteine aus der Probe (siehe Kapitel 5.2.3.). Ein Teil der Probe wird gemeinsam mit einer Matrix auf einer Metallplatte, der Quelle, kokristallisiert (siehe Abb. 5, 1. Schritt). Als Matrix wird in der Routinediagnostik meistens HCCA, *α -cyano-4-hydroxycinnamic acid*, verwendet, welche die Laserenergie absorbiert und in Wärmeenergie umwandeln kann. Diese Wärmeenergie wird benötigt, um die Probe zu verdampfen und so in die Gasphase zu überführen. Anschließend wird die Probe ionisiert, indem die Ladungen der Matrix auf die Probe übertragen wird. Als Laser wird in der Regel ein gepulster 337-Nanometer (nm)- Laserstrahl verwendet (siehe Abb. 5, 2. Schritt).

Die emittierten Ionen werden mit Hilfe einer Hochspannungsquelle durch einen Vakuumrohr in Richtung des Detektors beschleunigt und anhand ihrer Zeit bis zum Erreichen des Detektors (Time off light (TOF), ihrer Masse und des Ionisierungsgrades (Ladung) aufgetrennt (siehe Abb 5., 3. Schritt). Die Spektren werden visuell in einem Diagramm wiedergegeben. Auf der X-Achse ist das Masse–Ladungsverhältnis (m/z) in Dalton (Da) angegeben (siehe Abb. 5, 4. Schritt). Die Ladung ist meist $z=1$ und kann somit als untergeordnet angesehen werden. Die Identifizierung basiert somit vorrangig auf der molekularen Masse einzelner Proteine einer zu untersuchenden Probe. Die Y-Achse beschreibt die Häufigkeit eines Ions in der gemessenen Probe, dadurch kann die Intensität in einem Massebereich erfasst werden. Für einzelne Spezies lassen sich auf Grundlage der TOF und der Intensitäten charakteristische Proteinspektren generieren [20].

In einem Massenbereich von 2.000 Da bis 20.000 Da wird eine Identifizierung durch Mustervergleich der Spektrenspitzen mit denen einer Referenzspektren-Datenbank erreicht. Dieser Bereich entspricht hauptsächlich ribosomalen und anderen häufig vorhandenen strukturellen Proteinen [12][92].

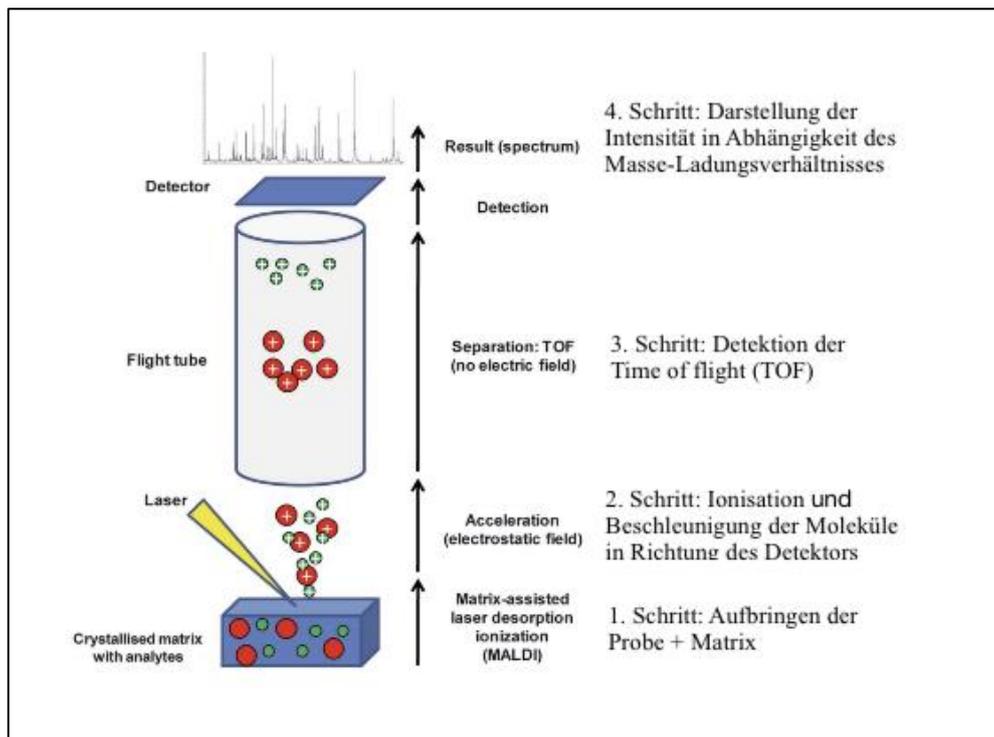


Abbildung 5: MALDI-TOF MS Funktionsweise [24]

In der Diagnostik von Bakterien ist MALDI-TOF MS schon seit mehr als 10 Jahren in der Routinediagnostik angekommen und hat dort alternative Verfahren, wie die auf biochemischen Reaktionen beruhende „bunte Reihe“, zur Erregeridentifikation nahezu vollständig verdrängt. Die kommerzielle Datenbank der Herstellerfirma Bruker Daltonics enthält neben Referenzspektren für Bakterien auch eine Datenbank zur Identifizierung von Pilzen. Im Vergleich zu anderen Methoden der Bakterien- und Pilzdiagnostik bieten besonders die Zeit- und Kostensparnis Vorteile. Eine Identifikation ist in der Regel in unter 24 Stunden möglich, da in vielen Fällen eine direkte Erregerdiagnostik aus einer positiven Blutkultur, ohne vorherige Anzucht durchgeführt werden kann. Bei Sputum-, Eiter-, Urin-, oder Stuhlproben ist eine vorherige Anzucht nötig. In der Kultur gewachsene Bakterienkolonien können dann massenspektrometrisch untersucht und identifiziert werden. Mit einem Preis von etwa 50 Cent pro Speziesidentifizierung ist MALDI-TOF MS sehr kostengünstig und kann viele teure Arbeitsschritte und Materialien einsparen [38]. Auch technisch ist die Durchführung der Analyse einfacher als z.B. eine PCR-Testung.

Eine Antibiotika- und Antimykotikaresistenztestung mittels MALDI-TOF MS ist Gegenstand aktueller Forschung [97].

Auch in der Diagnostik von Parasiten wurden in den letzten Jahren einige Fortschritte erzielt. So ist bereits die Identifikation von Zecken [27], Läusen und Mücken [58] [43][1] erfolgreich

durchgeführt worden. Ebenso wird die Erforschung von MALDI-TOF MS basierter Diagnostik von Helminthen weltweit vorangetrieben. Eine Identifikation von Nematoden und Trematoden konnte bereits in einigen Studien nachgewiesen werden und die Erforschung wird mit dem Ziel der Implementierung in die klinische Routinediagnostik weiterentwickelt. In den letzten Jahren wurden unter Anderem erfolgreiche Studien zur Identifikation von *Anisakis* spp. [66], *Dirofilaria* spp., *Ascaris* spp. [77], *Trichinella* spp. [53], *Taenia saginata* [108] und *Fasciola* spp. [99] durchgeführt.

4. Ziel der Arbeit / Fragestellung

Die MALDI-TOF MS als kostengünstiges, schnelles, speziesspezifisches Tool zur Diagnostik von Bakterien und Pilzen hat auch in der Parasitendiagnostik ein großes Potential. Das Hauptziel der vorliegenden Arbeit ist die Evaluation und Etablierung einer Methodik zur MALDI-TOF MS-gestützten Diagnostik verschiedener Stadien von *Fasciola* spp. und *Schistosoma* spp.

Die folgenden spezifischen Teilschritte wurden hierbei adressiert:

1. Etablierung eines einfachen Protokolls zur Proteinextraktion aus Helminthen-Proben zur nachfolgenden MALDI-TOF-Analytik.
2. Evaluierung der speziesspezifischen Identifizierung von *Fasciola* spp. und *Schistosoma* spp. und Klassifizierung mittels verschiedener Lernalgorithmen (LA) des „Machine Learning“ (ML, dt: maschinellen Lernens) (**Interspezies-Klassifikation**)
3. Untersuchung aller Einflussfaktoren mittels Anwendung verschiedener LA (**Intraspezies-Klassifikation**)
 - a. Untersuchung von folgenden Einflussfaktoren auf die Identifikation von *Fasciola* spp.:
 - i. Einfluss des Entwicklungsstadiums (Eier oder adulte Proben) von *F. gigantica*
 - ii. Einflusses des verwendeten Körperanteils der untersuchten adulten Proben auf die Zuverlässigkeit der Identifikation
 - iii. Einfluss der Lagerungsdauer von *Fasciola* spp. (12 Monate)
 - b. Untersuchung von folgenden Einflussfaktoren auf die Identifikation von *Schistosoma* spp.:
 - i. speziesspezifische Identifikation und der Einfluss speziesspezifischer Unterschiede von *S.mansoni* - und *S. japonicum* - Proben
 - ii. Einfluss des Lagerungsmediums (70% (v/v) Ethanol und RNAlater®)

5. *Material und Methoden*

5.1. *Versuchsaufbau*

Sowohl für die Arbeit mit *Fasciola*- als auch mit *Schistosoma*-Proben war zu Beginn eine Entwicklung einer in-house-Datenbank nötig, um Referenzspektren der jeweiligen Helminthen in der MALDI-TOF-Analytikdatenbank zu „hinterlegen“, anhand derer dann neue Proben identifiziert werden könnten. Hierfür wurde nach einem bereits etablierten Protokoll [99][108] eine Proteinextraktion der jeweiligen Proben durchgeführt. Nach der massenspektrometrischen Messung mit MALDI-TOF MS, unter Verwendung der FlexControl® Software Version 3.4 (Bruker Daltronic) durchgeführt, wurden alle Proben mit der kommerziellen Datenbank von Bruker Daltronic verglichen, um Kontaminationen mit Bakterien oder Pilzen auszuschließen.

Im nächsten Schritt wurde die BDAL-Datenbank von Bruker Daltronic um eine neue in-house-Datenbank erweitert. Hierzu mussten Referenzspektren erstellt werden, welche aus einigen der im ersten Schritt generierten Rohspektren erstellt werden. Dieser Prozess enthält einige Reinigungsschritte, welche mit der FlexAnalysis® Software Version 3.4 (Bruker Daltronic; Bremen, Germany) durchgeführt wurden. Mit der MALDI Biotyper Compass Explorer® Software Version 4.1 (Bruker Daltronic) konnten die gereinigten Rohspektren als MSPs in die in-house-Datenbank hinzugefügt werden und alle weiteren Proben mit ihnen verglichen werden.

Der dritte Teil der Arbeit bestand darin, unterschiedliche Einflussgrößen auf die Identifikation mittels MALDI-TOF MS zu detektieren und deren Einfluss zu untersuchen und zu vergleichen. Zudem wurden verschiedene Lernalgorithmen des ML angewendet, um eine eventuelle Überlegenheit eines LA zur Parasitendiagnostik mittels MALDI-TOF MS festzustellen.

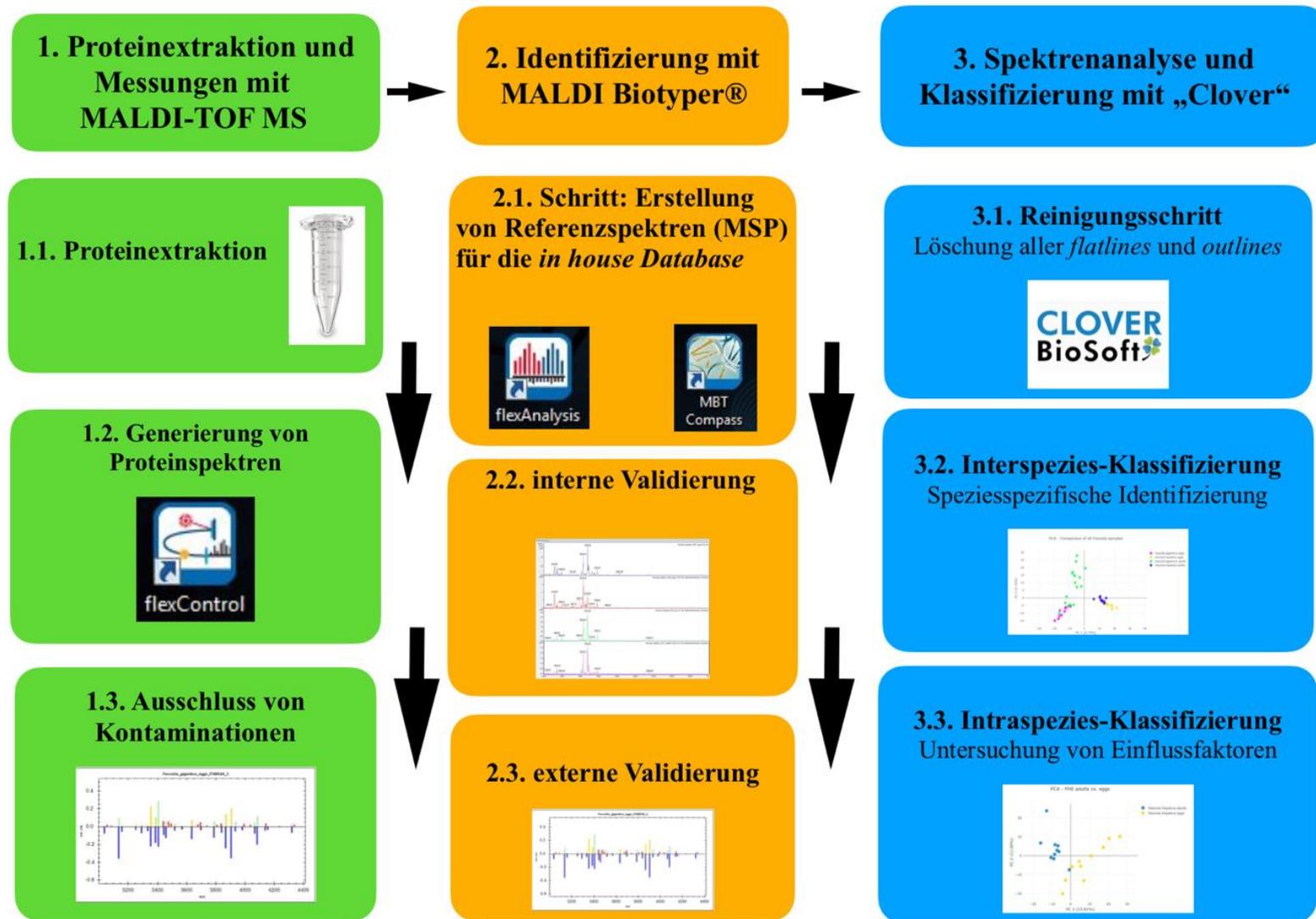


Abbildung 6: Workflow zur Arbeit mit *Fasciola* und *Schistosoma*

5.2. *Probenvorbereitung*

5.2.1. *Probenherkunft*

Fasciola gigantica: Die Proben der adulten *F. gigantica*-Würmer wurden aus Lebergewebe geschlachteter Rinder in Nigeria entnommen. Dafür wurde von der Leitung des Schlachthofes *Kubwa* (Abuja, Nigeria) eine schriftliche Genehmigung für die Entnahme und anschließende Analyse der Proben von geschlachteten Rindern eingeholt (Referenz Nr.: FCT/ARD/TRN/004, 17. Oktober 2019). Die Sammlung der Proben fand im Dezember 2019 und Januar 2020 statt. Alle Proben wurden in 70% (v/v) Ethanol bei Raumtemperatur gelagert und an das Institut für Medizinische Mikrobiologie und Hygiene in Homburg, Deutschland versendet, wo sie bei -20 °C gelagert wurden. Anhand der morphologischen und molekulargenetischen PCR-Untersuchung konnten alle Isolate als *F. gigantica* identifiziert werden.

Fasciola hepatica: Die adulten *F. hepatica*-Proben stammen von infizierten Rindern aus der Schweiz und wurden in einem Schlachthof in der Nähe von Zürich gewonnen. Sie wurden morphologisch beurteilt und identifiziert. Eine PCR-Sequenzierung wurde nicht durchgeführt, da eine Identifizierung und speziesspezifische Differenzierung zwischen *F. hepatica* und *F. gigantica* rein morphologisch möglich ist (siehe Kapitel 3.2.2.).

Schistosoma mansoni und Schistosoma japonicum: Die in dieser Studie verwendeten adulten Würmer von *S. mansoni* (puertoricanischer Stamm) und *S. japonicum* (japanischer Yamashiro-Stamm) wurden aus experimentell infizierten BALBc-Mäusen im Institut für Tropenmedizin und Parasitologie der Dokkyo Medical University (Tokyo, Japan) gewonnen. Die Proben wurden aus den Mesenterial- und Portalvenen der Versuchstiere entnommen und mehrmals mit Phosphatgepufferter Salzlösung (PBS, englisch phosphate-buffered saline) gewaschen. Alle Würmer konnten sowohl morphologisch als auch molekulargenetisch als *Schistosoma* spp. (40 *S. mansoni* und 22 *S. japonicum*) identifiziert werden. Von den 40 *S. mansoni*-Isolaten wurden 19 (7 gemischte Männchen/Weibchen, 6 Männchen und 6 Weibchen) in 70 % (v/v) Ethanol und 21 (11 gemischte Männchen/Weibchen, 5 Männchen und 5 Weibchen) in RNAlater® gelagert. Von den 22 *S. japonicum*-Isolaten wurden 11 (5 gemischte Männchen/Weibchen, 3 Männchen und 3 Weibchen) in 70 % (v/v) Ethanol und 11 (5 gemischte Männchen/Weibchen, 3 Männchen und 3 Weibchen) in RNAlater® gelagert. Für den Transport wurden alle Proben bei -40 °C gelagert, nach Erhalt der Proben in Homburg wurden diese bei -20 °C aufbewahrt.

5.2.2. Lagerungsmedien

RNAlater®: RNAlater® enthält einen organischen Eisenchelator, welcher Metalloproteasen durch Abspaltung von Zink und Kobalt deaktiviert. Zudem hat RNAlater® eine hohe Salzkonzentration, welche eine sehr gute Proteinerhaltung bewirkt und wahrscheinlich ebenso die proteolytischen Aktivitäten in den Zellen hemmt. Wie in der Arbeit von Saito et al. zu sehen ist, können DNA, RNA und weitere Proteine so gut erhalten werden (siehe Abb. 7) [92]. Ein weiterer Vorteil einer Lagerung in RNAlater® liegt auch darin, dass die Proben bis zu 2 Wochen auch bei hohen Temperaturen (28-32 °C) erhalten und in gekühltem Zustand (4 °C) mindestens 6 Monate gelagert werden können [107].

Ethanol: 70% (v/v) Ethanol ist schon seit vielen Jahren ein viel genutztes Konservierungsmedium für Parasiten. Es ist günstig in der Beschaffung und Proben können auch über einen längeren Zeitraum bei Raumtemperatur gelagert werden. Es wurde bereits in Studien mit unterschiedlichen Parasitenproben eine gute Konservierung von Proteinen in 70% (v/v) Ethanol nachgewiesen. Bei der Konservierung von *Teania saginata* und anschließender Untersuchung mit MALDI-TOF MS konnte eine sehr gute Erhaltung der Proteine und damit verbunden eine hoch qualitative Proteinspektren nachgewiesen werden [108]. Auch für *Trichnella* spp. Larven konnte nach längeren Lagerungszeiten hoch qualitative Proteinspektren (LSV-Scores 2.3–2.6) mit MALDI-TOF MS gemessen werden, allerdings wurde eine erhöhte Anzahl an Larven, im Vergleich zur Messung direkt nach der Probenentnahme, zur Proteinextraktion benötigt um eine ausreichende Proteinkonzentration zu erhalten [53].

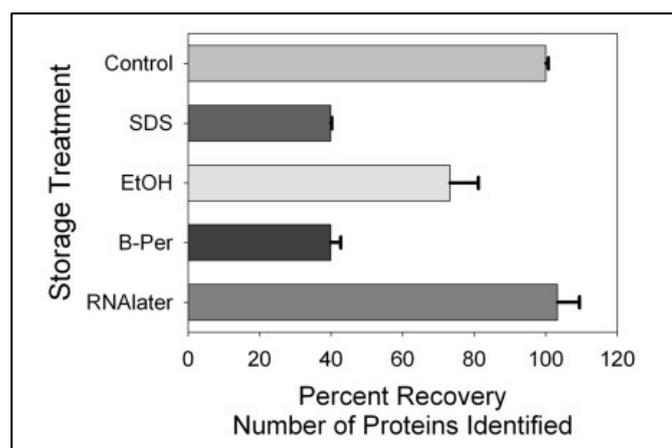


Abbildung 7: Vergleich der Anzahl der Proteinidentifizierungen vier Konservierungstechniken. trocken gefroren bei -80°C gelagert (Control), SDS-extraction buffer (SDS), 90% Ethanol (EtOH), B-PER extraction reagent (B-Per), RNAlater®: Die mit RNAlater® konservierte Probe wies die höchste Anzahl von Protein-Identifizierungen [92]

5.2.3. Proteinextraktion

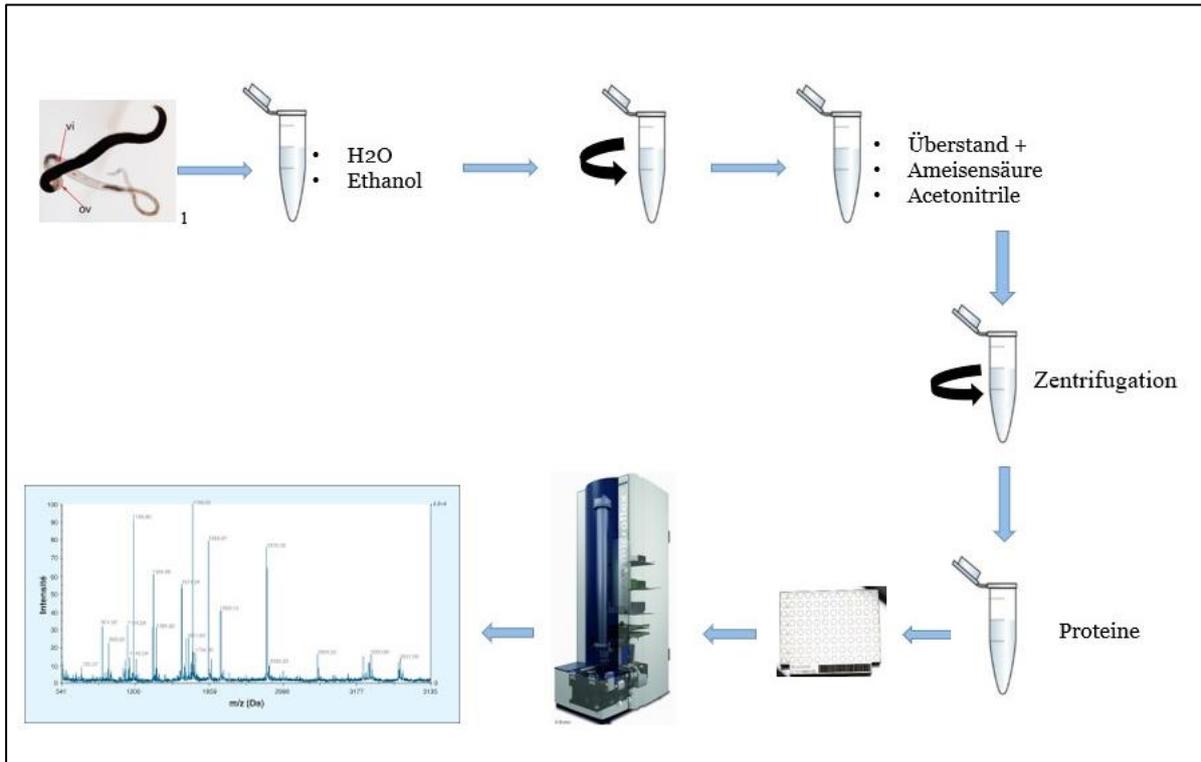


Abbildung 8: Protokoll Proteinextraktion

Adulte Fasciola-Würmer

Die Proben waren bei $-20\text{ }^{\circ}\text{C}$ eingefroren, zur Proteinextraktion wurden diese bei Raumtemperatur aufgetaut. Die adulten Würmer wurden auf einer Petrischale getrocknet und dann mit einem sterilen Skalpell am anterioren Ende ein etwa 0.5 cm großes Stück abgeschnitten (ca. 15 mg) und weiter in kleinere Stücke zerteilt. Die Stücke des anterioren Endes habe ich jeweils in ein 1,5 ml Eppendorf Gefäß gegeben (siehe Abb 6; Schritt 1.1.). Das weitere Verfahren orientierte sich an bereits angewendeten Protokollen zur Helminthendiagnostik [108][99]. In das Eppendorf Gefäß wurden zuerst 300 microliter (μl) Wasser in LC – MS - Qualität (Merck KG; Darmstadt, Deutschland) und 900 μl 100% (v/v) Ethanol absolut (Merck KG) zu der *Fasciola*-Probe hinzugefügt. Diese Mischung wurde für 30 Sekunden (sek) gevortext und anschließend für 2 Minuten (min) bei 1300 rpm (rounds per minute) zentrifugiert. Anschließend konnte die überschüssige Flüssigkeit abpipettiert und verworfen werden. Nach diesem Waschschrift wurde der Bodensatz durch 10-sekündiges vortexen mit 50 μl 70% iger Ameisensäure und 50 μl Acetonitril gelöst und anschließend bei 1300 rpm für 2 min zentrifugiert. Die gelösten Proteine befanden sich nun in der gewonnenen Lösung welche als klarer Überstand im Eppendorf Gefäß zu erkennen war. Diese wurde zur weiteren massenspektrometrischen Messung verwendet (siehe Abb. 8).

F. gigantica-Eier

Die Eier der *Fasciola* wurden durch Auspressen der Eier aus dem Uterus der weiblichen adulten Würmer gewonnen. Hierbei orientierte ich mich an einem Protokoll von Domingo et al.[29]. Dazu wurde der anteriore Anteil des Wurmes mit einem sterilen Skalpell abgetrennt, die Schnitfführung verlief durch den Uterus, welcher makroskopisch als dunkleres Areal am anterioren Ende des Wurmes zu erkennen ist (siehe Abb. 9, Pfeil). Der posteriore Anteil wurde auf eine Petrischale gelegt und durch Druck mit einer Glasscheibe 1,5 cm distal des abgetrennten Endes, die Eier aus dem Uterus gedrückt. Diesen wurde in einem Eppendorf Gefäß 0,5 ml steriles Wasser hinzugefügt und für 10 min bei 12.000 rpm zentrifugiert. Nach abpipetieren des flüssigen Überstandes wurde der Waschschrift nochmals wiederholt. In der Lösung haben wir Konzentrationen von ca. 1000 bis 40.000 Eiern pro ml feststellen können. Zur Konzentrationsbestimmung wurde 1 µl des flüssigen Überstandes mikroskopisch nach Eiern durchsucht und diese gezählt (siehe Abb. 10). In allen Proben haben sich Eier finden lassen. Zur Proteinextraktion wurde der wässrige Überstand abermals abpipettiert, sodass die Eier als trockener Bodensatz im Eppendorf Gefäß zurückblieben. Im Anschluss konnten die so gereinigten Eier nach dem gleichen Protokoll wie bei erwachsenen Würmern verwendet werden, mit dem Unterschied, dass auf Grund der geringeren Probenmenge nur jeweils 10-20 µl Acetonitril und 70%-iger Ameisensäure im letzten Schritt zur Proteinextraktion hinzugegeben werden [60].

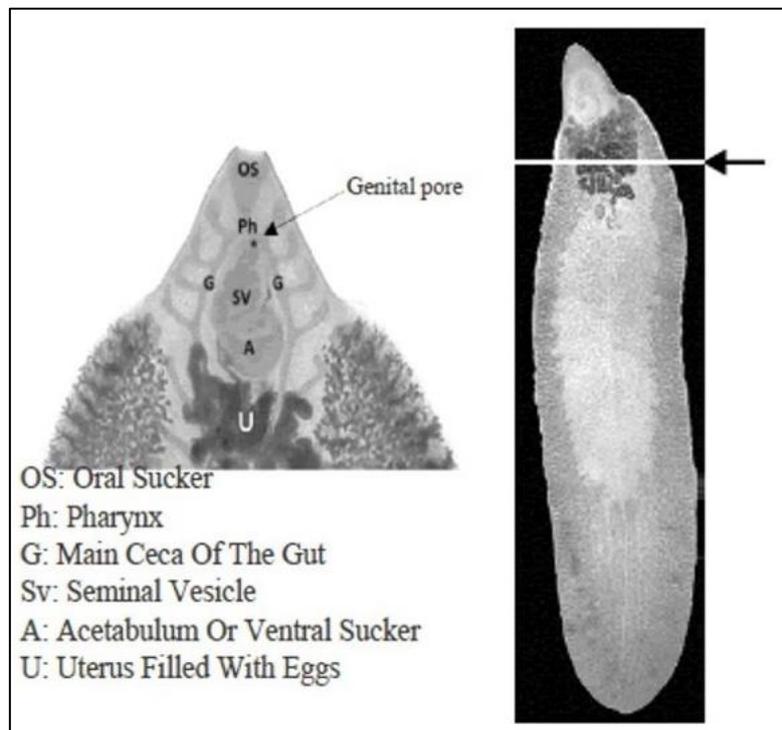


Abbildung 9: *F. gigantica*. Pfeil: Schnitfführung durch Uterus mit darin enthaltenen Eiern. OS: orale Sauger, Ph: Pharynx, G: Teile des Verdauungstraktes, SV: Samenblase, A: Acetabulum, U: Uterus gefüllt mit Eiern [29]



Abbildung 10: *F. gigantica*-Eier

Schistosoma

In den Versuchen wurden sowohl männlich, weibliche, als auch gepaarte (männlich und weiblich) Proben verwendet. Für jede Proteinextraktion waren 2 adulte Würmer nötig, um eine ausreichende Proteinkonzentration in der Lösung zu erhalten. Auf Grund der Größe (9 mm bis 16 mm) der adulten *Schistosoma*, haben wir den gesamten Körper zur Proteinextraktion verwendet. Die anschließenden Schritte folgten dem gleichen Protokoll wie bei der Proteinextraktion der *F. gigantica*-Eier.

5.3. *Massenspektrometrische Messung mittels MALDI-TOF MS*

5.3.1. *MALDI Target Plate Vorbereitung und Messung*

Beim MALDI Target handelt es sich um eine Metallplatte, hergestellt von Bruker Daltronic. Alle Messungen wurden auf einer *MSP 96 Polished Steel MALDI Target Plate* durchgeführt, mit 96 Vertiefungen, welche die einzelnen Proben voneinander trennen. In jede dieser Vertiefungen wurde 1 µl der gewonnenen Proteinlösung pipetiert. Danach musste eine Trockenzeit von 5 min bei Raumtemperatur abgewartet werden um anschließend jede Proteinprobe mit 1 µl Matrix bestehend aus alpha-Cyano-Hydroxy-Zimtsäure Säure in Lösung mit 2,5% Trifluoressigsäure und 50% Acetonitril in Wasser, zu bedecken und abermals eine Trockenzeit von 5 Minuten abzuwarten. Die Herstellung der Matrix orientiert sich an der von Bruker Daltronic empfohlenen Zusammensetzung. Anschließend wurde das MALDI-TOF Target in das Microflex LT Massenspektrometer (Bruker Daltronic) für MALDI-TOF MS gelegt und zunächst die Gerätekalibrierung gestartet. Diese wird standardmäßig mit dem BTS (Bacterial Test Standard) durchgeführt, einem *Escheria coli* DH5 alpha Extrakt, in Verbindung mit zwei hochmolekularen Proteinen, welches ein charakteristisches Protein und Peptid Profil im MALDI-TOF Massenspektroskopie zeigt [119]. Für jede Probe wurden, in einer von Bruker empfohlenen Reihenfolge, drei bis acht Wiederholungen (3 beziehungsweise (bzw). 8 Spots pro Isolat) je nach Verwendungszweck der Probe, zusätzlich zu einer BTS-Kalibrierungsspot auf die *MSP 96 Polished Steel MALDI Target Plate* aufgebracht (Siehe Abb. 11).

Im nächsten Schritt erfolgte eine automatische Gerätekalibrierung, welche sich an den Empfehlungen von Bruker Daltronic GmbH orientierte. Diese wurde mit einer BTS-Probe durchgeführt. Im AutoXecute-Modus wurden folgende Parameter zur Kalibrierung verwendet: „*Min. Number of assigned Peaks*“ (Mindestanzahl der zugewiesenen Peaks): 7; „*Max Peak Error*“ (Maximaler Peak-Fehler): 300 ppm; „*Peak Assignment Tolerance*“ (Toleranz bei der Peak Zuordnung): 1000 ppm [26].

Für die Erstellung von Referenzspektren war die Generierung von mindestens 32 Spektren nötig, also die Messung von 8 Spots pro Isolat. Für alle weitere Proben war eine Messung von lediglich 3 Spots pro Isolat ausreichend. Nach der Kalibrierung konnte die Messung der Proben gestartet werden (siehe Abb. 6, Schritt 1.2.).

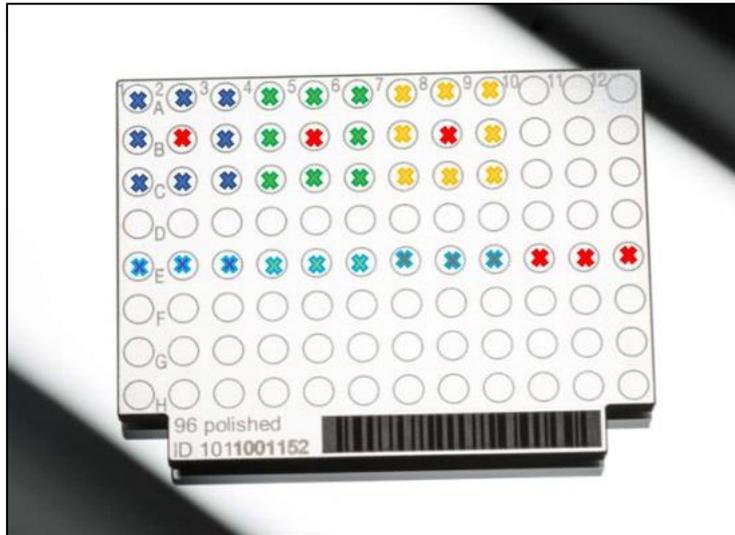


Abbildung 11: MSP 96 Polished Steel MALDI Target Plate: BTS (rot), Proben mit 8 Spots pro Isolat (blau, grün, gelb), Proben mit 3 Spots pro Isolat (mehrfarbige Kreuze)

5.3.2. MALDI-TOF MS Parameter

Die MALDI-TOF MS Spektren-Generierung wurde mit einem Microflex LT Massenspektrometer durchgeführt, welches von der FlexControl[®] Software Version 3.4 (Bruker Daltonics, Bremen, Deutschland) mit Detektion des Masse-Ladungs-Verhältnis von positiven Ionen im Bereich von 3000-20.000 Da (Dalton) kontrolliert wurde. Im Autoexecute Modus wurden pro Spot 240 Laserschüsse, je 40 auf 6 zufällige Regionen abgegeben und im Mittel ein Spektrum generiert. Pro Spot habe ich 4 Messungen durchgeführt und somit 12 beziehungsweise (bzw.) 32 Spektren pro Isolat erhalten. Die Laserfrequenz lag bei ca. 60 Hz, die Hochspannung bei 20 kV und die gepulste Ionenextraktion bei 180 Nanosekunden (ns). Die Intensität der Laserschüsse wurde bei 20 – 26 % gewählt.

5.4. Datenanalyse mit MALDI-TOF Biotyper

5.4.1. Spektrenaufbereitung und Erstellen einer haus-internen Datenbank

Die in den Messungen generierten 12 bzw. 32 Spektren pro Isolat werden als Rohspektren („Raw Spectra“) bezeichnet, da sie auf Grund der zufällig ausgewählten Messregionen teilweise *flatlines* oder *outlines* enthalten, also Messungen, an denen keine oder verfälschte Proteinspektren aufgezeichnet wurden.

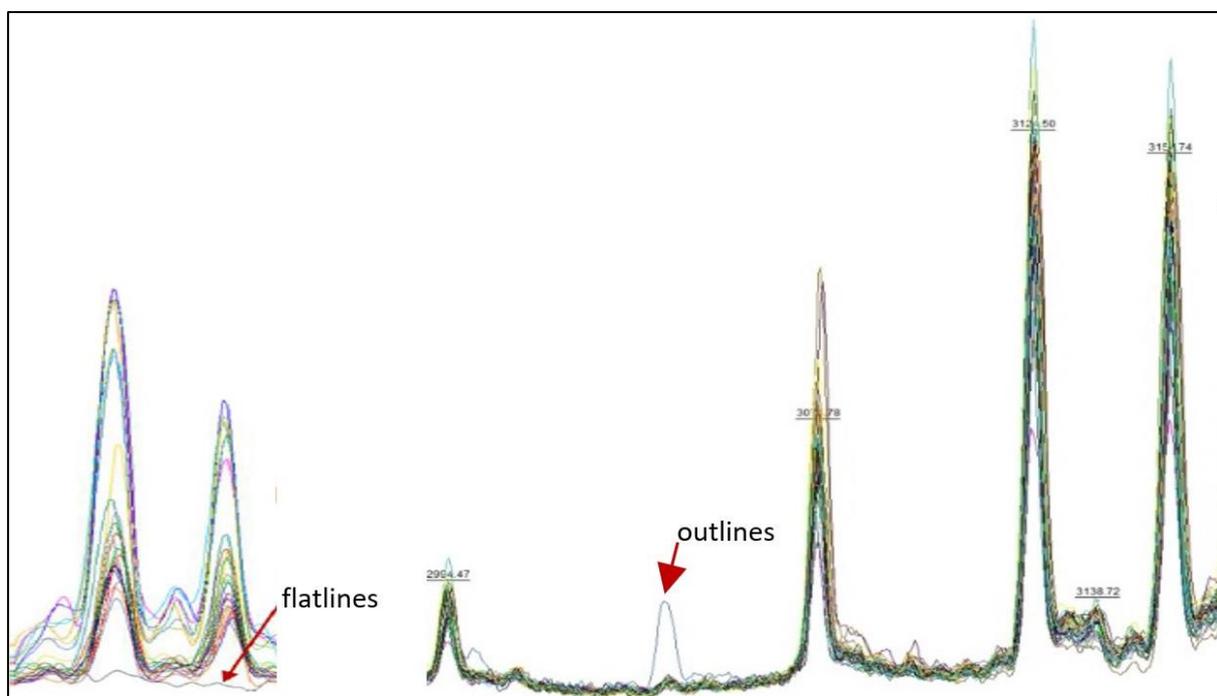


Abbildung 12: Spektrenaufbereitung: links: Rohspektrum mit *flatline* (keine Generierung von Protein Peaks) rechts: Rohspektrum mit *outline* (Peak Spektrum, welches in der Gesamtübersicht aller Spektren aus der Norm fällt)

Zur Erstellung von Referenzspektren (Main Spectrum Profil: MSP), welche zur Anlegung der Datenbank verwendet werden, können ausschließlich gereinigte Spektren verwendet werden, um eine Verfälschung der Datenbank durch Hintergrundgeräusche, z.B. der Matrix, zu verhindern und die Auflösung zu erhöhen. Die Reinigungsschritte werden mit der FlexAnalysis® software version 3.4 (Bruker Daltonics, Bremen, Germany) durchgeführt und umfassen eine Löschung von *flat-* und *outlines* (siehe Abb.12), eine Glättung der Intensitäten die Subtraktion von Basislinien und eine Festlegung der Peakverschiebung auf 500 ppm (parts per million). *Outlines* konnten im Vergleich der 32 Proteinspektren der einzelnen Proben detektiert werden. Mittels

einer automatisierten Excel Tabelle, wurden repräsentative Spektren ausgewählt, welche anschließend als MSP zusammengefasst wurden. Jedes genutzte Proteinspektrum musste mindestens über 8 Proteinpeaks im Bereich von 3 000 – 18 000 Da verfügen, der *maximum peak error* wurde mit < 300 ppm angegeben und die Differenz zwischen *before last fit* und *in last fit* musste < 300 ppm betragen. Eine Mindestanzahl von 27 Spektren pro MSP wurde von allen verwendeten Proben erreicht. Für die in-house-Datenbank wurden jeweils Referenzproben der verschiedenen Gattungen bzw. Spezies gereinigt und ausschließlich hochwertige MSPs für jede Probe ausgewählt. In einem MSP sind sowohl die Intensitäten als auch die deren Häufigkeitsverteilung, sozusagen als „Fingerabdruck“ der Probe verzeichnet. Zur Identifizierung einer unbekannt Probe werden die Peaks dieser Probe mit denen des MSPs verglichen. In der Abb. 13 kann man unter exakter Übereinstimmung (grün), annähernder Übereinstimmung (gelb) und fehlender Übereinstimmung (rot) mit dem MSP (blau) unterscheiden. Auf Grundlage dieses Vergleiches kann eine Identifikation erfolgen. Die Weiterverarbeitung und der Einschluss der neuen MSPs in die Datenbank wurde mit der MALDI Biotyper Compass Explorer® software version 3 (Bruker Daltonics) durchgeführt. Die in-house-Datenbank des Instituts für Medizinische Mikrobiologie und Hygiene in Homburg enthält bereits einige MSPs von Zestoden, Trematoden und Nematoden [36][108] und wurde nun im Bereich der Trematoden um MSPs von *Schistosoma* und *Fasciola* erweitert.

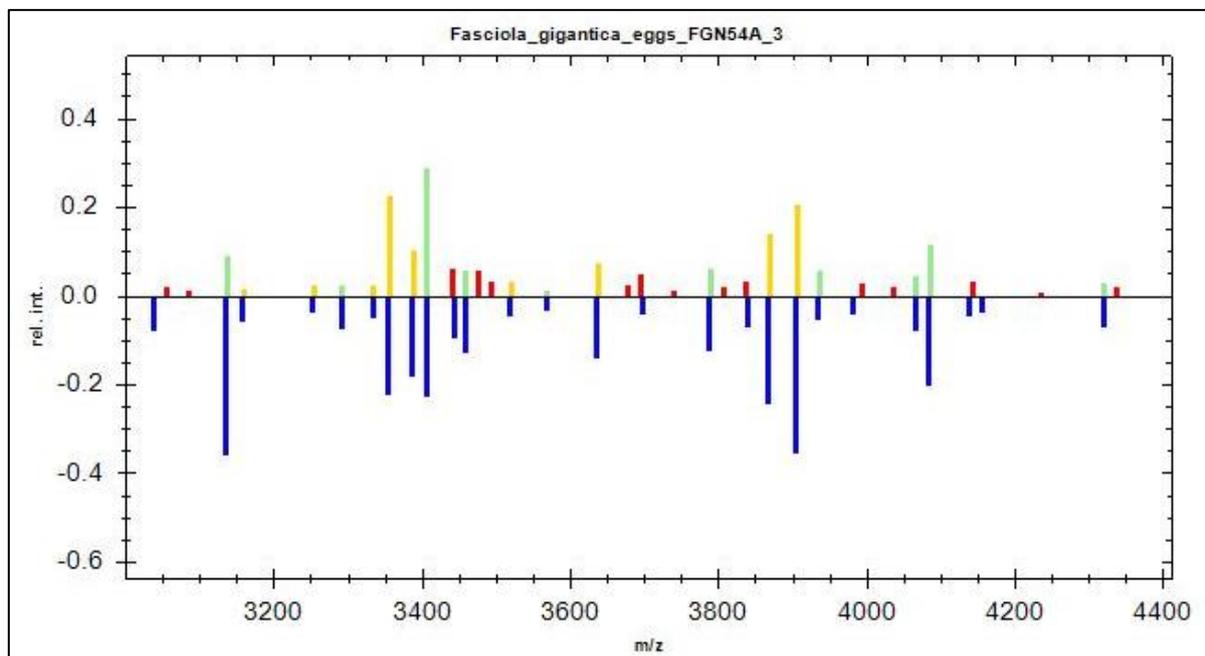


Abbildung 13: Beispielhaft eine Probe von *F. gigante*-Eiern. Abgleich des Proteinspektrums der Probe (obere Balken) mit MSP für *F. gigante*-Eier (untere Balken - blau): exakte Übereinstimmung (grün), annähernde Übereinstimmung (gelb) und fehlende Übereinstimmung (rot)

5.4.2. *Auswertung der Messergebnisse*

Der LSV, welcher mit der MALDI Biotyper Compass Explorer® Software Version 3 (Bruker Daltoniks) berechnet wurde, dient zur Einschätzung der Reliabilität der Identifizierung, auf der Grundlage der Vergleichbarkeit zwischen dem Referenz-MSP und den neu generierten Spektren der Test-Probe (siehe Abb.6, Schritt 2.1.). Der Cut-off für die LSV wurde auf der Grundlage der molekularen Identifizierung anhand einer Rezeptor-gesteuerten Kurve (ROC-Kurve) bestimmt, die durch logistische Regression (SAS 9.4, Grégy-sur-Yerres, Frankreich) berechnet wurde. Gemäß den Empfehlungen von Bruker Daltronic für die Identifikation von Bakterien wurden LSVs unter 1,7 als ungültige Identifikation angesehen. LSVs zwischen 1,7 und 2,0 wurden auf Gattungsebene als gültig angesehen (also z.B. *Schistosoma* spp.), und LSVs höher als 2,0 wurden als sichere Identifizierung der Art/Spezies gewertet (also z.B. *S.mansoni*) [49].

5.4.3. *Validierung der Datenbank*

Alle gemessenen Proben wurden gegen die kommerzielle BDAL-Datenbank von Bruker Daltronic getestet, um festzustellen, ob ein Spektrum mit bakteriellen Spektren übereinstimmt, also Verunreinigungen in der Probe durch Bakterien oder Pilze nachzuweisen sind (siehe Abb. 6, Schritt 1.3.).

Die neue in-house-Datenbase mit den enthaltenen Referenzspektren der hinzugefügten *Schistosoma*- und *Fasciola*-Stämme wurde zwei Validierungsschritten unterzogen. Zu Beginn wurde eine Interne Validierung durchgeführt, bei der die Rohspektren, welche zur MSP-Erstellung verwendet wurden, durch die in-house-Datenbase getestet wurden. Die Identifikation der Rohspektren fand somit durch ein MSP des identischen Isolats statt (siehe Abb. 6, Schritt 2.2.).

Der zweite Schritt umfasste die Identifikation neu gewonnener Proben durch die in-house-Datenbase. Dies ist der zentrale Schritt zu Validierung der Datenbank, da bei einer zuverlässigen Identifikation, diese auch auf unbekannte, unverarbeitete Proben übertragbar ist. Diese externe Validierung habe ich mit einer Kombination aus der BDAL-Datenbank und der in-house-Datenbase durchgeführt. Im Gegensatz zu den Rohspektren, welche zur Erstellung der MSPs verwendet wurden, sind für die externe Validierung lediglich die Messung von 3 Spots pro Isolat nötig (siehe Abb. 6, Schritt 2.3.).

Für *F. gigantica* enthielt die MALDI TOF MS-Datenbank am Institut für Medizinische Mikrobiologie und Hygiene bereits Referenzspektren [99]. Die bereits vorhanden MSPs habe ich zur

speziesspezifischen Identifikation und der Analyse des Einflusses der Lagerungsdauer und des verwendeten Anteils des adulten Wurmes verwendet.

Neue MSPs habe ich für *F. gigantica*-Eier, adulte *F. hepatica*, *F. hepatica*-Eier und *S. mansoni* und *S. japonicum* erstellt.

5.5. Klassifizierungs- und Vergleichsanalyse mittels „Clover“-Software

5.5.1. Parameter zur Vorverarbeitung

Zur weiteren Untersuchung unser Daten und der maschinellen Lernprozesse wurde eine kostenlose Online-Software, die Clover MS Data Analysis® Software (<https://platform.cloverms-dataanalysis.com/>, Clover BioSoft, Granada, Spanien) verwendet. Zur Vorverarbeitung der Spektren wurden Standardparameter verwendet, welche bereits in anderen Arbeiten mit MALDI-TOF MS angewendet wurden [18]. Zuerst mussten alle *flatlines* entfernt werden, um eine Verfälschung der Ergebnisse zu verhindern. Zur Glättung der Spektren wurde der Savitzky-Golay-Filter (Fensterlänge 11; Polynomordnung 3) angewendet, die Basislinie wurde mit der Top-Hat-Filtermethode (Faktor 0,02) entfernt. Um ein Durchschnittsspektrum pro Probe zu erstellen, das für die Klassifizierungs- und Vergleichsanalyse verwendet werden konnte, wurden die geglätteten Spektren nach folgenden Parametern ausgerichtet: zulässige Verschiebung, mittel; konstante Toleranz, 0,2 Da; lineare Toleranz, 2000 ppm (siehe Abb. 6, Schritt 3.1.).

Für die Analyse der *Schistosoma* konnten insgesamt 1657 bearbeitete Spektren zur weiteren Untersuchung verwendet werden. Für die Analyse der adulten *Fasciola* waren es 215 der *F. hepatica* Eier, 75 Spektren für die adulten *F. hepatica*-Proben. Es wurden weiterhin 287 Spektren zur Analyse der *F. gigantica*-Eier und 30 Spektren der adulten *F. gigantica*-Proben verwendet.

5.5.2. Klassifizierung und Algorithmen des maschinellen Lernens

Zur Vergleichsanalyse wurde eine Peak-Matrix verwendet, also eine Vorlage zur Analyse. Hierfür mussten für die vorverarbeiteten Spektren ein Peak-Matching, also ein Abgleich der

zugehörigen Peaks, durchgeführt werden, welches folgende Schritte umfasst: die Normalisierung des Gesamtionenstroms (TIC), gefolgt von einer „Schwellenwertmethode“ (Faktor 0.01), wodurch Peaks mit einer Intensität unter 1% der maximalen Intensität nicht berücksichtigt werden, eine konstante Toleranz von 0.5 Da und einer linearen Toleranz von 500 ppm [18].

ML ist ein Teilbereich der künstlichen Intelligenz, mit dessen Hilfe IT-Systeme in der Lage sind, auf der Grundlage vorhandener Daten und Algorithmen Muster und Regelmäßigkeiten zu erkennen und Lösungen zu entwickeln, wobei also künstliches Wissen aus Erfahrung generiert wird. Das aus den Daten gewonnene Wissen kann verallgemeinert und für neue Problemlösungen oder für die Analyse von bisher unbekanntem Daten genutzt werden. ML verwendet für seine Lernprozesse verschiedene Lernalgorithmen (LA), welche in zwei Hauptgruppen eingeteilt werden: unüberwachte und überwachte Lernalgorithmen. Der Hauptunterschied zwischen überwachten und unüberwachten LA besteht darin, dass überwachte LA Kategorien einen Trainingsdatensatz (Peak-Matrix) benötigen, im Gegensatz zu unüberwachten LA, bei denen das Ziel darin besteht, Datenpunkte in diskrete Gruppen einzuteilen und somit die „wahre Natur“ dieser Verteilung zu erfahren. So wird dabei innerhalb eines Datensatzes immer jeweils eine Probe als Peak Matrix verwendet und alle weiteren Proben als unbekannter Testdatensatz mit dieser Peak Matrix verglichen. Das Modell sucht selbständig nach Beziehungen zwischen Variablen und sagt diese Beziehungen dann anhand von nicht kategorisierten Testdaten voraus. Mit allen eingespeisten Proben des Datensatzes wird auf die gleiche Weise verfahren. Man erhält somit ein sehr viel detailreicheres Bild der Unterschiede innerhalb der verwendeten Daten. Beim überwachten Lernen enthalten die Trainingsdaten Kategorien, die das interessierende Ergebnis bezeichnen (z.B. die unterschiedlichen Lagerungslösungen für *S. mansoni* Proben). Das Modell wird auf diesen Daten trainiert und sagt dann die vordefinierten Ergebnisse von Interesse auf Testdaten voraus. Die Anwendung überwachter LA dient auch der genaueren Untersuchung von Unterschieden innerhalb eines Datensatzes. Für überwachte LA ist es möglich, eine externe Validierung durch neue unbekannte Proben durchzuführen. Die Peak Matrix setzt sich dabei aus allen Proben des ursprünglich verwendeten Datensatzes zusammen und wird durch weitere nicht klassifizierte Proben getestet. Abbildung 14 gibt eine Übersicht über die unterschiedlichen LA-Formen.

Zur Klassifizierung der verschiedenen Isolate wurden sowohl unüberwachte (Hauptkomponentenanalyse/principal component analysis, PCA) und überwachte (lineare Supervektormaschine (SVM), Partial Least Squares-Discriminant Analysis (PLS-DA), Random Forest (RF) und K-nearest neighbors (KNN)) maschinelle LA verwendet, um eine Überlegenheit eines bestimmten LA herausfinden zu können.

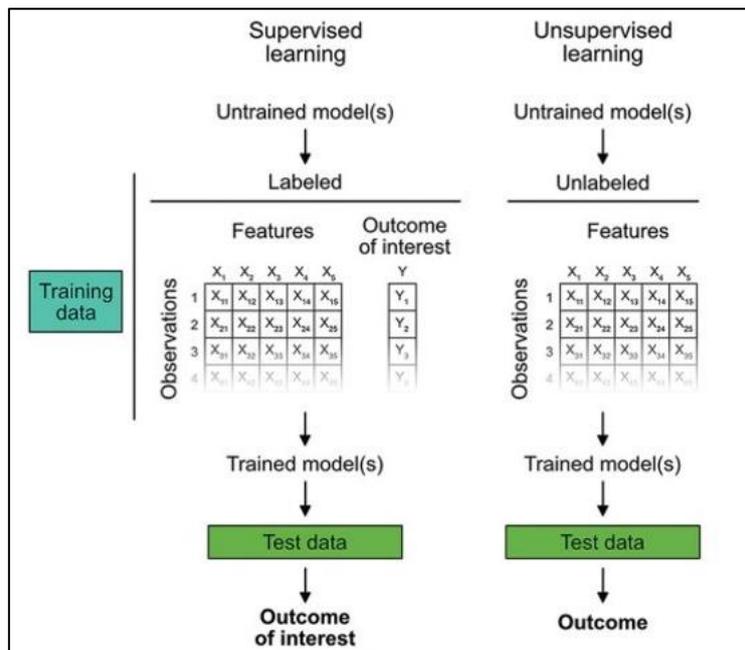


Abbildung 14: Unterschied überwachtes vs. unbeaufsichtigtes Lernen: Beim überwachten Lernen („supervised learning“) werden bestimmte Kategorien zur Einordnung vorbestimmt („Labeled“). Beim unbeaufsichtigten Lernen („unsupervised learning“) werden keine Kategorien zur Einordnung vorgegeben, die natürliche Datenverteilung wird dadurch untersucht. [6]

PCA ist ein Algorithmus zur Dimensionalitätsreduktion, was bedeutet, dass der hochdimensionale Datensatz auf eine Reihe von Koordinaten reduziert wird, indem Variablen entfernt werden, welche mit anderen korreliert sind, aber keinen großen Informationsverlust darstellen. So wird eine Strukturierung und Vereinfachung eines großen Datensatzes erreicht, die eine bessere Visualisierung der verschiedenen Cluster und Beziehungen zwischen den Proben ermöglichen und die Identifizierung von Untergruppen ermöglicht. Zusätzlich dient die PCA zur Erkennung von Ausreißern [40].

SVM kann die Objektklassen mit Hilfe von Trennebenen unterteilen. Diese Ebenen werden so gewählt, dass zwischen verschiedenen Klassen ein möglichst großer objektfreier Bereich verbleibt. Die Trennebene mit dem größten objektfreien Bereich wird als optimale Lösung angesehen. Der Vorteil dieses LA ist, dass sowohl lineare als auch nichtlineare Klassifizierung unterstützt werden.

PLS-DA detektiert Strukturen in Datensätzen und minimiert die Anzahl an deckungsgleichen Daten. Dieser LA ist robuster gegenüber Messunsicherheiten als vergleichbare LA, da Messunsicherheiten bereits in den Berechnungen eingeschlossen sind.

Die Basis für RF bilden viele einzelne Entscheidungsbäume. Ausgehend von der ersten Entscheidung werden immer mehr Verzweigungen gebildet, bis ein bestimmtes Ergebnisniveau

erreicht ist. Jeder einzelne Entscheidungsbaum in einem Random Forest erhebt nicht den Anspruch, jedem Datensatz die richtige Klasse zuzuordnen. Jeder Entscheidungsbaum liefert als kleines, unabhängiges Modell Input für die Gesamtentscheidung. Die zufällige Varianz der Bäume erhöht die Ergebnis- und Vorhersagegenauigkeit des RF-Algorithmus.

KNN ist eine der einfachsten Methoden zur Klassifizierung großer Datenmengen. Hierfür werden die Abstände zwischen den neuen Daten und den bereits klassifizierten Daten des kategorisierten Trainingsdatensatz zur Klassifizierung verwendet. So werden neue Punkte der Gruppe der ihnen am nächsten gelegenen Trainingsdatensätzen (Nachbarn) zugeordnet. Hierfür wird eine ganze Zahl (k) gewählt welche die Anzahl der entscheidenden nächsten Nachbarn festlegt. Der Euklidische Abstand ($k=2$) ist das am häufigsten verwendete Abstandsmaß, wurde in unseren Berechnungen verwendet.

Für die interne Validierung wurde die Methode der K -fachen Kreuzvalidierung („ k “ gibt dabei die Anzahl der Gruppen, in die eine bestimmte Datenprobe zur Validierung aufgeteilt werden soll an) verwendet, eine Wiederholungsauswahlmethode, welche die Leistung des ML-Modells schätzt [40]. Als Leistungsindikatoren der überwachten ML-Algorithmen wurde eine Konfusionsmatrix (die Werte wie Genauigkeit, Spezifität, Sensitivität, F1-Score, positiver Vorhersagewert (PPV) oder Präzision und negativer Vorhersagewert (NPV) generiert) sowie die Fläche unter der Receiver-Operating-Characteristic-Kurve (AUROC-Kurve) und die Fläche unter der Precision-Recall-Kurve (AUPR-Kurve) verwendet [83].

Der AUROC-Wert liegt immer zwischen 0 und 1 und ist ein Maß für die Klassifizierungsleistung des LA, hohe Werte lassen auf eine gute Leistung schließen (1 wäre eine perfekte Leistung), niedrigere Werte eher auf eine schlechtere Leistung.

Die Receiver-Operating-Characteristic-Kurve (ROC) wird in einem Diagramm aus dem negativ prädiktiven Wert (NPV) (x-Achse) und dem positiv prädiktiven Wert (PPV) (y-Achse) dargestellt. Aus der Kurve lassen sich die Genauigkeit und fast alle weiteren genannten Werte ableiten.

Die Sensitivität gibt an, wie hoch die Wahrscheinlichkeit ist, dass bei einem positiven Ergebnis dieses auch tatsächlich korrekt ist, beispielsweise eine Krankheit oder ein Erregernachweis durch den verwendeten LA gelingt. Die Spezifität wiederum beschreibt, mit welcher Genauigkeit ein negatives Ergebnis, also ein gesunder Mensch oder nicht vorhandener Erreger erkannt wird.

Der PPV lässt darauf schließen, wie verlässlich ein positives Ergebnis ist, also wie hoch die Wahrscheinlichkeit ist, dass bei einem positiven Ergebnis dieses auch tatsächlich korrekt ist.

Dies dient der Einschätzung der Aussagekraft eines Testverfahrens. Im Gegensatz dazu steht der NPV für die Vertrauenswürdigkeit eines negativen Ergebnisses.

Der F1 Score lässt sich aus dem geometrischen Mittel, also dem Mittelwert, den man mithilfe der n-ten Wurzel aus dem Produkt der betrachteten n positiven Zahlen erhält, aus Sensitivität und Präzision berechnet [40] [37]. Er gibt dabei Aufschluss über die ausgewogene Fähigkeit des Modells, sowohl positive Fälle zu erfassen (Sensitivität) als auch bei den Fällen, die es erfasst, genau zu sein (Präzision). Damit ist er einer der robustesten Gradmesser zur Einschätzung einer Modelleistung.

Fasciola gigantica und Fasciola hepatica

Interspezies Klassifizierung

In der Klassifizierungsanalyse von *Fasciola* wurde eine speziesspezifische Differenzierung mittels unüberwachter und überwachter LA dargestellt (siehe Abb. 6, Schritt 3.2.).

Intraspezies Klassifizierung – Entwicklungsstadium und verwendeter Körperanteil

Zur genaueren Untersuchung der entwicklungspezifischen Unterschiede der Proteinspektren wurden sowohl die Proben von *F. hepatica* als auch *F. gigantica* hinsichtlich der adulten („adults“) bzw. Ei-Proben („eggs“), im unüberwachten als auch mit überwachten LA untersucht (siehe Abb. 6, Schritt 3.3.).

Zur Differenzierung des verwendeten Körperanteils wurden die *F. gigantica*-Proben nochmals in *F. gigantica* anteriorer bzw. posteriorer Anteil aufgeteilt. Für diese Klassifizierung wurde nur ein unüberwachter LA (PCA) angewendet. Auf die Anwendung weiterer LA wurde an dieser Stelle verzichtet.

Schistoma mansoni und Schistosoma japonicum

Interspezies Klassifizierung

Die Klassifizierungsanalyse für *Schistosoma* wurde auf zwei Ebenen durchgeführt (siehe Abb. 6, Schritt 3.2.). Zunächst wurde die Interspezies-Ebene untersucht, also die Diskrimination aller Isolate zwischen *S. mansoni* und *S. japonicum*. Zur Anwendung kamen unüberwachte und überwachte LA.

Intraspezies Klassifizierung - Lagerungslösung

Für die Intraspezies-Klassifizierung wurden Proben derselben Spezies verglichen, um festzustellen, welche LA eine Differenzierung beider Lagerungslösungen darstellen konnten. ob sie in der Lage waren, die Wirkung der Lagerungslösungen (70% (v/v) Ethanol und RNAlater®) zu unterscheiden. Die Anwendung überwachter LA diente der genaueren Untersuchung von eventuellen Unterschieden und der Bestätigung der durch die PCA (unüberwachter LA) erhaltenen Daten (siehe Abb. 6, Schritt 3.3.).

6. Geräte, Reagenzien, Software

Target Reinigung

80 % Trifluoressigsäure	Merck KG; Darmstadt, Deutschland
70% (v/v) Ethanol	Merck KG; Darmstadt, Deutschland
HPLC – Reinstwasser	Sigma – Aldrich
KimWipe – Präzisionswischtuch	Kimberly – Clark

Matrix Aufarbeitung

2,5mg HCCA portioned	Bruker Daltonics
Acetonitril	Sigma - Aldrich
HPLC – Reinstwasser	Sigma - Aldrich
Trifluoressigsäure	Sigma - Aldrich
Vortexer	Scientific Industries

Lagerungsmedien und Probenaufarbeitung

70% (v/v) Ethanol	
RNAlater®	Ambion, Thermo Fisher

Messung und Auswertung

HCCA - Matrix (aufbereitet)	
Bacterial Test Standart (BTS)	Bruker Daltonics; Bremen, Germany
MALDI-TOF Microflex MS	Bruker Daltonics; Bremen, Germany
MicroScout Target (MSP 96 target polished steel)	Bruker Daltonics; Bremen, Germany
FlexAnalysis (Software)	Bruker Daltonics; Bremen, Germany
MALDI Biotyper Compass Explorer (Software)	Bruker Daltonics; Bremen, Germany
Clover MS Data Analysis® (Software)	Clover BioSoft, Granada, Spain
Excel (Software)	Microsoft
Word (Software)	Microsoft

7. Ergebnisse

7.1. *Fasciola gigantica* und *Fasciola hepatica*

7.1.1. Validierung der MALDI-TOF MS basierten Diagnostik von *F. hepatica* und *F. gigantica*-Eiern

Ausschluss von Kontaminationen und visuelle Spektrumanalyse

Zum Ausschluss von Kontaminationen wurden alle Proben zu Beginn mit der BDAL-Datenbank von Bruker Daltronic verglichen. Bei keiner Probe wurde eine Übereinstimmung mit dieser Datenbank gefunden (siehe Tabelle 1), eine Kontamination mit Pilzen oder Bakterien konnte somit ausgeschlossen werden.

<i>Fasciola</i> -Proben	Anzahl der Proben	Anzahl der Proteinspektren	Korrekte Identifikation	Durchschnitt LSV
<i>F. gigantica</i> anteriorer Anteil	50	284	0%	1.33
<i>F. gigantica</i> -Eier	32	1047	0%	1.33
<i>F. hepatica</i> posteriorer Anteil	10	126	0%	1.34
<i>F. hepatica</i> -Eier	10	251	0%	1.26

Tabelle 1: *Fasciola* spp. vs. BDAL-Datenbank von Bruker Daltronic

Validierung der MALDI-TOF MS basierten Diagnostik von *F. hepatica* und *F. gigantica*-Eiern

In der visuellen Analyse der Proteinspektren konnte optisch eine weitgehende Übereinstimmung der verschiedenen *F. gigantica*-Proben und der respektiven Peaks gefunden werden (siehe Abb. 15).

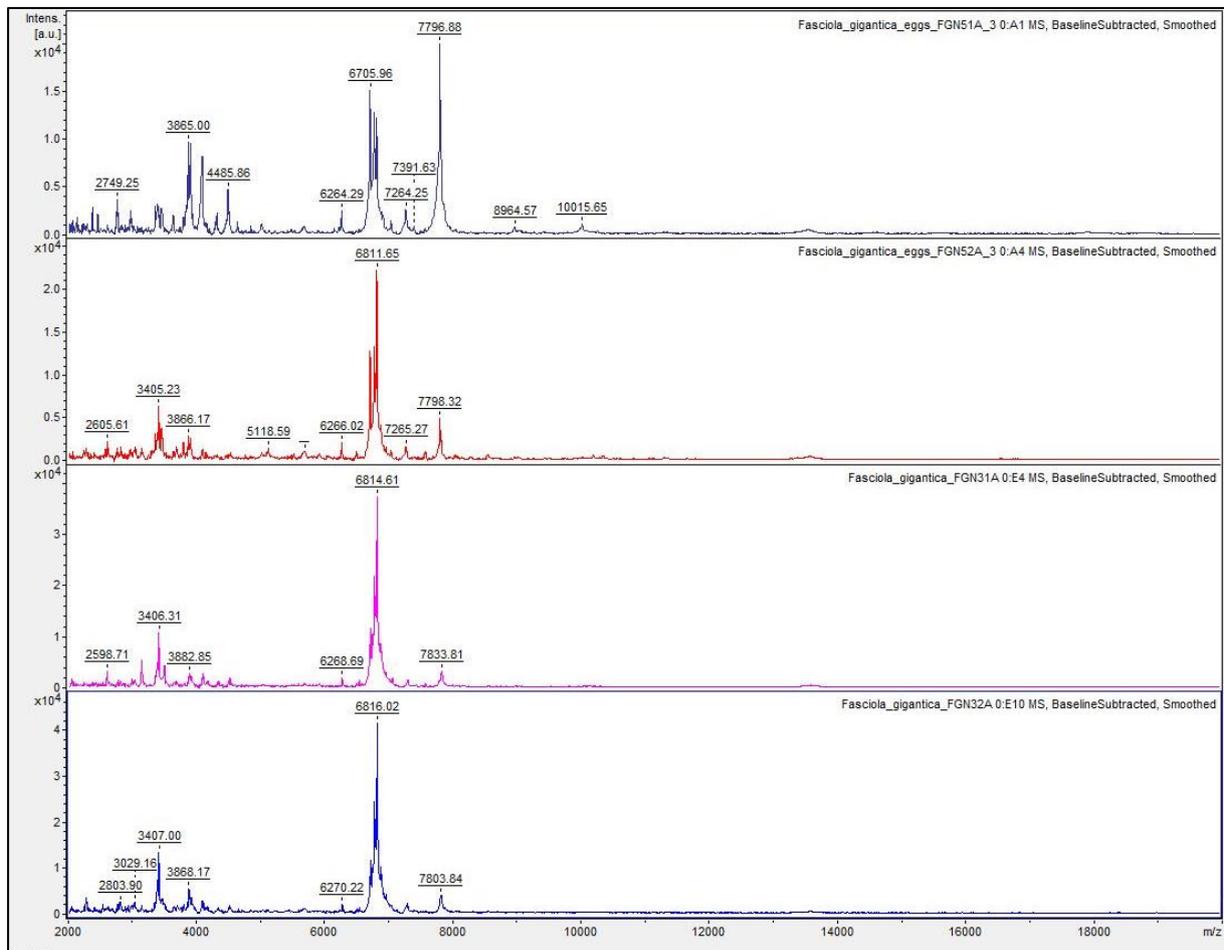


Abbildung 15: Proteinspektren von *F. gigantica* Eiern (dunkel blau und rot) und adulten *F. gigantica*-Proben (pink und Königsblau)

Für alle *Fasciola*-Proben konnten eindeutige Spektren generiert werden, welche alle zur Identifikationsanalyse verwendet wurden. Eine Differenzierung der Entwicklungsstadien von *F. hepatica* (adulte Proben vs. Eier) ist in der visuellen Analyse der Spektren nicht eindeutig zu erkennen (siehe Abb. 16). Im Vergleich zu den Proteinspektren der *F. gigantica*-Proben fällt jedoch auf, dass im Bereich 6200 m/z bis 6800 m/z deutlich mehr Peaks bei *F. hepatica*-Proben zu erkennen sind (siehe Abb. 17).

Validierung der MALDI-TOF MS basierten Diagnostik von *F. hepatica* und *F. gigantica*-Eiern

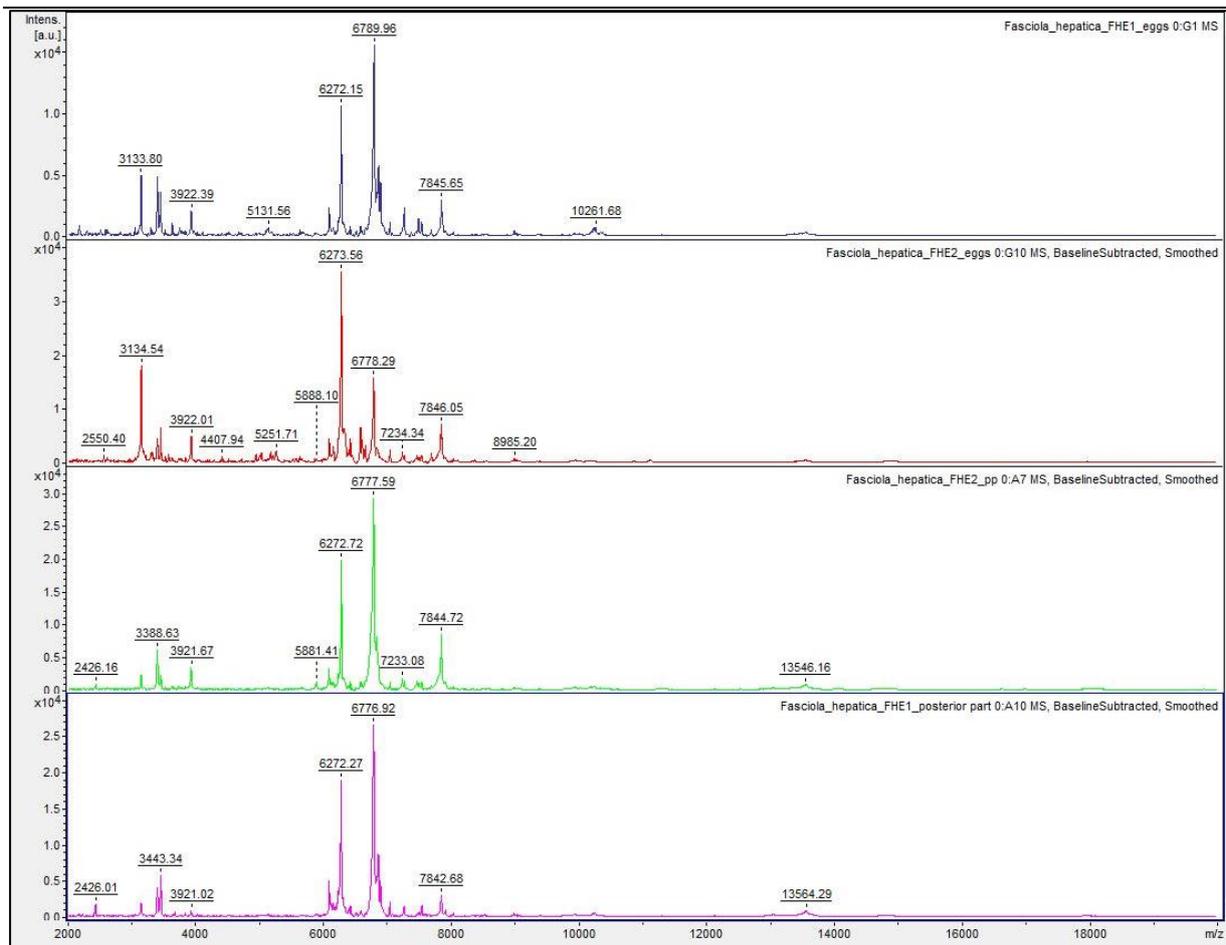


Abbildung 16: Proteinspektren von *F. hepatica*-Eiern (blau und rot) und adulten *F. hepatica*-Proben (grün und lila)

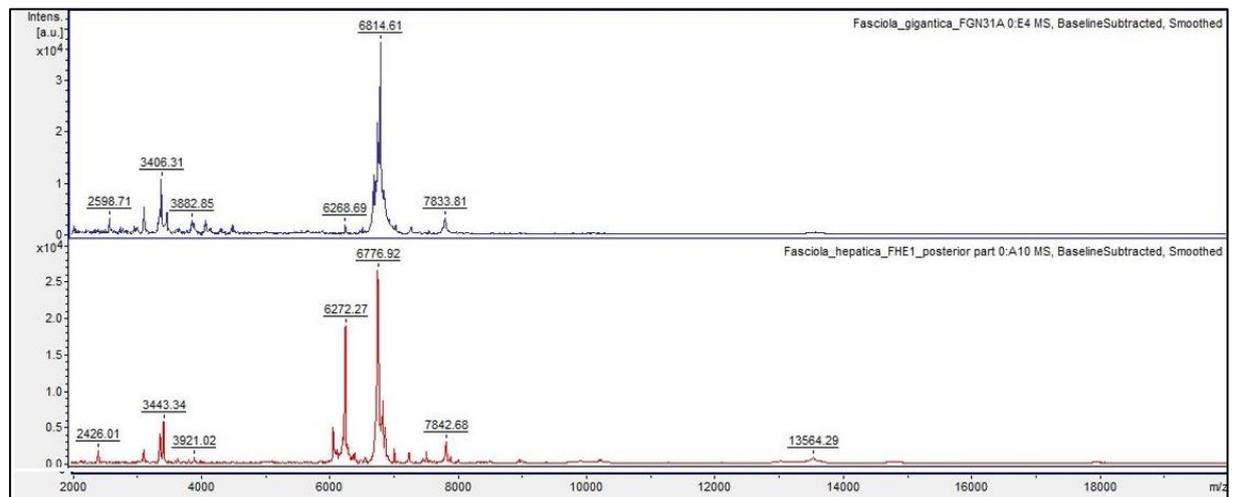


Abbildung 17: Vergleich der Proteinspektren von *F. gigantica* und *F. hepatica* mit der FlexControl® Software Version 3.4.

Validierung der In-house-Database

Alle für die MSP-Erstellung verwendeten Proben wurden einer internen Validierung unterzogen. Dabei ergaben sich ausschließlich Identifizierungen mit einem LSV über 2.0, was einer Identifizierung auf Speziesebene entspricht (siehe Tabelle 2, Abschnitt (A)). Für *F. gigantica*-Eier betrug der durchschnittliche LSV 2.54, die *F. hepatica* (posteriorer Anteil) konnte mit einem LSV von 2.67 erkannt werden und die beiden Proben der *F. hepatica*-Eier mit einem durchschnittlichen LSV von 2.74.

Im nächsten Schritt wurden alle weiteren, bis dahin unbekanntenen Proben mit der zuvor erweiterten Datenbank verglichen. Bei 108/111 (97.3%) aller Proben konnte eine Identifikation der Gattung ($LSV \geq 1.7$) mittels MALDI-TOF MS erreicht werden. Bei 94/111 (85.45%) war es möglich auch die Spezies mit MALDI-TOF MS zu ermitteln ($LSV \geq 2.0$). Die durchschnittlichen LSV-Werte lagen zwischen 2.02 und 2.29 (siehe Tabelle 2, Abschnitt (B)). Auch im LSV-Bereich von 1.7 bis 2.0 gab es keine Fehlklassifizierungen von *F. hepatica* als *F. gigantica* und vice versa.

Validierung der MALDI-TOF MS basierten Diagnostik von *F. hepatica* und *F. gigantica*-Eiern

Interne und externe Validierung der *Fasciola*-Proben:

		(A)	Interne	Validierung	
Spezies	Verwendeter Anteil	Anzahl Proben	der Durchschnitt LSV	LSV \geq 1.7	LSV \geq 2.0
<i>F. gigantica</i>	Anteriorer Anteil	-			
	Eier	10	2.54	100%	100%
<i>F. hepatica</i>	Posteriorer Anteil	1	2.67	100%	100%
	Eier	2	2.74	100%	100%
Total		13			

		(B)	Externe	Validierung	
Spezies	Verwendeter Anteil	Anzahl Proben	Durchschnitt LSV	LSV \geq 1.7	LSV \geq 2.0
<i>F. gigantica</i>	Anteriorer Anteil vs. MSP des posterioren Anteils	50	2.11	48/50 (96%)	42/50 (84%)
	Eier vs. MSP der adulten Proben	22	2.02	22/22 (100%)	15/22 (68.22%)
	Eier vs. MSP der adulten Proben und der Eier	22	2.29	22/22 (100%)	22/22 (100%)
<i>F. hepatica</i>	posteriorer Anteil	9	2.12	9/9 (100%)	8/9 (88.8%)
	Eier vs MSP der adulten Proben und der Eier	8	2.06	7/8 (87.5%)	7/8 (87.5%)
Total		111		108/111(97.3%)	94/111 (85.45%)

Tabelle 2: (A) Interne und (B) Externe Validierung: Identifizierung von *Fasciola*-Proben, in verschiedenen Entwicklungsstadien (adulte Proben, Eier), unter Verwendung erweiterter in-house Database, welche in der Software MBT compass explorer (Bruker Daltonic, Deutschland) implementiert

7.1.2. *Interspezies-Klassifikation: Speziesspezifische Differenzierung durch MALDI-TOF MS*

Eine Unterscheidung der beiden Spezies *F. gigantea* und *F. hepatica* konnte bereits in der visuellen Beurteilung der Spektren (siehe Abb. 17) erfolgen. 8/9 (88.8%) aller adulten *F. hepatica*-Proben wurden durch die in-house-Database, welche sowohl MSPs von *F. gigantea*, als auch *F. hepatica* enthielt, ihrer korrekten Spezies mit einem LSV ≥ 2.0 zugeordnet. Ebenso wurden von allen 50 anterioren Proben von *F. gigantea* 42/50 (84%) mit einem LSV über 2.0 identifiziert. Eine Fehlklassifizierung zur falschen Spezies fand nicht statt, bei 8/50 (16%) lag der LSV jedoch im Bereich 1.70-1.99 und konnte somit keiner Spezies zugeordnet werden.

Zur Klassifizierung mit der Clover MS Data Analysis[®] Software zur Untersuchung verschiedener LA wurden die in Kapitel 5.5.1. beschriebenen Parameter verwendet. In der Analyse der LA konnte im unüberwachten LA (PCA) eine klare Separierung der beiden Spezies festgestellt werden (siehe Abb. 18). Sowohl adulte als auch Proben im Eistadium konnten der Spezies zufolge klar getrennt werden. Auch alle überwachten Klassifizierungsalgorithmen (SVM, RF, PLS-DA, KNN) zeigten eine gute Unterscheidung zwischen *F. hepatica* und *F. gigantea* (siehe Abb. 20). Die Ergebnisse der zehnfachen Kreuzvalidierung zeigten Genauigkeitswerte zwischen 96.3% und 100 %, und F1-Werte (das harmonische Mittel aus Präzision und Sensitivität) zwischen 94,74 % und 100 % (Tabelle 5, Abschnitt C, im Anhang). Die LA SVM und RF zeigte dabei die besten Ergebnisse, ohne Fehlklassifizierungen, d. h. eine Genauigkeit von 100 % und einen F1-Wert von 100 %. Darüber hinaus wurden diese Werte durch die nahezu bei 1 liegenden Werte Fläche unter der AUROC-Kurve (0,995 für SVM, 0,994 für PLS-DA, 0,995 für RF und 0,945 für KNN) und Werte für die Fläche unter der AUPR-Kurve (1,00 für SVM, 0,99 für PLS-DA, 1,00 für RF und 0,968 für KNN) bestätigt (Siehe Abb. 32, im Anhang).

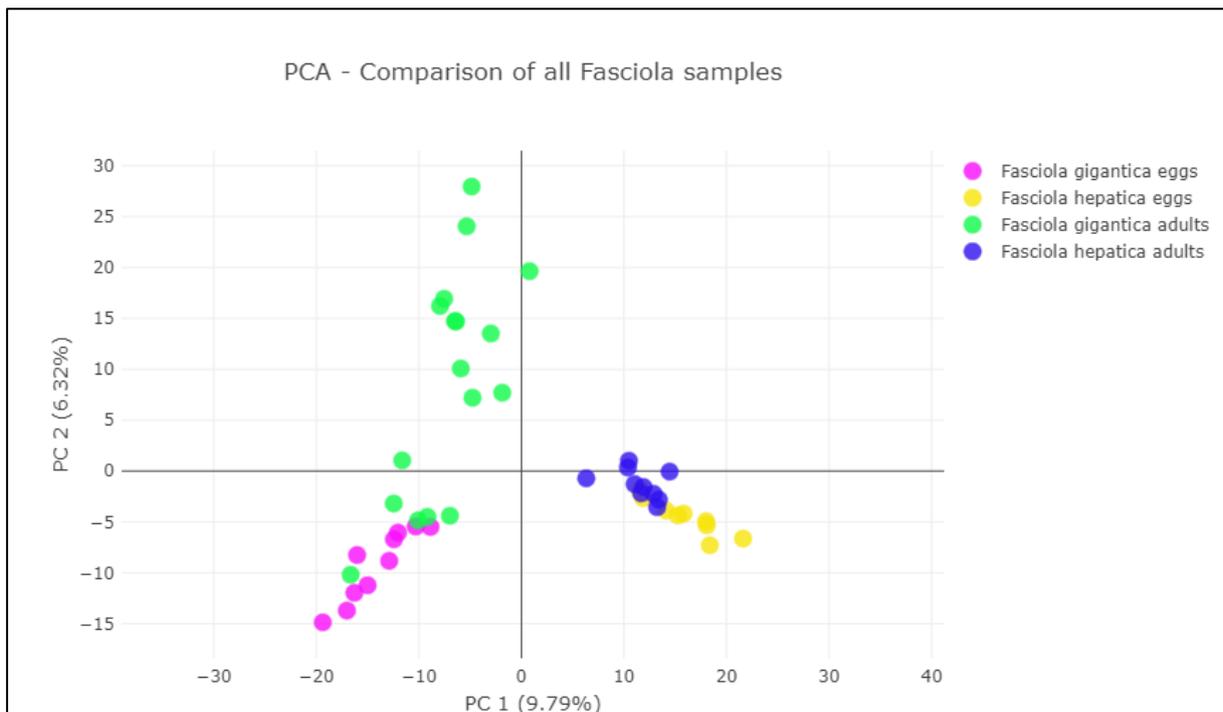


Abbildung 18: Unüberwachter LA (PCA): Darstellung aller *Fasciola*-Proben: adulte *F. gigantica* (posteriore und anteriore Anteile sind zusammengefasst) (grün), adulte *F. hepatica* (blau), *F. gigantica* Eier (pink), *F. hepatica* Eier (gelb)

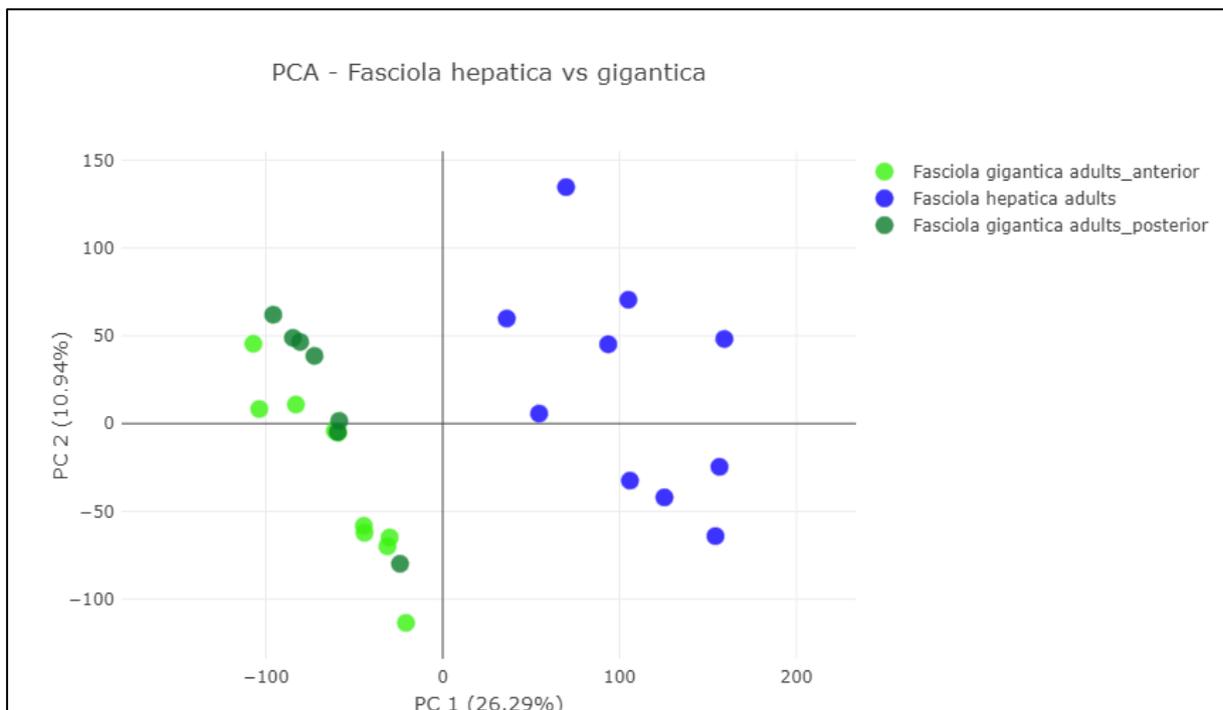


Abbildung 19: Unüberwachter LA (PCA): Darstellung adulte *F. hepatica*- (blau) und adulte *F. gigantica*-Proben, welche wiederum in *F. gigantica* anteriorer Anteil (hellgrün) und posteriorer Anteil (dunkelgrün) aufgeteilt sind.

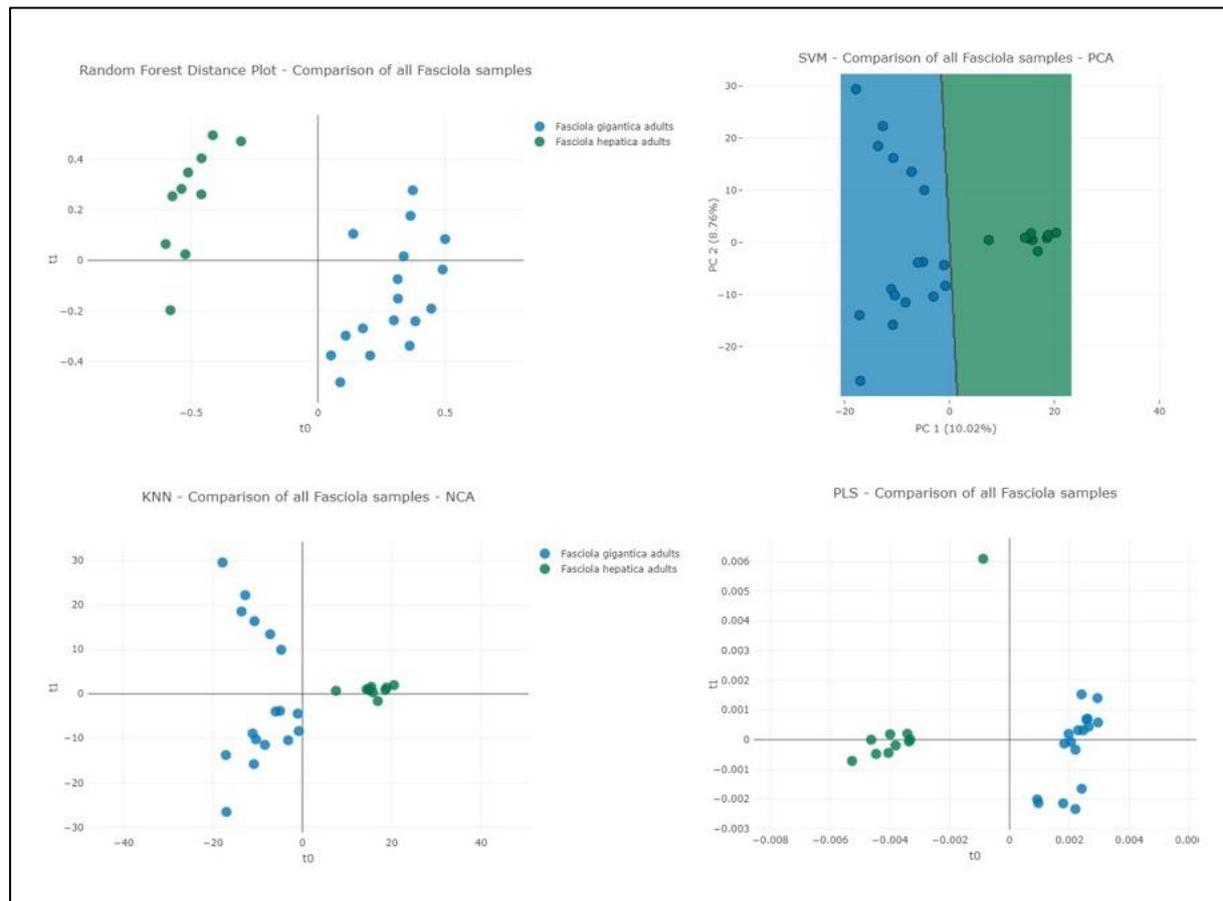


Abbildung 20: Überwachte LA (RF, SVM, KNN, PLS-DA) zur Darstellung von adulten *F. gigantica*-Proben (grün) und adulten *F. hepatica*-Proben (blau)

7.1.3. *Intraspezies-Klassifikation: Einfluss des Entwicklungsstadiums auf die MALDI-TOF basierte Diagnostik*

Um den Einfluss des Entwicklungsstadiums zu untersuchen, wurde die positive Identifikationsrate von adulten Proben und *Fasciola*-Eiern herangezogen. Zur Einschätzung, inwieweit sich die unterschiedlichen Entwicklungsstadien in ihren Proteinprofilen unterscheiden, wurde die Identifikationsrate durch die 7 MSPs adulter *F. gigantica*-Proben, welche aus dem posterioren Anteil der Würmer gewonnen wurde, verwendet, da diese keine Eier enthalten. Wie in der Tabelle 2 zu sehen ist, ergab sich in diesem Vergleich ein durchschnittlicher LSV von 2.02. Alle (22/22 (100%)) der Proben konnten auf dem Gattungslevel identifiziert werden und 68.22%, also 15 von 22 Proben auf dem Spezieslevel mit einem LSV über 2.0.

Im zweiten Schritt wurden 10 Ei-Proben zur Erstellung von Referenzspektren verwendet und zur in-house-Database hinzugefügt (siehe Kapitel 7.1.1). Im Vergleich zu den Identifikationsraten durch MSPs der adulten Würmer konnte eine deutliche Steigerung der Identifikation und der korrespondierenden LSV-Werte erreicht werden. Der durchschnittliche LSV lag nun bei 2.29 und es konnten 22/22 (100%) der Proben auf dem Spezieslevel erkannt werden.

Für *F. hepatica* ergaben sich bei der Identifizierung der Proben im Ei-Stadium ähnliche Ergebnisse. Hier wurden alle Proben mit der bereits um die 2 MSPs der Eier erweiterte Datenbank identifiziert. 7/8 (87,5%) der für den Blindtest verwendeten Proben konnten mit einem LSV über 2.0 identifiziert und somit ihrer Spezies korrekt zugeordnet werden. Lediglich eine Probe konnte nicht identifiziert werden.

Diese Ergebnisse ließen sich auch in der Analyse der LA darstellen. Sowohl bei *F. gigantica*, als auch bei *F. hepatica* konnte in der Übersicht aller *Fasciola*-Proben (siehe Abb. 18) nur eine uneindeutige Separierung des Entwicklungsstadiums erkannt werden. Daher wurde zur genaueren Untersuchung ein Vergleich innerhalb der jeweiligen Spezies durchgeführt und auch überwachte LA eingesetzt, um eventuelle Unterschiede sichtbar zu machen.

Für *F. gigantica* konnte bereits im unüberwachten LA (PCA) eine Differenzierung der Entwicklungsstadien erfolgen (siehe Abb. 21). In den überwachten LA bestätigte sich diese Beobachtung (siehe Abb. 22). Die Ergebnisse der zehnfachen Kreuzvalidierung zeigten Genauigkeitswerte zwischen 96.3% und 100 %, und F1-Werte zwischen 97,14 % und 100 % (siehe Tabelle 2, Abschnitt A, im Anhang). Dabei zeigten sich beim LA SVM die besten Ergebnisse, ohne Fehlklassifizierungen, d. h. eine Genauigkeit von 100 % und einen F1-Wert von 100 %.

In der Untersuchung der *F. hepatica*-Proben konnte ebenfalls eine Separierung des Entwicklungsstadiums sowohl im unüberwachten als auch in den überwachten LA erkannt werden (siehe Abb. 23 und 24). Hierbei lagen die Genauigkeitswerte zwischen 95% und 100 %, und die F1-Werte zwischen 90 % und 95,24 % (siehe Tabelle 2, Abschnitt B, im Anhang). Dabei war der LA SVM mit einem F1 Score etwas schlechter als PLS-DA, RF, und KNN.

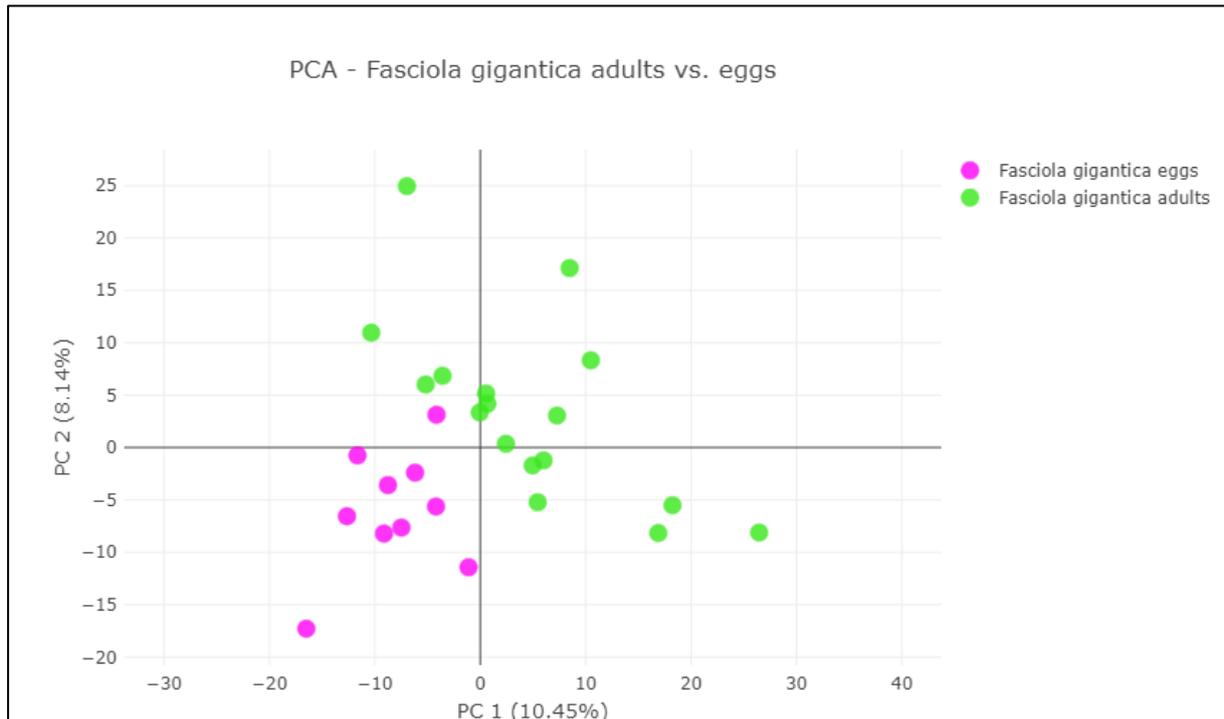


Abbildung 21: Unüberwachte LA (PCA): Darstellung *F. gigantica* adulte Proben (posteriore und anteriore Anteile sind zusammengefasst) (grün) und Eier (pink)

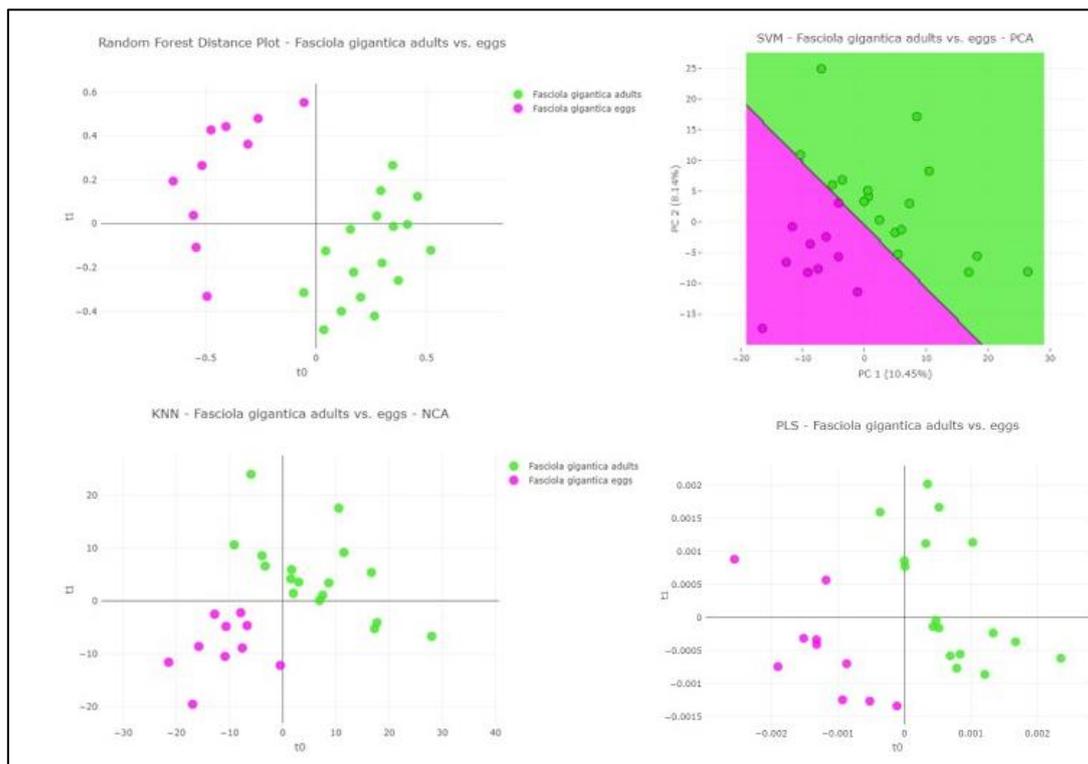


Abbildung 22: Überwachte LA (RF, SVM, KNN, PLS-DA) zur Darstellung von adulten *F. gigantica*-Proben (grün) und *F. gigantica*-Eiern (pink)

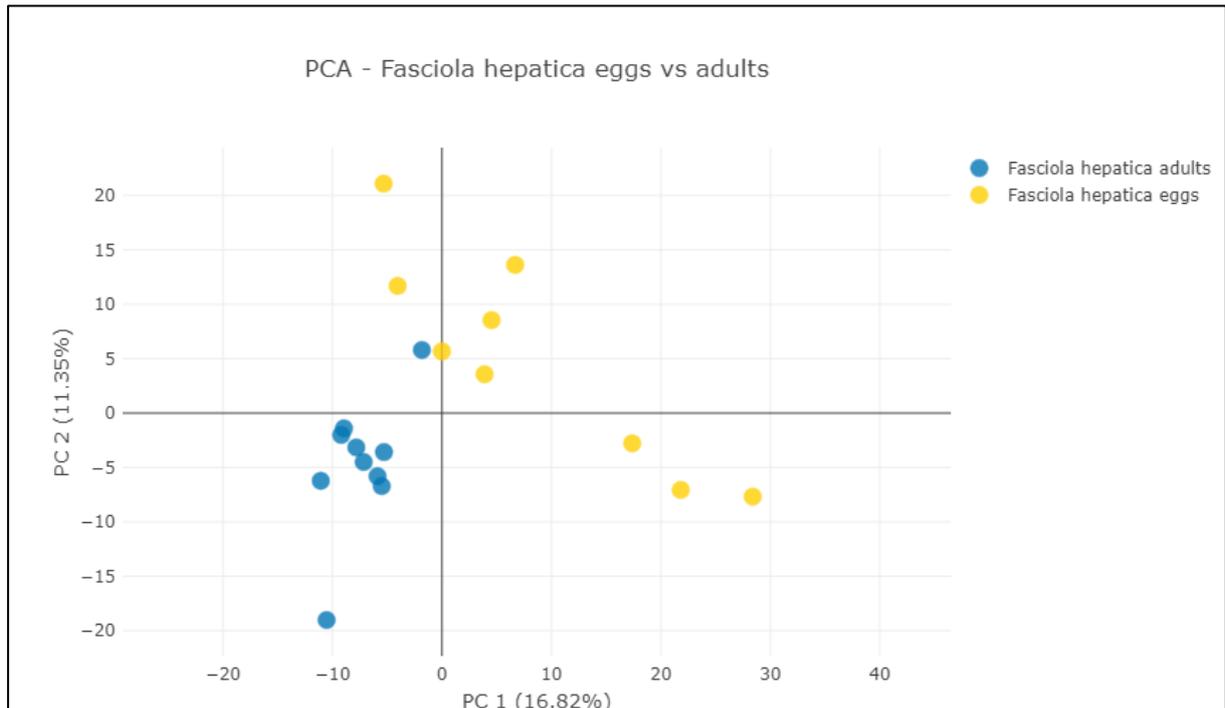


Abbildung 23: Unüberwachter LA (PCA) zur Darstellung adulter *F. hepatica*-Proben (blau) und *F. hepatica*-Eier (gelb)

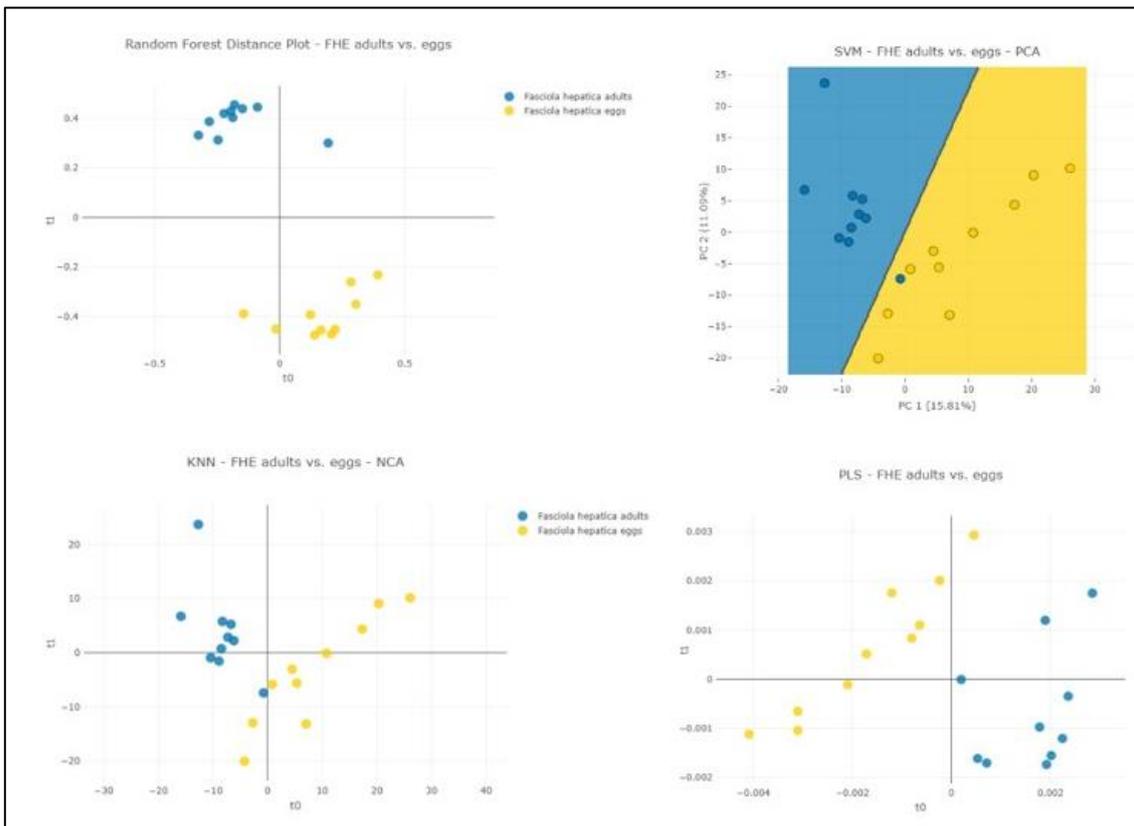


Abbildung 24: überwachter LA (RF, SVM, KNN, PLS-DA) zur Darstellung von adulten *F.hepatica*-Proben (blau) und *F.hepatica*-Eiern (gelb)

7.1.4. *Intraspezies-Klassifikation: Einfluss des verwendeten Körperanteils*

Alle für diese Arbeit verwendeten anterioren *F. gigantea*-Proben wurden mit MSPs, welche aus dem posterioren *F. gigantea* Anteil 12 Monate vor Messung der anterioren Anteile verarbeitet wurden, identifiziert. Dabei zeigte sich, dass 48/50 (96%) Proben auf dem Gattungslevel und 42/50 (84%) davon auf dem Spezieslevel ($LSV \geq 2.0$) identifiziert wurden.

Diese Beobachtung konnte durch die Anwendung unüberwachter LA (PCA) bestätigt werden. Dabei konnte keine Differenzierung zwischen dem verwendeten Körperanteil beobachtet werden, wohingegen ein klarer Unterschied der Spezies zu sehen war (siehe Abb. 18 und 19). Da bereits im unüberwachten LA keine Separierung dieser Proben zu sehen war, als auch eine Identifikation der anterioren Anteile durch die MSPs der posterioren Anteile gelang, wurden im Weiteren die *F. gigantea* anterior/posterior-Proben als *F. gigantea* adults (adulte Exemplare) zusammengefasst.

7.2. *Schistosoma mansoni* und *Schistosoma japonicum*

7.2.1. Validierung der MALDI-TOF MS basierten Diagnostik von *S. japonicum* und *S. mansoni*

Ausschluss von Kontaminationen und Spektrumanalyse

Bei einer Analytik aller Proben mit der kommerziellen BDAL-Datenbank, um eine mögliche Kontamination durch Bakterien oder Pilze zu erkennen, zeigte sich kein Hinweis auf das Vorhandensein eines in dieser kommerziellen Datenbank enthaltenen Mikroorganismus (siehe Tabelle 3).

<i>Schistosoma</i> Spezies	Anzahl der Proben	Anzahl der Proteinspektren	Korrekte Identifikation	Durchschnitt LSV
<i>S. mansoni</i> in 70% (v/v) Ethanol	19	561	0%	1.34
<i>S. mansoni</i> in RNAlater®	21	512	0%	1.27
<i>S. japonicum</i> in 70% (v/v) Ethanol	11	364	0%	1.35
<i>S. japonicum</i> in RNAlater®	11	392	0%	1.29

Tabelle 3: *Schistosoma* spp. vs. BDAL- Datenbank von Bruker Daltronic

Für alle *Schistosoma*-Proben konnten Spektren generiert werden, welche vollumfänglich zur Identifikationsanalyse verwendet wurden. Die Spektren der beiden untersuchten Spezies konnten bereits durch optische Beurteilung unterschieden werden und zeigten jeweils innerhalb der Spezies ähnliche Spektrenprofile. Die beiden Lagerungsmedien 70% (v/v) Ethanol und RNAlater®, wiesen ebenfalls in der optischen Beurteilung einige geringfügige Unterschiede untereinander auf (siehe Abb. 25).

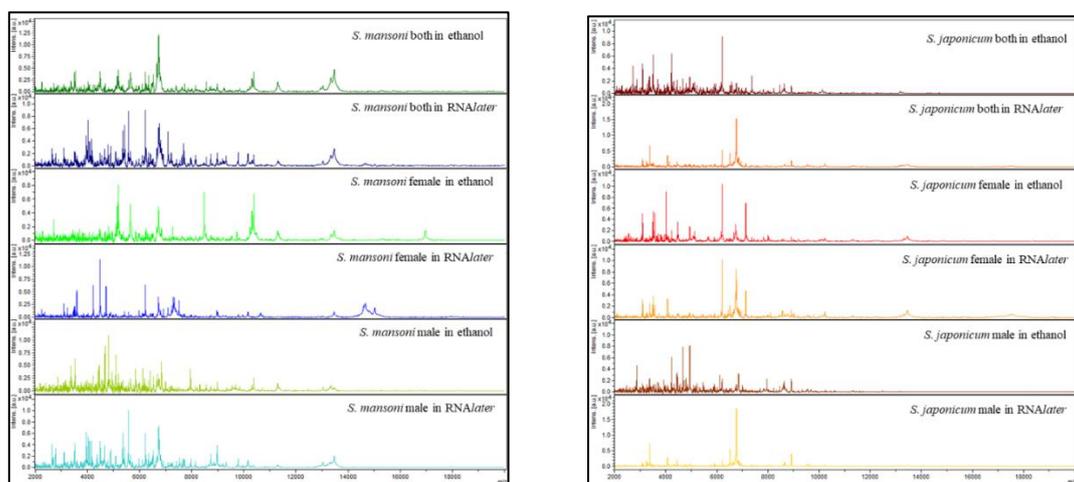


Abbildung 25: Proteinspektren von *S. mansoni*-Proben (links) und *S. japonicum*-Proben (rechts)

Validierung der In-house-Database

Es erfolgte sowohl eine interne als auch eine externe Validierung der erweiterten in-house-Database (siehe Kapitel 7.1.1.).

Die interne Validierung ergab eine korrekte Identifizierung von 100% der Proben. Es ergaben sich durchweg Identifizierungen auf dem Spezieslevel ($LSV \geq 2.0$) (siehe Tabelle 4, Abschnitt (A)).

Für die externe Validierung, bei welcher alle weiteren, bis dahin unbekanntes Proben mit der erweiterten Datenbank verglichen wurden (siehe Tabelle 4, Abschnitt (B)), konnte auf der Gattungsebene ($LSV \geq 1.7$) ebenfalls eine korrekte Identifizierung von allen 58 (100%) Proben erreicht werden. Auf der speziesspezifischen Ebene ($LSV \geq 2.0$) wurden 47/58 (81.03%) Proben korrekt identifiziert. Bei den gepaarten Proben war die Identifizierung zu 100% korrekt und damit genauer als bei den Proben, welche nach Geschlecht der Würmer getrennt waren (Tabelle 4, Abschnitt (B)). Die durchschnittlichen LSV-Werte lagen zwischen 2.08 und 2.54.

Interne und externe Validierung von *Schistosoma* – Proben

(A) Interne Validierung											
Ergebnisse der Identifikation											
Spezies	Lagerungs- medium	Anzahl der Proben	Spektren	LSV \geq 1.7		LSV \geq 2.0					
				♀+♂		♀+♂					
<i>S. mansoni</i>	70% (v/v) Ethanol	1	32	32/32 (100%)		32/32 (100%)					
	RNAlater®	1	27	27/27 (100%)		32/32 (100%)					
<i>S. japonicum</i>	70% (v/v) Ethanol	1	28	28/28 (100%)		32/32 (100%)					
	RNAlater®	1	32	32/32 (100%)		32/32 (100%)					
Total		4	119	119/119 (100%)		119/119 (100%)					

(B) Externe Validierung											
Ergebnisse der Identifikation											
Spezies	Lagerungs- medium	Anzahl der Proben	Ø LSV	LSV \geq 1.7				LSV \geq 2.0			
				♀+♂	♀	♂	Total	♀+♂	♀	♂	Total
<i>S. mansoni</i>	70% (v/v) Ethanol	18	2.08	6/6 (100%)	6/6 (100%)	6/6 (100%)	18/18 (100%)	6/6 (100%)	5/6 (83.33%)	0/6 (0%)	11/18 (61.11%)
	RNAlater®	20	2.54	10/10 (100%)	5/5 (100%)	5/5 (100%)	20/20 (100%)	10/10 (100%)	4/5 (80%)	5/5 (100%)	19/20 (95%)
<i>S. japonicum</i>	70% (v/v) Ethanol	10	2.23	4/4 (100%)	3/3 (100%)	3/3 (100%)	10/10 (100%)	4/4 (100%)	0/3 (0%)	3/3 (100%)	7/10 (70%)
	RNAlater®	10	2.54	4/4 (100%)	3/3 (100%)	3/3 (100%)	10/10 (100%)	4/4 (100%)	3/3 (100%)	3/3 (100%)	10/10 (100%)
Total		58	2.34	26/26 (100%)	17/17 (100%)	17/17 (100%)	58/58 (100%)	24/24 (100%)	12/17 (70.59%)	11/17 (64.71%)	47/58 (81.03%)

Tabelle 4: (A) Interne und (B) Externe Validierung: Identifizierung von *Schistosoma*-Proben, die in verschiedenen Lagerungsmedien (70% (v/v) Ethanol und RNAlater®) aufbewahrt wurden, unter Verwendung erweiterter in-house-Database, welche in der Software MBT compass explorer (Bruker Daltonic, Deutschland) implementiert ist

7.2.2. *Interspezies-Klassifikation: Speziesspezifische Differenzierung durch MALDI-TOF MS*

Zur speziesspezifischen Untersuchung von *Schistosoma*-Proben wurden 38 *S. mansoni* und 20 *S. japonicum*-Proben mit der um 2 MSPs von *S. mansoni* und 2 MSPs von *S. japonicum* erweiterten in-house-Databas verglichen. Eine Identifikation als *Schistosoma* spp., auf dem Gattungslevel mit einem LSV von mindestens 1.7, war bei allen getesteten Proben möglich. Da ausschließlich gepaarte Proben zur Erstellung von MSPs verwendet wurden, war auch bei der Identifikation von rein männlichen oder rein weiblichen Proben keine Minderung der Identifikationsrate auf dem Gattungslevel zu erkennen.

Auf dem Spezieslevel konnte bei 81.03% aller Proben eine Identifikation mit einem LSV von über 2.0 erreicht werden, dies entspricht 47 aller 58 gemessenen Proben. Der durchschnittliche LSV-Wert lag bei 2.34. Dabei konnten anteilig mehr Proben von *S. japonicum* (17/20 (85 %)) als von *S. mansoni* (30/38 (79 %)) korrekt ihrer Spezies (LSV>2.0) nach identifiziert werden. 11 Proben (8 *S. mansoni*-Proben und 3 *S. japonicum*-Proben) konnten lediglich ihrer Gattung *Schistosoma* spp. zugeordnet werden, es fanden dabei aber keine Fehlklassifikationen zu einer falschen Spezies statt.

Zur Klassifizierung mit der Clover MS Data Analysis® Software zur Untersuchung verschiedener LA wurden die in Kapitel 5.5.1. beschriebenen Parameter verwendet. Es ergaben sich in allen angewendeten LA klare Trennungen zwischen *S. mansoni* und *S. japonicum* (siehe Abb. 26 und 27).

Die überwachten LA sind als Abstandsdiagramme dargestellt. In der 10-fachen Kreuzvalidierung konnte eine geringfügige Überlegenheit des RF LA festgestellt werden (siehe Tabelle 6, im Anhang). So war keine Fehlklassifizierung in diesem LA zu beobachten, der F1 Score lag bei 100% und die Genauigkeit ebenfalls bei 100%. Die Werte der drei weiteren Klassifizierungsalgorithmen (SVM, PLS-DA, KNN) lagen beim F1 Score zwischen 94.45% und 100% und bei der Genauigkeit zwischen 96.77% und 100%. Auch diese Werte lassen auf eine speziesspezifische Spektrumsprofile schließen. Diese Beobachtung wird auch durch die AUROC-Werte (0,955 für SVM; 0,995 für PLS-DA; 0,995 für RF; und 0,983 für KNN) und die AUPR-Werte (0,98 für SVM; 1,00 für PLS-DA; 1,00 für RF; und 0,97 für KNN) gestützt (siehe Abb. 33, im Anhang).

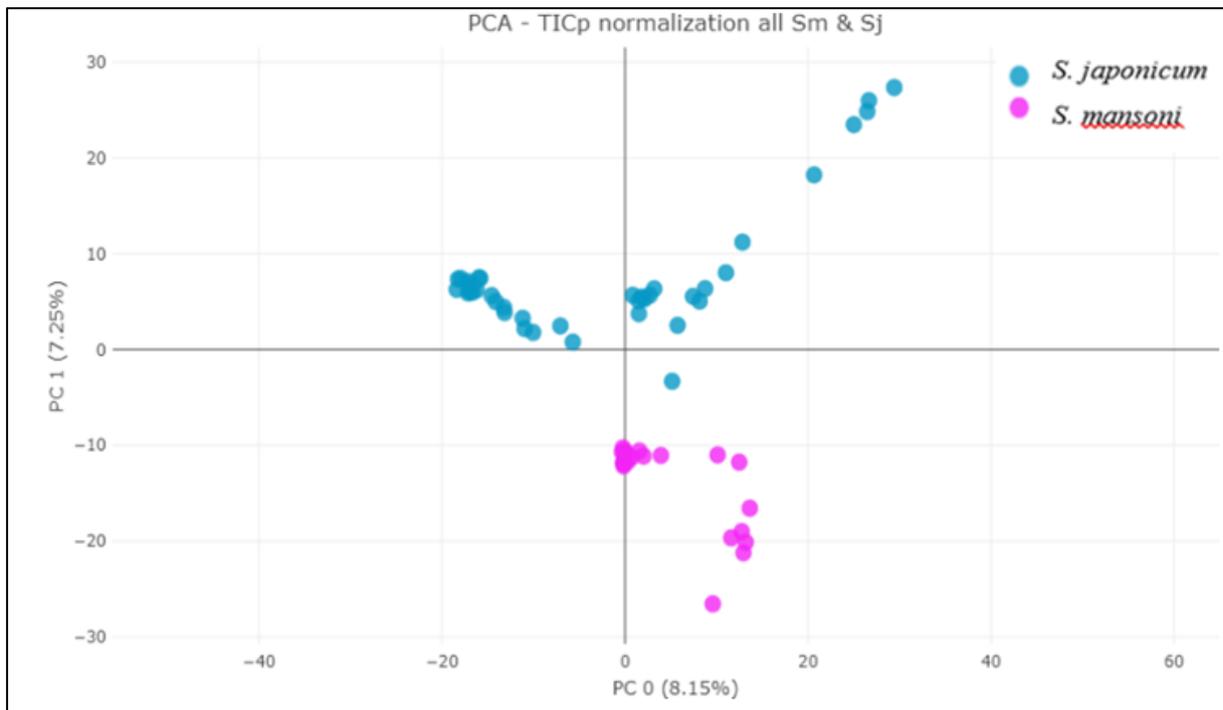


Abbildung 26: unüberwachter LA (PCA) zur Darstellung von *S. mansoni* (pink) und *S. japonicum* (blau)

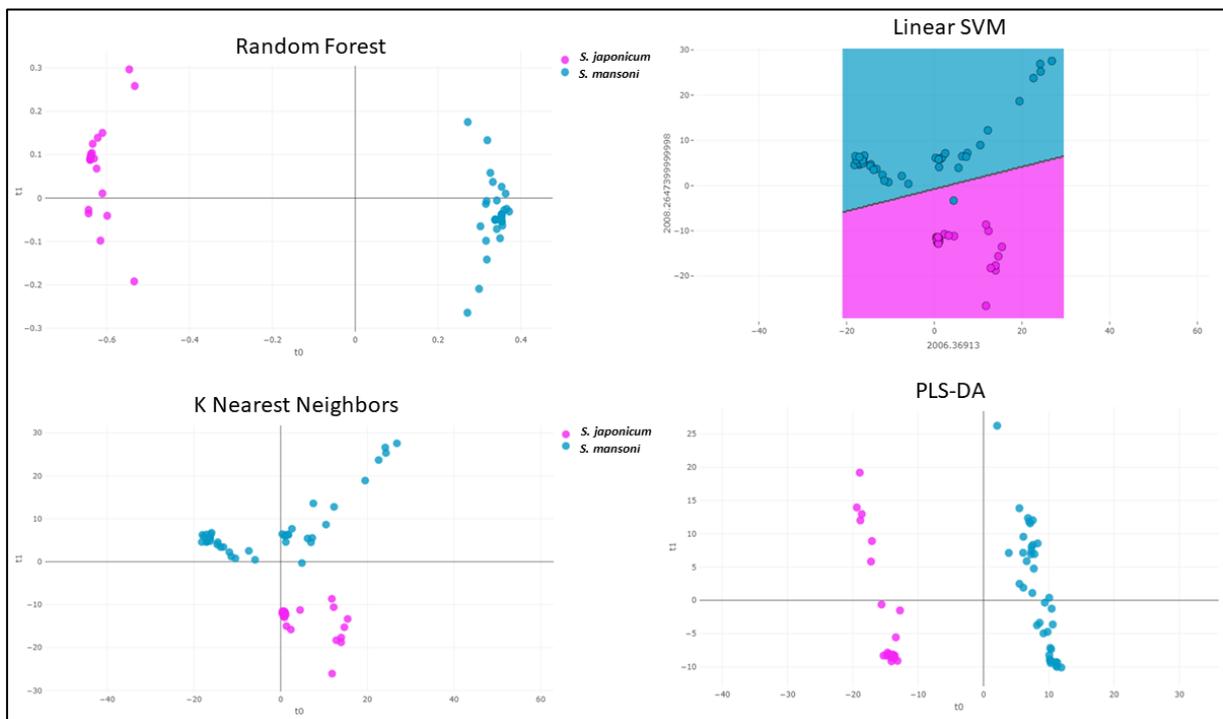


Abbildung 27: überwachter LA (RF, SVM, KNN, PLS-DA) zur Darstellung von *S. mansoni* (pink) und *S. japonicum* (blau)

7.2.3. *Intraspezies-Klassifikation: Einfluss des Lagerungsmediums auf MALDI-TOF MS basierte Diagnostik*

Zur Untersuchung von Einflussfaktoren wurden beide Spezies getrennt und innerhalb einer Spezies der Einfluss der Lagerungslösungen untersucht. In der Auswertung der LSV-Werte zeigte sich, dass bei Lagerung in RNAlater® mehr Proben, 96.6% (29/30) mit einem LSV über 2.0 identifiziert wurden als bei Lagerung in Ethanol, 35.7% (10/28 Proben).

Mittels PCA-Analyse sowie mit überwachten LA (SVM, PLS-DA, KNN) ließ sich in der Betrachtung der Grafiken eine deutliche Differenzierung der gemessenen Proteinspektren in Abhängigkeit des Lagerungsmediums (70% (v/v) Ethanol oder RNAlater®) erkennen (siehe Abb. 28-31). Hierbei war keine Überlegenheit eines einzelnen LA zu erkennen und alle wiesen in der 10-fachen Kreuzvalidierung eine Genauigkeit von 100% und einen F1 Score von ebenfalls 100% auf, mit Ausnahme des SVM, der mit dem *S. mansoni*-Datensatz getestet wurde (Genauigkeit 97,5%, F1-Score 97,6%) (siehe Tabelle 6, im Anhang).

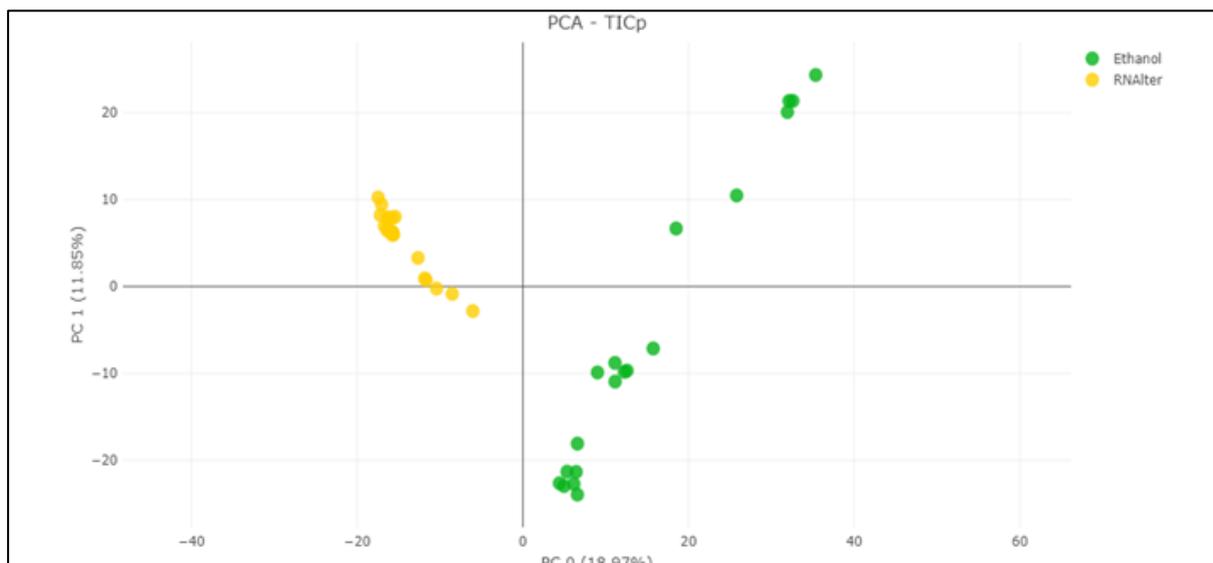


Abbildung 28: unüberwachter LA (PCA) zur Darstellung von *S. mansoni* in 70% (v/v) Ethanol (grün) und RNAlater® (gelb)

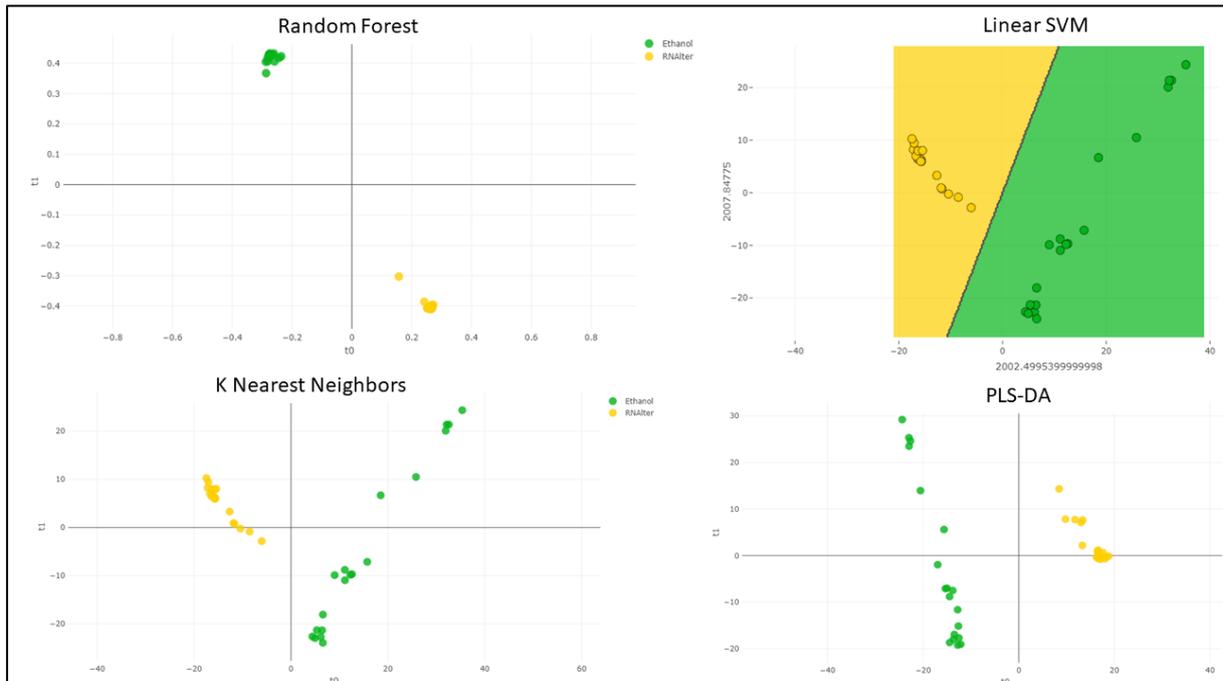


Abbildung 29 überwachter LA (RF, SVM, KNN, PLS-DA) zur Darstellung von *S. mansoni* in 70% (v/v) Ethanol (grün) bzw. RNAlater® (gelb)

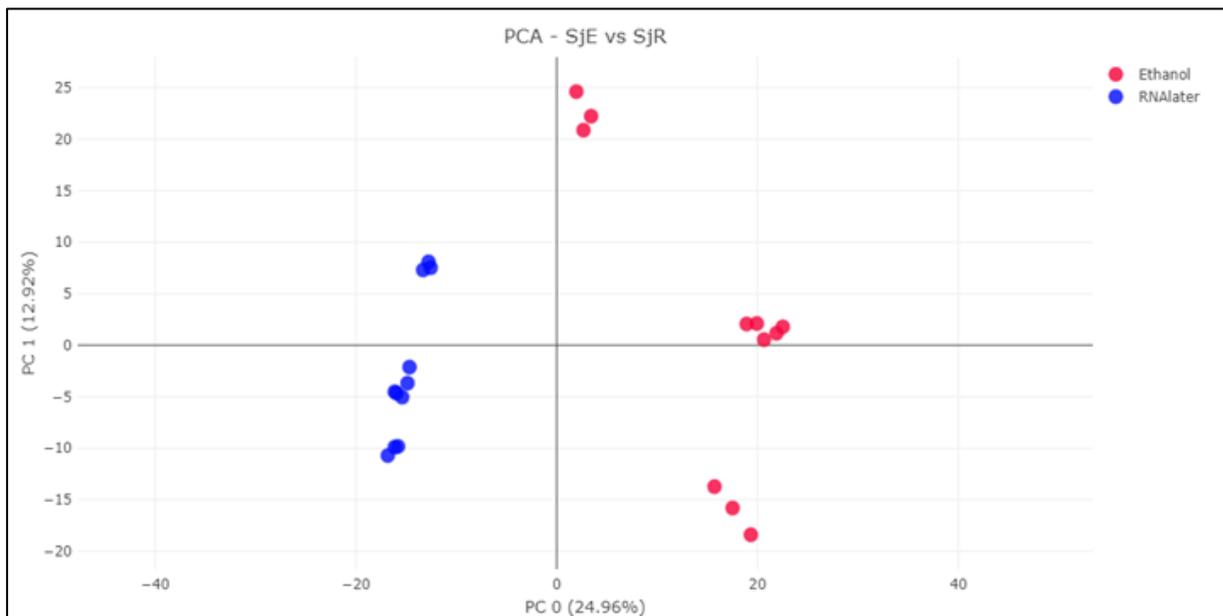


Abbildung 30: unüberwachter LA (PCA) zur Darstellung von *S. japonicum* in 70% (v/v) Ethanol (rot) bzw. RNAlater® (blau)

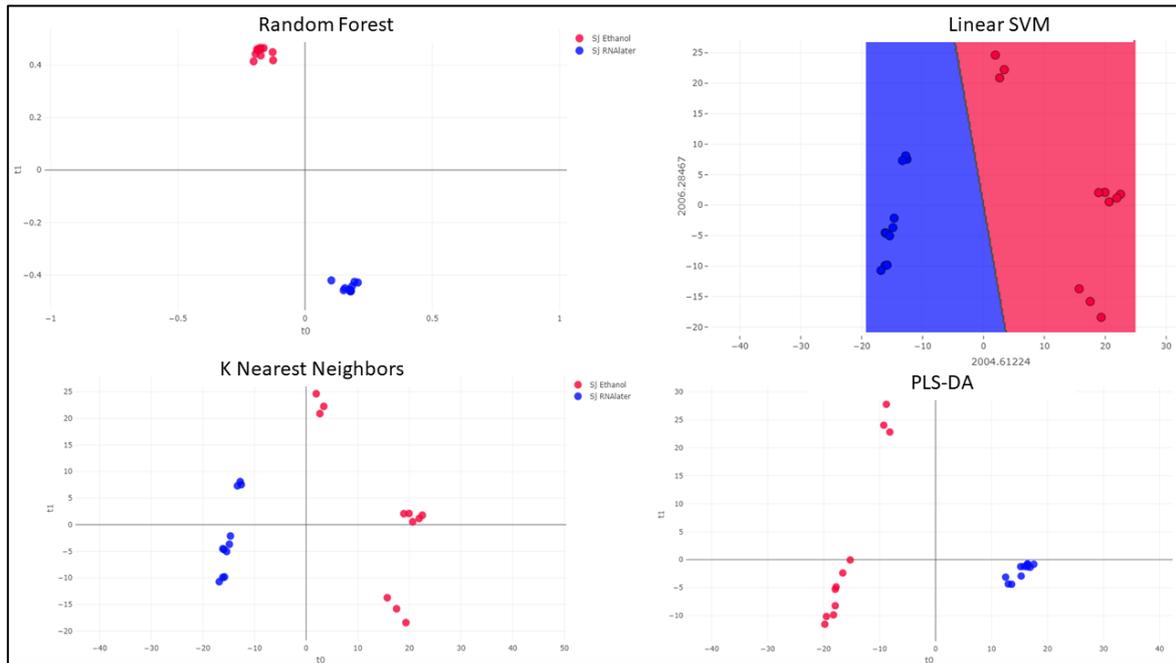


Abbildung 31: überwachter LA (RF, SVM, KNN, PLS-DA) zur Darstellung von *S. japonicum* in 70% (v/v) Ethanol (rot) bzw. RNAlater® (blau)

8. Diskussion

8.1. Kritische Betrachtung der eigenen Untersuchung

Gewählte Methodik

In früheren Studien zur Verwendung von MALDI-TOF MS zur Identifizierung von Trematodenarten konnten vielversprechende Ergebnisse erzielt werden, welche die mögliche Anwendung dieser Technik als Diagnostikmethode belegen [99][49]. Diese Ergebnisse decken sich mit den in der vorliegenden Arbeit erlangten Erkenntnissen. So waren in den Referenzspektren sowohl der *Schistosoma*-Proben als auch der *Fasciola*-Proben die Peaks zum großen Anteil im Bereich von 3000 kDa bis 8000 kDa verortet. Bemerkenswert ist, dass es bei den *Schistosoma*-Proben zu keinen Fehlidentifizierung auf der Gattungsebene kam und 100% der Proben der korrekten Art zugeordnet wurden. Für die *Fasciola*-Proben wurden ähnliche Ergebnisse erreicht. Nur 2 Proben (*F. gigantica*, anteriorer Anteil) konnten mit der in-house-Database nicht identifiziert werden und erreichten keinen LSV über 1.7. Auch wurde keine Fehlidentifizierung zwischen den Spezies beobachtet.

In dieser Arbeit wurden ausschließlich Proben verwendet, welche zu Forschungszwecken aus Tieren entnommen wurden. So stammten die *Schistosoma*-Proben aus experimentell infizierten Mäusen, die *F. gigantica*-Proben aus Lebergewebe geschlachteter Rinder in Nigeria und die *F. hepatica*-Proben aus Rindern in der Schweiz. Zur Untersuchung der Entwicklungsstadien wurden die Eier der *Fasciola*-Proben aus den Uteri der adulten Würmer extrahiert. Dies entspricht nicht der für die Routinediagnostik verfügbaren Extrahierungsmethoden. So werden in der Regel Stuhl-, Urin-, Serum-, oder Vollblutproben entnommen, welche die Eier der Parasiten enthalten und auf diese untersucht. Es bleibt also offen, ob eine Identifizierung der Eier direkt aus einer Stuhlprobe mittels MALDI-TOF MS möglich ist oder ob vorangehende Reinigungsschritte des Kots nötig sind, um lediglich die Eier zur Proteinextraktion zu verwenden. In der Diagnostik von Bakterien ist es bereits möglich, direkt Blut- oder Urinproben, ohne vorherige Anzucht und Proteinextraktion, massenspektrometrisch zu untersuchen [82]. In diesem Gebiet ist auch für die Parasitendiagnostik weitere Forschung nötig.

Für die Proteinextraktion wird eine ausreichende Menge an Material benötigt. Für die adulten *Fasciola*-Proben war ein lediglich 0.5 cm langes Stück des Körpers notwendig, bei den Eiern haben wir eine geschätzte Mindestanzahl von 500 Eiern pro Proteinextraktion verwendet. Bei

den *Schistosoma*-Proben wurden mindestens 2 adulte Proben zur Proteinextraktion benötigt. Für die klinische Diagnostik sind also bestimmte Mindestmengen an parasitärem Material notwendig, beispielsweise eine Mindestanzahl von EPG in Stuhlproben. Dies könnte mitunter zu einem Problem werden. So finden sich bei großen Stuhlmengen oder leichten Infektionen oft eine verminderte Konzentration an Eiern im Kot [67].

Die Erstellung von Referenzspektren muss unbedingt von morphologisch und molekulargenetisch identifizierten Exemplaren gewonnen werden. Die Qualität der MSPs ist zur korrekten und speziesspezifischen Identifikation mittel MALDI-TOF MS von zentraler Bedeutung [87]. Bei Proben von der Größe der adulten *F. gigantica*-Proben ist dies theoretisch möglich, indem ein anderes Körperteil für die molekulare Analyse verwendet wird. Zur Identifikation der adulten *Schistosoma*-Proben ist dies schwieriger, da der gesamte Körper zur Proteinextraktion verwendet wird. Daher müssen Methoden entwickelt werden, die es zuverlässig ermöglichen, sowohl die Proteinextraktion zur Erstellung von MSPs als auch die molekulargenetische Untersuchung zur Sicherung der Art und Spezies mit exakt dem gleichen Ausgangsmaterial durchzuführen. Obwohl bei der Proteinextraktion in der Regel Bedingungen herrschen, die die DNA schädigen, war es in einer Studie von Bredtmann et al. möglich, aus dem in Acetonitril/Ameisensäure unlöslichen Material DNA in ausreichender Qualität für einen PCR Test zu extrahieren [12]. Für weiterführende Forschung und Implementierung in eine kommerzielle Datenbank wäre eine Erstellung von Referenzspektren mit gleichzeitiger molekulargenetischer Identifizierung derselben Exemplare anzustreben.

Beim ML weisen geringfügige Unterschiede, welche zwischen den verschiedenen LA gefunden wurden, darauf hin, dass auch unter Verwendung unterschiedlicher Algorithmen des ML eine gute Differenzierung der Proteinspektren möglich ist. Die Leistung der ML-Systeme hängt von der Qualität der Daten ab, die ihnen zur Verfügung gestellt werden. Alle Einschränkungen, welche für die Rohdaten der Proben beschrieben sind oder im Folgenden beschrieben werden (tierische statt humaner Proben, molekulargenetische Untersuchung ausschließlich von einzelnen Exemplaren einer Entnahmegruppe etc.) wirken sich somit direkt auf die Qualität der trainierten Modelle aus.

Untersuchte Einflussfaktoren

Entwicklungsstadien

In meiner Arbeit habe ich die Untersuchung von 2 verschiedenen Entwicklungsstadien von *F. gigantea* und *F. hepatica* untersucht. Für *F. gigantea* konnte eine Identifizierung der Eier durch MSPs von adulten Proben mit einem LSV von mindestens 1.7 erreicht werden und bei fast 70% der Proben auch auf dem Spezieslevel, sodass eine Identifizierung ebenfalls mit einem MSP einer anderen Entwicklungsstufe möglich ist. Eine Verbesserung der Identifikation wird durch die Inkludierung von MSPs verschiedener Entwicklungsstadien erreicht. Damit konnten für *F. gigantea*-Eier eine Identifikation von 100% der Proben erreicht werden und für *F. hepatica* Eier bei 87.5% der Proben. Dies deckt sich mit den Ergebnissen von Karger et. al. der postulierte, dass die Art die primäre Determinante zur Identifizierung ist und erst innerhalb einer Art die Differenzierung der Entwicklungsstufe möglich ist [54]. Auch ließen sich diese Ergebnisse mit unseren Erkenntnissen in der Darstellung der LA des ML bestätigen. So war auffällig, dass das Entwicklungsstadiums eine klare Differenzierung der Eier gegen die adulten Proben in der Intraspezies-Klassifikation zuließ, die Unterschiede in der Interspezies-Klassifizierung aber nicht mehr augenscheinlich waren. Dies lässt die Schlussfolgerung zu, dass bei der Identifizierung neuer Proben weniger das Entwicklungsstadiums, als vielmehr die Spezies entscheidend für eine korrekte Identifizierung ist, allerdings zur Verbesserung der Ergebnisse, Einflussfaktoren wie verschiedene Entwicklungsstadien durchaus eine Rolle spielen.

Körperanteil

In der Analyse des Einflusses des verwendeten Körperanteils bei der Identifikation von *F. gigantea*-Proben ergaben sich keine signifikanten Unterschiede. Die Untersuchung des anterioren Anteils der Würmer fand 12 Monate nach der Untersuchung des posterioren Anteils statt, die Ergebnisse wurden so auch auf den Einfluss der Lagerungsdauer übertragen. Auf Grund der geringfügigen Veränderungen der Identifikationsraten lässt sich der Rückschluss ziehen, dass sowohl die Lagerungsdauer als auch der verwendete Anteil des Wurmes sich nicht in relevanter Weise auf die diagnostische Genauigkeit auswirken. Diese Beobachtungen decken sich mit einer Studie von Rivero et al. zur Identifikation von *Trichuris* spp., welche keinen signifikanten Einfluss des verwendeten Körperanteils auf die Identifikation festgestellt konnten [88]. Ebenso konnte in einer Studie zu *Taenia saginata* von Wendel et al. nachgewiesen werden, dass die Proben auch nach einer 24-wöchigen Lagerung noch zuverlässig identifiziert werden konnten [108]. In der vergleichenden Untersuchung der *F. gigantea* und *F. hepatica*-Proben fiel im unüberwachten LA (PCA) keine Separierung von *F. gigantea*-Proben je nach verwendetem

Körperanteil auf. Diese Ergebnisse wurden durch die erfolgreiche Identifikation der anterioren Anteile durch die MSPs der posterioren Anteile bestätigt. Dies legt nahe, dass es für die Identifikation unerheblich ist welcher Anteil einer adulten Probe zur Identifikation bzw. zur Erstellung einer Datenbank verwendet wird. Für den klinischen Alltag ist dieses Ergebnis vorteilhaft, da teilweise bei Extraktion der Proben Anteile fehlen können oder nicht verwertbar sind. Es wäre jedoch zur genaueren Untersuchung interessant, denselben Zeitpunkt zur Proteinextraktion der unterschiedlichen Körperanteile zu wählen, um einen eventuellen Einfluss der Lagerungsdauer auszuschließen. Ebenso sollte überprüft werden, ob der gleiche Körperanteil auch nach einer längeren Lagerungszeit eine ähnliche Identifikationsgenauigkeit ermöglicht.

Lagerungsmedium

Bei der Untersuchung der *Schistosoma*-Proben in Abhängigkeit von der verwendeten Aufbewahrungslösung (RNAlater® oder 70% (v/v) Ethanol) konnte eine unterschiedliche Intensität der Peaks beobachtet werden. Eine weitere Diskussion dieses Einflusses auf die Aufbewahrung der Proben ist im Unterpunkt 8.3. zu finden.

Geschlecht

Für getrenntgeschlechtliche Proben ist die Untersuchung von abweichenden Proteinprofilen nach Geschlecht des Isolats für nachfolgende Forschung wichtig. Es war in der Untersuchung der LA für *Schistosoma*-Proben möglich, eine Clusterung von männlichen, weiblichen und gemischtgeschlechtlichen Proben zu erkennen, jedoch ist es auf Grund der geringen Anzahl der Proben nicht möglich eine eindeutige Schlussfolgerung aus diesen Daten zu ziehen

Zusammenfassend ist zu sagen, dass vorrangig die Spezies entscheidend zur Differenzierung der Proben beitrug. Bei genauerer Betrachtung (Intraspezies-Klassifikation) und Anwendung überwachter LA ist aber auch eine klare Separierung der Entwicklungsstadien, und der verwendeten Lagerungsmedien festzustellen. Der verwendete Körperanteil und das Geschlecht scheinen den geringsten Einfluss auf die Proteinspektren zu haben.

8.2. *Praktische Anwendung der Ergebnisse*

2020 legte die WHO einen neuen Arbeitskatalog mit der Zielsetzung zur Bekämpfung und Eliminierung der 20 NTDs vor. Zu dieser Gruppe der Erkrankungen gehört auch die Schistosomiasis und die humane Fasziole. Dabei wurde von der WHO auch auf die Dringlichkeit der Entwicklung neuer Diagnostikmethoden hingewiesen. Die Implementierung einer neuen, günstigen und zuverlässigen Diagnostikmethode, wie sie bereits für Bakterien mittels MALDI-TOF MS zur Verfügung steht, könnte somit ein wichtiger Schritt in diese Richtung sein.

Es ist nun von besonderer Bedeutung eine Diagnostikmethode zu etablieren, welche breit verfügbar ist, eine einfache Durchführung, auch durch ungelertes Personal erlaubt und gleichzeitig kostengünstig ist, um auch in Ländern des globalen Südens eine Anwendung zu finden. MALDI-TOF MS ermöglicht eine speziesspezifische Identifikation von humanmedizinisch relevanten Helminthen wie z.B. *Schistosoma*, was von großem Vorteil ist, da sich sowohl das Symptomspektrum als auch die Verbreitung der unterschiedlichen Spezies unterscheiden können. Bei der Fasziole und der Schistosomiasis stellt sich allerdings das Problem der Hybridisation, da durch die teils ähnlichen Endemiegebiete Hybride zwischen den einzelnen Spezies entstehen können und somit eine eindeutige Differenzierung der Spezies deutlich erschwert ist. Auch für zukünftige Forschung ist eine speziesspezifische Unterscheidung verschiedener Proben von Bedeutung, da sich teilweise auch Behandlungsansätze und Krankheitsverläufe stark unterscheiden können [75][108].

Zur Validierung der Datenbank ist es wichtig, in zukünftigen Arbeiten die in-house-Datenbank mit Proben von Hybrid-Spezies und humanen Proben auf eine korrekte Identifizierung zu überprüfen. Eine sichere Identifizierung auf der speziesspezifischen Ebene ist nur möglich, wenn verschiedene Stämme und unterschiedliche Entnahmebedingungen der Isolate in der Datenbank als MSP vorhanden sind [59]. Dabei ist die Anzahl der verfügbaren Proteinspektren in der Datenbank zur Identifikation von neuen Proben sehr wichtig [87]. In dieser Arbeit waren sowohl die Proben, welche zur MSP-Erstellung genutzt wurden, als auch die Proben, welche zur Testung der Datenbank verwendet wurden, zum selben Zeitpunkt, am selben Standort und den gleichen Endwirten entnommen worden. Um die Anwendbarkeit der MSPs auch für die Routinediagnostik zu überprüfen, ist es wichtig, unterschiedliche Entnahmebedingungen und mögliche hybride Spezies auf ihren Einfluss der Identifizierungsrate zu untersuchen. Zudem ist eine Inkludierung von *S. haematobium* in die in-house-Datenbank für weiterführende Projekte von

großer Bedeutung, um alle hauptsächlich relevanten, humanpathogenen Spezies der *Schistosoma* abzudecken. Zudem wäre ein Detektieren von spezifischen Biomarkern vorteilhaft, welche möglichst unabhängig von genetischer Variabilität sind, um eine Übertragbarkeit der Ergebnisse auf unbekannte Proben zu gewährleisten. In dieser Arbeit wurden keine spezifischen Biomarker untersucht.

Für diese Arbeit wurden Referenzspektren von *Schistosoma* und *Fasciola* der in-house-Datenbank hinzugefügt und sind somit bis jetzt nur im Institut für Mikrobiologie in Homburg verfügbar. Die Erstellung einer in-house-Datenbank ist mit der MALDI-TOF Biotyper Software in einer einfachen Weise möglich. Wie bereits dargelegt, konnte eine gute Anwendbarkeit der in-house-Datenbank nachgewiesen werden. Eine komplette oder partielle Weitergabe dieser Datenbank ist zwar prinzipiell möglich, eine breite Anwendung wäre jedoch nur möglich, wenn Parasitenspektren auch in kommerziell verfügbare Datenbanken integriert und global zur Verfügung gestellt würden [77].

8.3. *Vergleich der eigenen Ergebnisse mit denen aus der Literatur*

MALDI-TOF MS

MALDI-TOF MS ist in den letzten Jahren zu einem der vielversprechendsten Diagnostikinstrumenten in der Mikrobiologie geworden. Auf Grund seiner schnellen und sehr spezifischen Identifikation von Bakterien und Pilzen findet MALDI-TOF MS bereits Anwendung in vielen Laboren. Die offizielle Bruker Database, ausschließlich von Bruker Daltonik (Bremen) vertrieben, beinhaltet > 4000 Referenzspektren von mehr als 3000 Arten aus 540 Gattungen von Bakterien und Pilzen[103].

In der Forschung zur Untersuchung von Parasiten mittels MALDI-TOF MS wurden in den letzten Jahren große Fortschritte erzielt und auch für andere Helminthenarten wie die Nematoden *Trichuris suis* [88], *Ascaris* spp., *Dirofilaria* spp. [77], *Trichnella* spp. [73] und *Anisakis* spp. [66] und den Zestoden *Taenia saginata* [108] wurde diese Technik weiterentwickelt. Dabei konnten in allen Studien erfolgreiche Ergebnisse erreicht werden und für *Anisakis* spp., *Ascaris* spp., *Dirofilaria* spp. und *Taenia saginata* auch spezifische Biomarker detektiert werden. In der Arbeit von Nagorny et al. zu *Ascaris* spp und *Dirofilaria* spp. konnte anhand der Proteinspektren die Annahme nahegelegt werden, dass die Art determinierenden Peaks im Bereich von 2 kDa bis 6 kDa lagen und die speziesspezifischen Peaks im sich im Bereich von 8 kDa bis 20 kDa unterschieden [77]. Zu ähnlichen Ergebnissen gelangten Rivero et al. und konnten die meisten Peaks in einem Bereich von 2 kDa bis 7 kDa detektieren [88]. In dieser Arbeit lagen die meisten Peaks der *Fasciola*- und *Schistosoma*-Proben in einem Bereich von 3 kDa bis 8 kDa. Allerdings besteht Unklarheit über eine Standardisierung des Analyseprotokolls und auch die Auswirkung verschiedener präanalytischer Verfahren ist noch nicht hinreichend geklärt.

Zur breiten Anwendung, auch in weiteren Bereichen der Mikrobiologie, ist eine stetige Weiterentwicklung und eine Behebung von Einschränkungen notwendig. So sind MS/MS-basierte Methoden aktuell noch aufwendig in Bezug auf Zeit, Aufwand und Fachwissen, die für die Probenvorbereitung, Analyse und Datenverarbeitung erforderlich sind. Jedoch konnte in einer Studie von Tan et.al. für eine Diagnostik mittels MALDI-TOF MS von Hefen und Bakterien eine zeitliche Einsparung von 1.45 Tagen im Vergleich zum Standardprotokoll zur Identifizierung erreicht werden [102]. In dieser Arbeit habe ich ein einfaches, leicht zu erlernendes Protokoll zur Probenvorbereitung verwendet, dies sollte auch das Ziel für zukünftige Proteinextraktionen sein. Grundsätzlich ist eine Protokollstandardisierung anzustreben, welche jedoch

aktuell auf Grund von Genvariationen der Proben oft noch durch hauseigene Standardarbeitsanweisungen ergänzt wird [25][22]. Ein weiterführendes Ziel ist es, an dieser Stelle auch andere Proben, wie Blut, Urin, Sputum ohne vorherige Aufbereitung direkt zur Diagnostik zu verwenden. Es ist in diesem Zusammenhang auch wichtig die Entwicklung einer benutzerfreundlicheren Software voranzutreiben, um den Umgang mit den gewonnenen Daten zu erleichtern. Die Beschaffung der Materialien zur Verarbeitung und Messung der Proben ist günstig, so konnten die Reagenzien- und Arbeitskosten in einer Studie zu Identifizierung von Bakterien und Hefen mit MALDI-TOF MS um 56.9% im Vergleich zum Standardprotokoll vermindert werden [102]. Jedoch ist die Anschaffung der hochauflösenden Instrumente auf Grund von hohen Kosten noch immer ein Problem [95]. Von zentraler Bedeutung zur Anwendung von MALDI-TOF MS ist eine dichtere und genauere Referenzdatenbank mit einer stabilen Taxonomie wichtig, wobei auch auf eine stetige Aktualisierung der Datenbank geachtet werden muss, um neue Arten oder neu auftretende Krankheitserreger miterfassen zu können [41]. Zu einem Problem können falsch klassifizierte Daten in der Referenzdatenbank werden. Die Fehlklassifizierungen entstehen beispielsweise durch Kreuzkontamination von Genomsequenzen oder Fehlern bei der Handhabung der Proben und der taxonomischen Charakterisierung [85]. Die Identifikation von neuen Proben kann ebenfalls erschwert sein, da Proteinprofile durch das Alter der Proben, die Anzahl der verwendeten Laserschüsse, die gemittelten Proteinspektren pro Messung, Personalschulungen etc. beeinflusst werden [25]. Die komplexe biologische Natur der Lebensformen der Parasiten und deren komplexe Lebenszyklen, aus denen die Proteinprobe für die MALDI-TOF MS-Analyse extrahiert werden, stellt einen weiteren Einflussfaktor auf die Proteinprofile dar und eine besondere Herausforderung im Gegensatz zur Diagnostik von Bakterien oder Pilzen [75]. Genetische Variationen wurden sowohl in Bakterien [59], Pilzen [7] als auch in neusten Studien bei Parasiten in der Anwendung von MALDI-TOF MS festgestellt. Dabei können genetische Variationen die Proteinprofile verändern [59]. Diese Variationen ließen sich in einer Arbeit von Sy et.al für *Fasciola*-Proben in einer Hauptkomponentenanalyse widerspiegeln [99]. An dieser Stelle könnten ML Algorithmen zur weiteren Untersuchung dieser Variationen vielversprechend sein. Auch geringere Unterschiede wie die unterschiedliche Lagerungslösung der *Schistosoma*-Proben oder der verwendete Körperanteil der adulten *F. gigantica*-Proben konnten in dieser Arbeit bereits bei der Untersuchung der LA des ML detektiert werden.

Lagerungsmedium

Die Lagerung von Parasitenproben ist schon Gegenstand anderer Arbeiten zur Diagnostik mit MALDI-TOF MS gewesen. Meist wurde Natriumchlorid (über *Anisakis* spp. [66], *Dirofilaria* spp. und *Ascaris* spp. [77] oder 70% (v/v) Ethanol (*Fasciola* spp. [99], *Cyathotominae* spp. [12] und *Läusen* [81]) verwendet. In einer kürzlich erschienen Arbeit von Wendel et al. wurde die Wirkung von vier verschiedenen Lagermedien (70% (v/v) Ethanol, Natriumchlorid 0,45 % (v/v), Flüssigchromatographie (LC) MS-grade Wasser und Formalin 37 % (v/v)) zur Identifizierung von *T. saginata*-Proglottiden untersucht. In dieser Studie wurde über äquivalente Ergebnisse in Bezug auf Spektrenprofile und Identifizierungsraten für alle Medien bis zu einer Lagerungsdauer von 24 Wochen berichtet. Ausgenommen davon war Formalin 37 % (v/v), da keine Spektren generiert werden konnten [108]. Mayer-Scholl et.al. untersuchten die Auswirkungen verschiedener Konservierungsbedingungen von *Trichinella* spp. bei anschließender Messung mit MALDI-TOF MS (Einfrieren der Proben und Konservierung in 70% (v/v) Ethanol). Dabei ergaben sich keine signifikanten Unterschiede bei den generierten Proteinspektren, lediglich von geringfügigen Abweichungen, welche die Identifizierungsrate, gemessen durch den LSV-Wert, aber nicht wesentlich beeinflussten, wurde berichtet [73]. Bei einer Studie zur Untersuchung von *Anthropoden* von Nebbak et al. konnte eine geringfügige Überlegenheit des Einfrierens gegenüber der Lagerung in 70% (v/v) Ethanol nachgewiesen werden [78]. Zur Anwendung in der Diagnostik in ländlicheren Gebieten ist eine Proteinextraktion aus frischen Proben oder eine Lagerung bei Raumtemperatur in günstigen Konservierungsmitteln wie RNAlater® anzustreben, da die Proben bis zu 2 Wochen auch bei tropischen Temperaturen (28-32 °C) erhalten und in gekühltem Zustand (4 °C) mindestens für 6 Monate konserviert werden können [107].

In dieser Arbeit konnten bei den *Schistosoma*-Proben unterschiedliche Peaks und eine unterschiedliche Peakintensität in Abhängigkeit von der verwendeten Aufbewahrungslösung (RNAlater® oder 70% (v/v) Ethanol) beobachtet werden. Wurde RNAlater® eingesetzt, war eine korrekte Identifizierung von 96,6% (29/30) der Proben möglich, wohingegen nur 35,7% (10/28) nach Lagerung in 70% (v/v) Ethanol korrekt identifiziert wurden. Sollte eine Probe zur Diagnostik oder Forschung versendet werden und daher über einen gewissen Zeitraum gelagert werden müssen, bietet sich nach unseren Ergebnissen für *Schistosoma*-Proben eher RNAlater® als 70% (v/v) Ethanol an, da eine bessere speziesspezifisch Identifizierungsrate festgestellt wurde. Dies entspricht der Arbeit von Saito et al. zur Konservierung von *Synechococcus*

WH8102 (ein marines Cyanobakterium), in welcher im Vergleich von 5 verschiedenen Lagerungsmedien (u.a. Ethanol und RNAlater®) eine Überlegenheit von RNAlater® gegenüber anderen Medien aufweist. Dies ist vermutlich auf eine hohe Salzkonzentration in RNAlater® zurückzuführen, welche eine sehr gute Proteinerhaltung bewirkt und wahrscheinlich ebenso die proteolytischen Aktivitäten in den Zellen hemmt [92]. In Zusammenschau mit vorrangegangenen Arbeiten scheint aber auch 70% (v/v) Ethanol und NaCl ein geeignetes Lagerungsmittel. Eine vergleichende Übersichtsarbeit aller gängigen Lagerungsmittel ist für weiterführende Forschung von Bedeutung. Dabei kann die Anwendung von Algorithmen des ML hilfreich sein. In dieser Arbeit konnten vielversprechende Ergebnisse dazu erreicht werden, da es für die *Schistosoma*-Proben möglich war, eine klare Differenzierung der Speichermedien in verschiedenen LA zu erkennen.

Maschinelles Lernen

Die Anwendung von klassifizierten ML-Algorithmen in der Mikrobiologie wurde in den letzten Jahren mehrfach erfolgreich durchgeführt. Neben der Erforschung von ML-Algorithmen zur Subtypisierung von Bakterienarten [13] oder zum Nachweis von Antibiotikaresistenzen [19][36], konnte auch eine Anwendbarkeit der LA zu Identifizierung von Parasiten nachgewiesen werden. Zur automatischen Identifizierung von *Monogeneans* (Plattwürmer, welche auf Kiemen und der Haut von Fischen leben) konnte eine Gesamtklassifizierungsgenauigkeit von 90 % unter Verwendung von K-nearest neighbor und eine Genauigkeit von 91,25 % bei der Kreuzvalidierung mit der Leave One Out (LOO)-Methode erreicht werden [113]. Bei der LOO-Methode handelt es sich um eine Kreuzvalidierung bei der nur eine Probe als Testset verwendet wird, während der Rest zum Trainieren des Modells dient.

In einer weiteren Studie wurde von Ali et al. die Unterscheidung von Mitgliedern der Fisch-Ektoparasiten-Gattung *Gyrodactylus* mittels ML-Algorithmen untersucht. Dabei lag besonders die Unterscheidung verschiedener LA im Fokus. Von 4 LA wurden auch KNN, als linearer LA und SVM als nicht linearer LA in den Vergleich gesetzt. Dabei konnte der KNN mit einer Genauigkeit von 98.53% etwas bessere Ergebnisse liefern als der SVM mit einer Genauigkeit von 95.89%. Zur Validierung wurde ebenfalls die 10-fache Kreuzvalidierung verwendet [5]. In meiner Arbeit waren ebenso nur sehr geringe Unterschiede zwischen den LA zu erkennen (F1-Werte zwischen 94.5% und 100%), wobei für *Fasciola* der LA SVM in den meisten Fällen leicht überlegen war und bei *Schistosoma*-Proben der LA RF.

Auch in der Diagnostik von Helmintheneiern in Stuhlproben wurden bereits ML-Algorithmen angewendet. So konnten diese anhand von Größe, Form und Beschaffenheit der Schale morphometrisch ihrer Art zugeordnet werden. In der 10-fachen Kreuzvalidierung wurden korrekte Klassifikationsraten von 86% bis 90% erreicht [112]. Zur taxonomischen Bewertung von *Strongyloides fuelleborni* und *Strongyloides stercoralis* konnte unter Verwendung der Nukleotidsequenzanalyse ebenfalls ein Nutzen zur Anwendung von ML-Algorithmen nachgewiesen werden [8].

In meiner Arbeit konnten diese positiven Ergebnisse bestätigt werden. Sowohl die Klassifizierung von Proteinspektren mit unbeaufsichtigten (PCA), als auch mit überwachten (SVM, PLS-DA, RF und KNN) ML-Algorithmen konnten die zuverlässige Unterscheidung der Spezies (*S. mansoni* und *S. japonicum*) bei den *Schistosoma*-Proben und die Unterscheidung der Spezies (*F. hepatica* und *F. gigantica*) und des Entwicklungsstadiums unter Beweis stellen. In der Analyse ist aufgefallen, dass die Spezies die wichtigste Determinante für die Massenspektren war, gefolgt von den Entwicklungsstadien und Lagerungsmedien, welche erst im Vergleich innerhalb der jeweiligen Spezies differenziert wurden. Dies deckt sich mit den Ergebnissen von Karger et al. zur Anwendung von MALDI-TOF MS bei Zecken [54].

Um einen breiteren Zugang zu ML-Modellen zu ermöglichen und damit die Anwendung von ML in der Mikrobiologie zunehmend in die diagnostische Routine zu implementieren, ist eine Weitergabe von validierten, hochqualitativen ML-Modellen wichtig. Beispielsweise wurde das ML-Modell in der Untersuchung antimikrobieller Resistenzen in einer einfach interpretierbaren Weise der Öffentlichkeit zur Verfügung gestellt [31]. Von besonderer Bedeutung ist, dass die ML-Modelle interpretierbar sind und klar ersichtlich ist an welcher Stelle Fehler auftreten und wie sie zu beheben sind, bevor neue Daten validiert werden [83].

Resistenztestungen

Zur Behandlung von Faziolosen wird eine einmalige Gabe von Triclabendazol angewendet und für die Behandlung der Schistosomiasis erfolgt nach WHO-Leitlinie eine einmalige Gabe von Praziquantel. Bei beiden Wirkstoffen handelt es sich um sehr wirksame Medikamente. Jedoch ist in den letzten Jahren die Resistenzbildung gegen diese Medikamente aufgrund von stark steigendem Gebrauch, auch in der Veterinärmedizin, immer mehr zu einem Problem geworden. Die WHO erklärte die Resistenztestung von *Schistosoma* gegen Praziquantel daher zu einem der zentralen Ziele zur Bekämpfung von NTDs (siehe Abb. 1). Mittels MALDI-TOF MS werden bereits erfolgreich Resistenztestungen gegen Antibiotika durchgeführt [93] und anhand der erhaltenen Spektren wird versucht, das Vorhandensein von Virulenzfaktoren oder die Resistenz

gegen antimikrobielle Mittel vorherzusagen indem spezifische Proteine oder präzise Biomarker identifiziert werden [79][21]. Auch eine Identifizierung parasitärer Resistenzdeterminanten mittels MALDI-TOF MS sowie eine Resistenztestung von Parasiten wäre für weitere Fortschritte von zentraler Bedeutung und sollte daher erforscht werden.

8.4. *Schlussfolgerung*

Diese Arbeit unterstreicht die Möglichkeiten einer praktischen Anwendung von MALDI-TOF MS als effizientes, schnelles und exaktes Werkzeug für eine zuverlässige Identifizierung und Differenzierung von humanpathogenen Trematoden-Arten. Dies ist die erste Arbeit, welche zur Identifizierung von adulten *Schistosoma*-Würmern und zur Untersuchung des Einflusses verschiedener Entwicklungsstadien von *Fasciola*-Spezies durchgeführt wurde. Hierfür war, in Ermangelung einer kommerziell erhältlichen Datenbank zur Parasitenidentifizierung mittels MALDI-TOF, die Erstellung und Validierung einer in-house-Datenbank notwendig. Dabei wurden ML-basierte Klassifizierungsalgorithmen als Vorhersagemodelle für die Speziesunterscheidung sowie hinsichtlich deren Einflusses von Lagerungsmedien und des Einflusses des Entwicklungsstadiums angewendet.

Es sollten dieser Arbeit weitere Studien angeschlossen werden, die vor allem eine größere Probenanzahl zur Abdeckung von Genvariationen und weiteren Einflussfaktoren, wie Lagerungslösung und Lagerungsdauer (frische Proben vs. gelagerte Proben), beinhalten sollte, um eine breit und umfangreich zu validierende Referenzdatenbank erstellen zu können. Zur Identifikation von *Schistosoma* sollten auch MSPs von *S. haematobium* der Datenbank hinzugefügt werden, um alle klinisch führenden humanpathogenen Spezies abzudecken. Schließlich sollten weitere Helminthen und Protozoen hinsichtlich ihrer Identifizierbarkeit mittels MALDI-TOF MS untersucht werden.

9. Literaturverzeichnis

Literatur gem. Vorgaben des Medizinischen Fakultät der UdS.

1. A N, AC W, I B, D R, P P, L A (2016) Standardization of sample homogenization for mosquito identification using an innovative proteomic tool based on protein profiling. *Proteomics* 16:3148–3160
2. Adela Valeroa M, Dolores Bargaesa M, Khoubbanea M, Artigasa P, Quesadaa C, Berinde L, Ubeira FM, Mezo M, Hernandez JL, Agramunt VH, Mas-Coma S (2016) Higher physiopathogenicity by *Fasciola gigantica* than by the genetically close *F. hepatica*: experimental long-term follow-up of biochemical markers. *Trans R Soc Trop Med Hyg* 110:55–66
3. Ahammed AS, S.M.A. A (2019) Immature *Fasciola gigantica*: Time-dependent ultrastructural changes following in vivo treatment with triclabendazole. *Acta Trop* 196:15–21
4. Alasaad S, Soriguer RC, Abu-Madi M, El Behairy A, Jowers MJ, Baños PD, Píriz A, Fickel J, Zhu XQ (2011) A TaqMan real-time PCR-based assay for the identification of *Fasciola* spp. *Vet Parasitol* 179:266–271
5. Ali R, Hussain A, Bron JE, Shinn AP (2012) The use of ASM feature extraction and machine learning for the discrimination of members of the fish ectoparasite genus *gyrodactylus*. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)* 7666 LNCS:256–263
6. Anahtar MN, Yang JH, Kanjilal S (2021) Applications of Machine Learning to the Problem of Antimicrobial Resistance: an Emerging Model for Translational Research.
7. Bader O (2013) MALDI-TOF-MS-based species identification and typing approaches in medical mycology. *Proteomics* 13:788–799
8. Barratt JLN, Sapp SGH (2020) Machine learning-based analyses support the existence of species complexes for *Strongyloides fuelleborni* and *Strongyloides stercoralis*. *Parasitology* 147:1184–1195

9. Bezerra FSM, Leal JKF, Sousa MS, Pinheiro MCC, Ramos AN, Silva-Moraes V, Katz N (2018) Evaluating a point-of-care circulating cathodic antigen test (POC-CCA) to detect *Schistosoma mansoni* infections in a low endemic area in north-eastern Brazil. *Acta Trop* 182:264–270
10. Booth M (2018) Climate Change and the Neglected Tropical Diseases. *Adv Parasitol* 100:39–126
11. Bosch F, Palmeirim Id MS, Ali SM, Ame SM, Id JH, Keiser Id J (2021) Diagnosis of soil-transmitted helminths using the Kato-Katz technique: What is the influence of stirring, storage time and storage temperature on stool sample egg counts?
12. Bredtmann CM, Krücken J, Murugaiyan J, Kuzmina T, von Samson-Himmelstjerna G (2017) Nematode species identification—current status, challenges and future perspectives for cyathostomins. *Front Cell Infect Microbiol* 7:1–8
13. Calderaro A, Buttrini M, Martinelli M, Farina B, Moro T, Montecchini S, Arcangeletti MC, Chezzi C, De Conto F (2021) Rapid Classification of *Clostridioides difficile* Strains Using MALDI-TOF MS Peak-Based Assay in Comparison with PCR-Ribotyping. *Microorganisms* 9:
14. Calvani NED, George SD, Windsor PA, Bush RD, Šlapeta J (2018) Comparison of early detection of *Fasciola hepatica* in experimentally infected Merino sheep by real-time PCR, coproantigen ELISA and sedimentation. *Vet Parasitol* 251:85–89
15. Calvani NED, Ichikawa-Seki M, Bush RD, Khounsy S, Šlapeta J (2020) Which species is in the faeces at a time of global livestock movements: single nucleotide polymorphism genotyping assays for the differentiation of *Fasciola* spp. *Int J Parasitol* 50:91–101
16. Calvani NED, Šlapeta J (2021) *Fasciola* Species Introgression: Just a Fluke or Something More? *Trends Parasitol* 37:25–34
17. Calvo-Urbano B, Léger E, Gabain I, De Dood CJ, Diouf D, Borlase A, Rudge JW, A M Corstjens PL, Sène M, Van Dam GJ, Walker M, Webster JP Sensitivity and specificity of human point-of-care 2 circulating cathodic antigen (POC-CCA) test in African 3 livestock for rapid diagnosis of schistosomiasis: a Bayesian 4 latent class analysis.

18. Candela A, Arroyo MJ, Sánchez-molleda Á, Méndez G, Quiroga L, Ruiz A, Cercenado E, Marín M, Muñoz P, Mancera L, Rodríguez-temporal D, Rodríguez-sánchez B (2022) Rapid and Reproducible MALDI-TOF-Based Method for the Detection of Vancomycin-Resistant *Enterococcus faecium* Using Classifying Algorithms. *Diagnostics* 2022, Vol 12, Page 328 12:328
19. Candela A, Zvezdanova ME, Arroyo MJ, Gema M, Mancera L, Rodríguez JG, Serra JL, Jim R, Mu P, Lozano I, Castro C, Escribano P (2022) Detection of azole resistance in *Aspergillus fumigatus* complex isolates using MALDI-TOF mass spectrometry nchez on behalf of the. 28:260–266
20. Carbonnelle E, Mesquita C, Bille E, Day N, Dauphin B, Beretti JL, Ferroni A, Gutmann L, Nassif X (2011) MALDI-TOF mass spectrometry tools for bacterial identification in clinical microbiology laboratory. *Clin Biochem* 44:104–109
21. Carbonnelle É, Nassif X (2011) Utilisation en routine du MALDI-TOF-MS pour l'identification des pathogènes en microbiologie médicale. *médecine/sciences* 27:882–888
22. Clark AE, Kaleta EJ, Arora A, Wolk DM (2013) Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry: a Fundamental Shift in the Routine Practice of Clinical Microbiology. *Clin Microbiol Rev* 26:547
23. Cringoli G, Rinaldi L, Maurelli MP, Utzinger J (2010) FLOTAC: new multivalent techniques for qualitative and quantitative copromicroscopic diagnosis of parasites in animals and humans. *Nat Protoc* 2010 53 5:503–515
24. Croxatto A, Prod'hom G, Greub G (2012) Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS Microbiol Rev* 36:380–407
25. Cuénod A, Foucault F, Pflüger V, Egli A (2021) Factors Associated With MALDI-TOF Mass Spectral Quality of Species Identification in Clinical Routine Diagnostics. *Front Cell Infect Microbiol* 11:
26. Daltonik GmbH B Revision 1 MALDI Biotarget 48-Instructions for Use.
27. Diarra AZ, Almeras L, Laroche M, Berenger J-M, Koné AK, Bocoum Z, Dabo A,

- Doumbo O, Raoult D, Parola P (2017) Molecular and MALDI-TOF identification of ticks and tick-associated bacteria in Mali. *PLoS Negl Trop Dis* 11:
28. Dietrich CF, Kabaalioglu A, Brunetti E, Richter J (2015) Fasciolosis. *Z Gastroenterol* 53:285–290
29. Domingo CYJ, Paraguison Alili R, Alvaran IM, Aquino RJ (2018) Prepatent RT-LAMP Detection of *Fasciola gigantica* in Snails (*Lymnaea* spp.) and Goats (*Capra hircus*) Targeting Cathepsin B3 gene. *J Vet Sci Med Diagn* 07:
30. Dr med. Matthias Vennemann (2016) Rural development – the underestimated health hazard? - www.rural21.com. URL: <https://www.rural21.com/english/search/detail/article/rural-development-the-underestimated-health-hazard.html>
31. Drouin A, Letarte G, Raymond F, Marchand M, Corbeil J, Laviolette F (2019) Interpretable genotype-to-phenotype classifiers with performance guarantees. *Sci Rep* 9:
32. Ebersbach JC, Sato MO, de Araújo MP, Sato M, Becker SL, Sy I (2023) Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for differential identification of adult *Schistosoma* worms. *Parasit Vectors* 16:20
33. Eckert, Johannes; Friedhoff KT (2008) 3.2 Helminthen(I). Thieme Verlag
34. Elliott DE (1996) Schistosomiasis. Pathophysiology, diagnosis, and treatment. *Gastroenterol Clin North Am* 25:599–625
35. Engels D, Zhou X-N Neglected tropical diseases: an effective global response to local poverty-related disease priorities.
36. Feucherolles M, Poppert S, Utzinger J, Becker SL (2019) MALDI-TOF mass spectrometry as a diagnostic tool in human and veterinary helminthology: A systematic review. *Parasites and Vectors* 12:1–13
37. Feucherolles M, Nennig M, Becker SL, Martiny D, Losch S, Penny C, Cauchie H, Ragimbeau C (2022) Combination of MALDI-TOF Mass Spectrometry and Machine Learning for Rapid Antimicrobial Resistance Screening : The Case of. *12:1–16*

38. Gaillot O, Blondiaux N, Loïez C, Wallet F, Lemaître N, Herwegh S, Courcol RJ (2011) Cost-effectiveness of switch to matrix-assisted laser desorption ionization-time of flight mass spectrometry for routine bacterial identification. *J Clin Microbiol* 49:4412
39. Gandhi P, Schmitt EK, Chen CW, Samantray S, Venishetty VK, Hughes D (2019) Triclabendazole in the treatment of human fascioliasis: a review. *Trans R Soc Trop Med Hyg* 113:797–804
40. Goodswen SJ, Barratt JLN, Kennedy PJ, Kaufer A, Calarco L, Ellis JT (2021) Machine learning and applications in microbiology. *FEMS Microbiol Rev* 45:1–19
41. Grenga L, Pible O, Armengaud J (2019) Pathogen proteotyping: A rapidly developing application of mass spectrometry to address clinical concerns.
42. Gryseels B, Polman K, Clerinx J, Kestens L (2006) Human schistosomiasis. *Lancet* 368:1106–1118
43. Gryseels B, Polman K, Clerinx J, Kestens L (2006) Human schistosomiasis. *Lancet* 368:1106–1118
44. Guegan H, Fillaux J, Charpentier E, Robert-Gangneux F, Chauvin P, Guemas E, Boissier J, Valentin A, Cassaing S, Gangneux JP, Berry A, Iriart X (2019) Real-time PCR for diagnosis of imported schistosomiasis. *PLoS Negl Trop Dis* 13:
45. Häslér B, Cornelsen L, Bennani H, Rushton J (2014) A review of the metrics for One Health benefits. *OIE Rev Sci Tech* 33:453–464
46. Hatz CFR (2001) The use of ultrasound in schistosomiasis. *Adv Parasitol* 48:225–284
47. Hof H, Schlüter D, Dörries R H (2019) *Duale Reihe Medizinische Mikrobiologie*. 7. edition. Thieme, Stuttgart
48. Huang Y, Li W, Lu W, Xiong C, Yang Y, Yan H, Liu KC, Cao P (2016) Cloning and in vitro characterization of a *Schistosoma japonicum* aquaglyceroporin that functions in osmoregulation. *Sci Rep* 6:1–8
49. Huguenin A, Depaquit J, Villena I, Ferte H (2019) MALDI-TOF mass spectrometry: A new tool for rapid identification of cercariae (Trematoda, Digenea). *Parasite* 26:

50. Jones ME, Mitchell RG, Leen CLS (1992) Long seronegative window in *schistosoma* infection. *Lancet* (London, England) 340:1549–1550
51. Kabaalioglu A, Çubuk M, Şenol U, Çevikol C, Karaali K, Apaydin A, Sindel T, Lüleci E (2000) Fascioliasis: US, CT, and MRI findings with new observations. *Abdom Imaging* 25:400–404
52. Kaewkes S (2003) Taxonomy and biology of liver flukes. *Acta Trop* 88:177–186
53. Karadjian G, Bilska-Zajac E, Bahn P, Py JS, John A, Gassilloud B, Rózycki M, Cencek T, Mayer-Scholl A, Vallée I (2020) Species identification of *Trichinella* originated from various host and different geographical location by MALDI-TOF. *Exp Parasitol* 213:107890
54. Karger A, Kampen H, Bettin B, Dautel H, Ziller M, Hoffmann B, Süß J, Klaus C (2012) Species determination and characterization of developmental stages of ticks by whole-animal matrix-assisted laser desorption/ionization mass spectrometry. *Ticks Tick Borne Dis* 3:78–89
55. Keiser J, Utzinger J (2005) Emerging foodborne trematodiasis. *Emerg Infect Dis* 11:1507–1514
56. Keiser J, Utzinger J (2009) Food-borne trematodiasis. *Clin Microbiol Rev* 22:466–483
57. Kokaliaris C, Garba A, Matuska M, Bronzan RN, Colley DG, Dorkenoo AM, Ekpo UF, Fleming FM, French MD, Kabore A, Mbonigaba JB, Midzi N, Mwinzi PNM, N’Goran EK, Polo MR, Sacko M, Tchuem Tchuente LA, Tukahebwa EM, Uvon PA, Yang G, Wiesner L, Zhang Y, Utzinger J, Vounatsou P (2022) Effect of preventive chemotherapy with praziquantel on schistosomiasis among school-aged children in sub-Saharan Africa: a spatiotemporal modelling study. *Lancet Infect Dis* 22:136–149
58. Laroche M, Almeras L, Pecchi E, Bechah Y, Raoult D, Viola A, Parola P (2017) MALDI-TOF MS as an innovative tool for detection of *Plasmodium* parasites in *Anopheles* mosquitoes. *Malar J* 16:1–10
59. Lartigue MF, Héry-Arnaud G, Haguenoer E, Domelier AS, Schmit PO, Van Der Meer-Marquet N, Lanotte P, Mereghetti L, Kostrzewa M, Quentin R (2009) Identification of

- Streptococcus agalactiae* isolates from various phylogenetic lineages by matrix-assisted laser desorption ionization-time of flight mass spectrometry. J Clin Microbiol 47:2284–2287
60. Le TH, Nguyen KT, Nguyen NTB, Doan HTT, Le XTK, Hoang CTM, Van De N (2012) Development and evaluation of a single-step duplex PCR for simultaneous detection of *Fasciola hepatica* and *Fasciola gigantica* (family Fasciolidae, class Trematoda, phylum Platyhelminthes). J Clin Microbiol 50:2720–2726
61. Leerapun A, Puasripun S, Kijdamrongthum P, Thongsawat S (2021) Human fascioliasis presenting as liver abscess: clinical characteristics and management. Hepatol Int 15:804–811
62. Lewis FA, Tucker MS (2014) Schistosomiasis. Adv Exp Med Biol 766:47–75
63. Lopez M, White AC, Cabada MM (2012) Burden of *Fasciola hepatica* infection among children from Paucartambo in Cusco, Peru. Am J Trop Med Hyg 86:481–485
64. Lovis L, Mak TK, Phongluxa K, Soukhathammavong P, Sayasone S, Akkhavong K, Odermatt P, Keiser J, Felger I (2009) PCR Diagnosis of *Opisthorchis viverrini* and *Haplorchis taichui* Infections in a Lao Community in an Area of Endemicity and Comparison of Diagnostic Methods for Parasitological Field Surveys. J Clin Microbiol 47:1517
65. Marcos L, Maco V, Terashima A (2020) Triclabendazole for the treatment of human fascioliasis and the threat of treatment failures. <https://doi.org/10.1080/1478721020211858798> 19:817–823
66. Marzano V, Pane S, Foglietta G, Mortera SL, Vernocchi P, Muda AO, Putignani L (2020) Mass spectrometry based-proteomic analysis of anisakis spp.: A preliminary study towards a new diagnostic tool. Genes (Basel) 11:1–18
67. Mas-coma S, Bargues MD, Valero MA (2014) Diagnosis of human fascioliasis by stool and blood techniques: update for the present global scenario.
68. Mas-Coma S (2005) Epidemiology of fascioliasis in human endemic areas. J Helminthol 79:207–216

69. Mas-Coma S, Valero MA, Bargues MD (2009) Chapter 2 Fasciola, *Lymnaeids* and Human Fascioliasis, with a Global Overview on Disease Transmission, Epidemiology, Evolutionary Genetics, Molecular Epidemiology and Control. *Adv Parasitol* 69:41–146
70. Mas-Coma S, Valero MA, Bargues MD (2009) Climate change effects on trematodiasis, with emphasis on zoonotic fascioliasis and schistosomiasis. *Vet Parasitol* 163:264–280
71. Mas-Coma S, Bargues MD, Valero MA (2018) Human fascioliasis infection sources, their diversity, incidence factors, analytical methods and prevention measures. *Parasitology* 145:1665–1699
72. Mas-Coma S, Valero MA, Bargues MD (2019) Fascioliasis. *Adv Exp Med Biol* 1154:71–103
73. Mayer-Scholl A, Murugaiyan J, Neumann J, Bahn P, Reckinger S, Nöckler K (2016) Rapid Identification of the Foodborne Pathogen *Trichinella* spp. by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *PLoS One* 11:152062
74. McManus DP, Dunne DW, Sacko M, Utzinger J, Vennervald BJ, Zhou XN (2018) Schistosomiasis. *Nat Rev Dis Prim* 2018 41 4:1–19
75. Mouri O, Morizot G, Van der Auwera G, Ravel C, Passet M, Chartrel N, Joly I, Thellier M, Jauréguiberry S, Caumes E, Mazier D, Marinach-Patrice C, Buffet P (2014) Easy identification of *leishmania* species by mass spectrometry. *PLoS Negl Trop Dis* 8:
76. Mucheka VT, Lamb JM, Pfukenyi DM, Mukaratirwa S (2015) DNA sequence analyses reveal co-occurrence of novel haplotypes of *Fasciola gigantica* with *F. hepatica* in South Africa and Zimbabwe. *Vet Parasitol* 214:144–151
77. Nagorny SA, Valentinovna Aleshukina A, Sergeevna Aleshukina I, Alexandrovna Ermakova L, Yurievna Pshenichnaya N (2019) The application of proteomic methods (MALDI-toff MS) for studying protein profiles of some nematodes (*dirofilaria* and *ascaris*) for differentiating species. *Int J Infect Dis* 82:61–65
78. Nebbak A, El Hamzaoui B, Berenger JM, Bitam I, Raoult D, Almeras L, Parola P (2017) Comparative analysis of storage conditions and homogenization methods for tick and flea species for identification by MALDI-TOF MS. *Med Vet Entomol* 31:438–448

79. Nordmann P (2010) Résistance aux carbapénèmes chez les bacilles à Gram négatif. *médecine/sciences* 26:950–959
80. Normand AC, Cassagne C, Gautier M, Becker P, Ranque S, Hendrickx M, Piarroux R (2017) Decision criteria for MALDI-TOF MS-based identification of filamentous fungi using commercial and in-house reference databases. *BMC Microbiol* 17:1–17
81. Ouarti B, Laroche M, Righi S, Meguini MN, Benakhla A, Raoult D, Parola P (2020) Development of MALDI-TOF mass spectrometry for the identification of lice isolated from farm animals. *Parasite* 27:
82. Patel R (2015) MALDI-TOF MS for the Diagnosis of Infectious Diseases. *Clin Chem* 61:100–111
83. Peiffer-Smadja N, Delli Ere S, Rodriguez C, Birgand G, Lescure F-X, Fourati S, Rupp E E (2020) Machine learning in the clinical microbiology laboratory: has the time come for routine practice *Clin Microbiol Infect* 26:1300–1309
84. Perrodin S, Walti L, Gottstein B, Kim-Fuchs C, Candinas D, Banz V (2019) Fasciola hepatica in a country of low incidence: a tricky diagnosis. *Hepatobiliary Surg Nutr* 8:597–603
85. Pible O, Armengaud J (2015) Improving the quality of genome, protein sequence, and taxonomy databases: A prerequisite for microbiome meta-omics 2.0. *Proteomics* 15:3418–3423
86. Rana MAA, Roohi N, Khan MA FASCIOLIASIS IN CATTLE-A REVIEW. *J Anim Plant Sci* 24:2014
87. Rettinger A, Krupka I, Grünwald K, Dyachenko V, Fingerle V, Konrad R, Raschel H, Busch U, Sing A, Straubinger RK, Huber I (2012) *Leptospira* spp. strain identification by MALDI TOF MS is an equivalent tool to 16S rRNA gene sequencing and multi locus sequence typing (MLST). *BMC Microbiol* 12:
88. Rivero J, Zurita A, Cutillas C, Callejón R (2022) The Use of MALDI-TOF MS as a Diagnostic Tool for Adult *Trichuris* Species. *Front Vet Sci* 9:
89. Ross AG, Vickers D, Olds GR, Shah SM, McManus DP (2007) Katayama syndrome.

- Lancet Infect Dis 7:218–224
90. Saba R, Korkmaz M, Inan D, Mamikoğlu L, Turhan Ö, Günseren F, Čevikol C, Kabaalioğlu A (2004) Human fascioliasis. Clin Microbiol Infect 10:385–387
 91. Saba R, Korkmaz M, Inan D, Mamikoğlu L, Turhan Ö, Günseren F, Čevikol C, Kabaalioğlu A (2004) Human fascioliasis. Clin Microbiol Infect 10:385–387
 92. Saito MA, Bulygin V V., Moran DM, Taylor C, Scholin C (2011) Examination of microbial proteome preservation techniques applicable to autonomous environmental sample collection. Front Microbiol 2:1–10
 93. Schubert S, Kostrzewa M (2017) MALDI-TOF MS in the Microbiology Laboratory: Current Trends. Curr Issues Mol Biol 2017, Vol 23, Pages 17-20 23:17–20
 94. Secor WE (2012) The effects of schistosomiasis on HIV/AIDS infection, progression and transmission. Curr Opin HIV AIDS 7:254–259
 95. Singhal N, Kumar M, Viridi JS (2016) MALDI-TOF MS in clinical parasitology: Applications, constraints and prospects. Parasitology 143:1491–1500
 96. Siqueira LMV, Senra C, de Oliveira AA, Carneiro NF de F, Gomes LI, Rabello A, Coelho PMZ, Oliveira E (2021) A Real-Time PCR Assay for the Diagnosis of Intestinal Schistosomiasis and Cure Assessment After the Treatment of Individuals With Low Parasite Burden. Front Immunol 11:
 97. Sören Schubert K, Kostrzewa M MALDI-TOF MS in Microbiology.
 98. Steinmann P, Keiser J, Bos R, Tanner M, Utzinger J (2006) Schistosomiasis and water resources development: systematic review, meta-analysis, and estimates of people at risk. Lancet Infect Dis 6:411–425
 99. Sy I, Margardt L, Ngbede EO, Adah MI, Yusuf ST, Keiser J, Rehner J, Utzinger J, Poppert S, Becker SL (2021) Identification of adult *fasciola* spp. Using matrix-assisted laser/desorption ionization time-of-flight (maldi-tof) mass spectrometry. Microorganisms 9:1–11
 100. Takeuchi-Storm N, Denwood M, Petersen HH, Enemark HL, Stensgaard AS, Sengupta

- ME, Beesley NJ, Hodgkinson J, Williams D, Thamsborg SM (2018) Patterns of *Fasciola hepatica* infection in Danish dairy cattle: Implications for on-farm control of the parasite based on different diagnostic methods. *Parasites and Vectors* 11:1–18
101. Talaie H, Emami H, Yadegarinia D, Nava-ocampo AA, Massoud J, Azmoudeh M, Mascoma S (2004) RANDOMIZED TRIAL OF A SINGLE , DOUBLE AND TRIPLE DOSE OF 10 MG / KG OF A HUMAN FORMULATION OF TRICLABENDAZOLE IN PATIENTS WITH FASCIOLIASIS. 777–782
102. Tan KE, Ellis BC, Lee R, Stamper PD, Zhang SX, Carroll KC (2012) Prospective Evaluation of a Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry System in a Hospital Clinical Microbiology Laboratory for Identification of Bacteria and Yeasts: a Bench-by-Bench Study for Assessing the Impact on Time to Identification and Cost-Effectiveness.
103. Torres-sangiao E, Leal Rodriguez C, García-riestra C (2021) Application and perspectives of maldi–tof mass spectrometry in clinical microbiology laboratories. *Microorganisms* 9:1–19
104. Utzinger J, Becker SL, Knopp S, Blum J, Neumayr AL, Keiser J, Hatz CF (2012) Neglected tropical diseases: diagnosis, clinical management, treatment and control. *Swiss Med Wkly* 2012 47 142:
105. Valero A, Hamed K, Villegas F, Grueninger H, Ault SK, Montresor A, Engels D, Mascoma S, Gabrielli F (2012) Administration of Triclabendazole Is Safe and Effective in Controlling Fascioliasis in an Endemic Community of the Bolivian Altiplano. 6:6–12
106. Webb CM, Cabada MM (2018) Recent developments in the epidemiology, diagnosis, and treatment of *Fasciola* infection. *Curr Opin Infect Dis* 31:409–414
107. Webster BL (2009) Isolation and preservation of *schistosome* eggs and larvae in RNAlater® facilitates genetic profiling of individuals.
108. Wendel TP, Feucherolles M, Rehner J, Poppert S, Utzinger J, Becker SL, Sy I (2021) Evaluating Different Storage Media for Identification of *Taenia saginata* Proglottids Using MALDI-TOF Mass Spectrometry. *Microorganisms* 9:

109. WHO (2020) Ending the neglect to attain the Sustainable Development Goals: a road map for neglected tropical diseases 2021–2030.
110. Wichmann D (2015) Diagnostik der Schistosomiasis (Bilharziose) – Notwendige Untersuchungen bei Reiserückkehrern und Patienten aus Endemiegebieten. *Flugmedizin · Tropenmedizin · Reisemedizin - FTR* 22:73–77
111. World Health Organization (2021) WHO model list of essential medicines - 22nd list, 2021. Tech Doc 2021
112. Yoon Seok Yang, Duck Kun Park, Hee Chan Kim, Choi MH, Chai JY (2001) Automatic identification of human helminth eggs on microscopic fecal specimens using digital image processing and an artificial neural network. *IEEE Trans Biomed Eng* 48:718–730
113. Yousef Kalafi E, Tan WB, Town C, Dhillon SK (2016) Automated identification of Monogeneans using digital image processing and K-nearest neighbour approaches. *BMC Bioinformatics* 17:
114. Z DA, Vlamincck J, Levecke B, Briones-Montero A, Geldhof P (2019) Comparison of Kato-Katz Thick Smear, Mini-FLOTAC, and Flukefinder for the Detection and Quantification of *Fasciola hepatica* Eggs in Artificially Spiked Human Stool. *Am J Trop Med Hyg* 101:59–61
115. Zou L, Ruan S (2015) Schistosomiasis transmission and control in China. *Acta Trop* 143:51–57
116. A ROADMAP FOR IMPLEMENTATION ACCELERATING WORK TO OVERCOME THE GLOBAL IMPACT OF NEGLECTED TROPICAL DISEASES A ROADMAP FOR IMPLEMENTATION, World Health Organization
117. CDC - Schistosomiasis - Resources for Health Professionals. URL: https://www.cdc.gov/parasites/schistosomiasis/health_professionals/index.html#dx,
118. Leitlinie: Diagnostik und Therapie der Schistosomiasis (Bilharziose) Ziele und Kontext., AWMF-Register Nr. 042/005, Klasse S1
119. Instructions for Use Bruker Bacterial Test Standard. Bruker Daltronic

120. (2008) Circulating Antigen. *Encycl Parasitol* 251–251

121. (2018) CDC - *Fasciola* - Biology. URL:
<https://www.cdc.gov/parasites/fasciola/biology.html>

10. *Publikation*

Ebersbach JC, Sato MO, de Araújo MP, Sato M, Becker SL, Sy I (2023) Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for differential identification of adult *Schistosoma* worms. **Parasit Vectors** 16:20. doi: 10.1186/s13071-022-05604-0.

11. Danksagung

Ich möchte mich an dieser Stelle bei allen Menschen bedanken, die mich in der Arbeit an meiner Dissertation unterstützt haben. Auf dem für mich neuen Feld des wissenschaftlichen Arbeitens, war die erfahrene, jederzeit hilfsbereite und wertschätzende Unterstützung meines Doktorvater Dr. med. Dr. phil. Sören Becker eine große Hilfe.

Ganz besonderen Dank möchte ich auch Issa Sy aussprechen, welcher mich bei jedem Schritt der Arbeit begleitet, beraten und motiviert hat und damit meinen Einstieg in das Thema unglaublich erleichtert hat. Auch allen Mitarbeitenden des Instituts für medizinische Mikrobiologie und Hygiene, Homburg möchte ich für Hilfsbereitschaft und den außerordentlich netten Umgang danken. Die Arbeit hat mit Ihnen allen viel Spaß gemacht.

Allen Menschen aus meinem privaten Umfeld möchte ich für ihre Unterstützung danken. Ganz besonders möchte ich diesen an Joanne Buinier, Yael Lohse und Elisabeth Nesor für ihre Durchsicht der Arbeit, andauernde Unterstützung und ihren Rückhalt richten. Ein besonderer Dank geht auch an Maximilian Emmerich für seine unermüdliche Geduld und Ermutigung. Und zu guter Letzt möchte ich meinen Eltern, Almut und Georg Ebersbach und meinem Bruder Enno danken für ihren lieben Zuspruch und ihre so hilfreiche Beratung.

12. Lebenslauf

Aus datenschutzrechtlichen Gründen wird der Lebenslauf in der elektronischen Fassung der Dissertation nicht veröffentlicht.

13. Anhang

10-fache Kreuzvalidierung: *F. gigantea* und *F. hepatica*

(A) <i>F. gigantea</i> (adult) vs. (Eier)			
Support Vector Machine (SVM)			
Tatsächliche Gruppe	Vorhergesagte Gruppe		
	<i>F. gigantea</i> (adult)	<i>F. gigantea</i> (Eier)	% Korrekt
<i>F. gigantea</i> (adult)	17 (TP)	0 (FN)	100 (Sensitivität)
<i>F. gigantea</i> (Eier)	0 (FP)	10 (TN)	100 (Spezifität)
	100% (PPV)	100% (NPV)	100 (Genauigkeit)
			100 (F1 Score)
Partial least square discriminant analysis (PLS-DA)			
Tatsächliche Gruppe	Vorhergesagte Gruppe		
	<i>F. gigantea</i> (adult)	<i>F. gigantea</i> (Eier)	% Korrekt
<i>F. gigantea</i> (adult)	17 (TP)	0 (FN)	100 (Sensitivität)
<i>F. gigantea</i> (Eier)	0 (FP)	10 (TN)	100 (Spezifität)
	100% (PPV)	100% (NPV)	100 (Genauigkeit)
			97.14 (F1 Score)
Random forest (RF)			
Tatsächliche Gruppe	Vorhergesagte Gruppe		
	<i>F. gigantea</i> (adult)	<i>F. gigantea</i> (Eier)	% Korrekt
<i>F. gigantea</i> (adult)	17 (TP)	0 (FN)	100 (Sensitivität)
<i>F. gigantea</i> (Eier)	1 (FP)	9 (TN)	90 (Spezifität)
	94.44% (PPV)	100% (NPV)	96.3 (Genauigkeit)
			97.14 (F1 Score)
K nearest neighbor (KNN)			
Tatsächliche Gruppe	Vorhergesagte Gruppe		
	<i>F. gigantea</i> (adult)	<i>F. gigantea</i> (Eier)	% Korrekt
<i>F. gigantea</i> (adult)	17 (TP)	0 (FN)	100 (Sensitivität)
<i>F. gigantea</i> (Eier)	1 (FP)	9 (TN)	90 (Spezifität)
	94.44 % (PPV)	100% (NPV)	96.3 (Genauigkeit)
			97.14 (F1 Score)
(B) <i>F. hepatica</i> (adult) vs. (Eier)			
Support Vector Machine (SVM)			
Tatsächliche Gruppe	Vorhergesagte Gruppe		
	<i>F. hepatica</i> (adult)	<i>F. hepatica</i> (Eier)	% Korrekt
<i>F. hepatica</i> (adult)	9 (TP)	1 (FN)	90 (Sensitivität)
<i>F. hepatica</i> (Eier)	1 (FP)	9 (TN)	90 (Spezifität)
	90% (PPV)	90% (NPV)	90 (Genauigkeit)
			90 (F1 Score)
Partial least square discriminant analysis (PLS-DA)			
Tatsächliche Gruppe	Vorhergesagte Gruppe		
	<i>F. hepatica</i> (adult)	<i>F. hepatica</i> (Eier)	% Korrekt
<i>F. hepatica</i> (adult)	9 (TP)	1 (FN)	90 (Sensitivität)
<i>F. hepatica</i> (Eier)	0 (FP)	10 (TN)	100 (Spezifität)
	100% (PPV)	90.91% (NPV)	95 (Genauigkeit)

Anhang

95.24 (F1 score)			
Random forest (RF)			
Vorhergesagte Gruppe			
Tatsächliche Gruppe	<i>F. hepatica</i> (adult)	<i>F. hepatica</i> (Eier)	% Korrekt
<i>F. hepatica</i> (adult)	9 (TP)	1 (FN)	90 (Sensitivität)
<i>F. hepatica</i> (Eier)	0 (FP)	10 (TN)	100 (Spezifität)
	100% (PPV)	90.91% (NPV)	95 (Genauigkeit)
			95.24 (F1 Score)
K nearest neighbor (KNN)			
Vorhergesagte Gruppe			
Tatsächliche Gruppe	<i>F. hepatica</i> (adult)	<i>F. hepatica</i> (Eier)	% Korrekt
<i>F. hepatica</i> (adult)	9 (TP)	1 (FN)	90 (Sensitivität)
<i>F. hepatica</i> (Eier)	0 (FP)	10 (TN)	100 (Spezifität)
	100% (PPV)	90.91% (NPV)	95 (Genauigkeit)
			95.24 (F1 Score)
C) <i>F.gigantica</i> vs. <i>F.hepatica</i>			
Support Vector Machine (SVM)			
Vorhergesagte Gruppe			
Tatsächliche Gruppe	<i>F. hepatica</i>	<i>F. gigantica</i>	% Korrekt
<i>F. hepatica</i>	10 (TP)	0 (FN)	100 (Sensitivität)
<i>F. gigantica</i>	0 (FP)	17 (TN)	100 (Spezifität)
	100% (PPV)	100% (NPV)	100 (Genauigkeit)
			100 (F1 Score)
Partial least square discriminant analysis (PLS-DA)			
Vorhergesagte Gruppe			
Tatsächliche Gruppe	<i>F. hepatica</i>	<i>F. gigantica</i>	% Korrekt
<i>F. hepatica</i>	9 (TP)	1 (FN)	90 (Sensitivität)
<i>F. gigantica</i>	0 (FP)	17 (TN)	100 (Spezifität)
	100% (PPV)	94.44% (NPV)	96.3 (Genauigkeit)
			94.74 (F1 Score)
Random forest (RF)			
Vorhergesagte Gruppe			
Tatsächliche Gruppe	<i>F. hepatica</i>	<i>F. gigantica</i>	% Korrekt
<i>F. hepatica</i>	10 (TP)	0 (FN)	100 (Sensitivität)
<i>F. gigantica</i>	0 (FP)	17 (TN)	100 (Spezifität)
	100% (PPV)	100% (NPV)	100 (Genauigkeit)
			100 (F1 Score)
K nearest neighbor (KNN)			
Vorhergesagte Gruppe			
Tatsächliche Gruppe	<i>F. hepatica</i>	<i>F. gigantica</i>	% Korrekt
<i>F. hepatica</i>	9 (TP)	1 (FN)	90 (Sensitivität)
<i>F. gigantica</i>	0 (FP)	17 (TN)	100 (Spezifität)
	100% (PPV)	94.44% (NPV)	96.3 (Genauigkeit)
			94.74 (F1 Score)

Tabelle 5: 10-fache Kreuzvalidierung *Fasciola* spp.

TP, wahr positiv; TN, wahr negativ, FP, falsch positiv; FN, falsch negativ; PPV, positive prädiktive Werte; NPV, negative prädiktive Werte.

10-fache Kreuzvalidierung: *S. mansoni* und *S. japonicum*

(A) <i>S. mansoni</i> vs <i>S. japonicum</i>			
Support Vector Machine (SVM)			
Tatsächliche Gruppe	Vorhergesagte Gruppe		
	<i>S. japonicum</i>	<i>S. mansoni</i>	% Korrekt
<i>S. japonicum</i>	22 (TP)	0 (FN)	100 (Sensitivität)
<i>S. mansoni</i>	2 (FP)	38 (TN)	95 (Spezifität)
	91.67 (PPV)	100 (NPV)	96.77 (Genauigkeit)
			95.7 (F1 Score)
Partial least square discriminant analysis (PLS-DA)			
Tatsächliche Gruppe	Vorhergesagte Gruppe		
	<i>S. japonicum</i>	<i>S. mansoni</i>	% Korrekt
<i>S. japonicum</i>	21 (TP)	1 (FN)	95.45 (Sensitivität)
<i>S. mansoni</i>	1 (FP)	39 (TN)	97.5 (Spezifität)
	95.45% (PPV)	97.5% (NPV)	96.77% (Genauigkeit)
			95.5 (F1 Score)
Random forest (RF)			
Tatsächliche Gruppe	Vorhergesagte Gruppe		
	<i>S. japonicum</i>	<i>S. mansoni</i>	% Korrekt
<i>S. japonicum</i>	22 (TP)	0 (FN)	100 (Sensitivität)
<i>S. mansoni</i>	0 (FP)	40 (TN)	100 (Spezifität)
	100 (PPV)	100 (NPV)	100 (Genauigkeit)
			100 (F1 Score)
K nearest neighbor (KNN)			
Tatsächliche Gruppe	Vorhergesagte Gruppe		
	<i>S. japonicum</i>	<i>S. mansoni</i>	% Korrekt
<i>S. japonicum</i>	21 (TP)	1 (FN)	100 (Sensitivität)
<i>S. mansoni</i>	1 (FP)	39 (TN)	100 (Spezifität)
	95.45% (PPV)	97.5% (NPV)	100 (Genauigkeit)
			95.5 (F1 Score)
(B) <i>S. mansoni</i>, Ethanol vs RNAlater®			
Support Vector Machine (SVM)			
Tatsächliche Gruppe	Vorhergesagte Gruppe		
	RNAlater®	Ethanol	% Korrekt
RNAlater®	20 (TP)	0 (FN)	100 (Sensitivität)
70% (v/v) Ethanol	0 (FP)	19 (TN)	100 (Spezifität)
	100% (PPV)	100% (NPV)	100 (Genauigkeit)
			97.6 (F1 Score)
Partial least square discriminant analysis (PLS-DA)			
Tatsächliche Gruppe	Vorhergesagte Gruppe		
	RNAlater®	Ethanol	% Korrekt
RNAlater®	20 (TP)	0 (FN)	100 (Sensitivität)
70% (v/v) Ethanol	0 (FP)	19 (TN)	100 (Spezifität)
	100% (PPV)	100% (NPV)	100 (Genauigkeit)
			100 (F1 Score)
Random forest (RF)			
Tatsächliche Gruppe	Vorhergesagte Gruppe		
	RNAlater®	Ethanol	% Korrekt
RNAlater®	20 (TP)	0 (FN)	100 (Sensitivität)
70% (v/v) Ethanol	0 (FP)	19 (TN)	100 (Spezifität)
	100% (PPV)	100% (NPV)	100 (Genauigkeit)
			100 (F1 Score)
K nearest neighbor (KNN)			
Tatsächliche Gruppe	Vorhergesagte Gruppe		
	RNAlater®	Ethanol	% Korrekt
RNAlater®	20 (TP)	0 (FN)	100 (Sensitivität)
70% (v/v) Ethanol	0 (FP)	19 (TN)	100 (Spezifität)
	100% (PPV)	100% (NPV)	100 (Genauigkeit)
			100 (F1 Score)

Anhang

Tatsächliche Gruppe	RNAlater®	Ethanol	% Korrekt
RNAlater®	20 (TP)	0 (FN)	100 (Sensitivität)
70% (v/v) Ethanol	0 (FP)	19 (TN)	100 (Spezifität)
	100% (PPV)	100% (NPV)	100 (Genauigkeit)
			100 (F1 Score)
C) <i>S. japonicum</i>, Ethanol vs RNAlater			
Support Vector Machine (SVM)			
Vorhergesagte Gruppe			
Tatsächliche Gruppe	RNAlater®	Ethanol	% Korrekt
RNAlater®	11 (TP)	0 (FN)	100 (Sensitivität)
70% (v/v) Ethanol	0 (FP)	11 (TN)	100 (Spezifität)
	100% (PPV)	100% (NPV)	100 (Genauigkeit)
			100 (F1 Score)
Partial least square discriminant analysis (PLS-DA)			
Vorhergesagte Gruppe			
Tatsächliche Gruppe	RNAlater®	Ethanol	% Korrekt
RNAlater®	11 (TP)	0 (FN)	100 (Sensitivität)
70% (v/v) Ethanol	0 (FP)	11 (TN)	100 (Spezifität)
	100% (PPV)	100% (NPV)	100 (Genauigkeit)
			100 (F1 Score)
Random forest (RF)			
Vorhergesagte Gruppe			
Tatsächliche Gruppe	RNAlater®	Ethanol	% Korrekt
RNAlater®	11 (TP)	0 (FN)	100 (Sensitivität)
70% (v/v) Ethanol	0 (FP)	11 (TN)	100 (Spezifität)
	100% (PPV)	100% (NPV)	100 (Genauigkeit)
			100 (F1 Score)
K nearest neighbor (KNN)			
Vorhergesagte Gruppe			
Tatsächliche Gruppe	RNAlater®	Ethanol	% Korrekt
RNAlater®	11 (TP)	0 (FN)	100 (Sensitivität)
70% (v/v) Ethanol	0 (FP)	11 (TN)	100 (Spezifität)
	100% (PPV)	100% (NPV)	100 (Genauigkeit)
			100 (F1 Score)

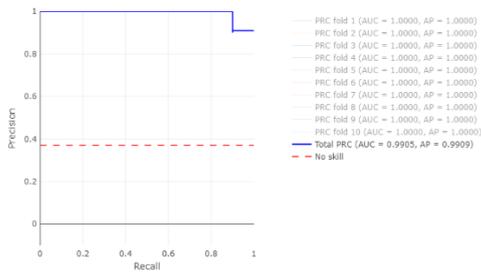
Tabelle 6: 10-fache Kreuzvalidierung *Schistosoma* spp.

TP, wahr positiv; TN, wahr negativ, FP, falsch positiv; FN, falsch negativ; PPV, positive prädiktive Werte; NPV, negative prädiktive Werte.

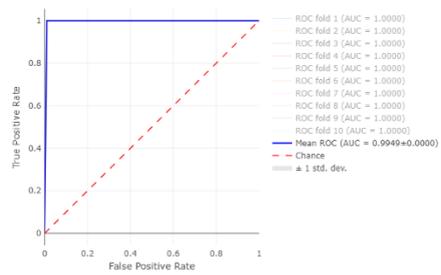
AUROC und ARCP - *Fasciola spp.*

1. PLS-DA

Precision-Recall Curve. Positive category: *Fasciola hepatica* adults

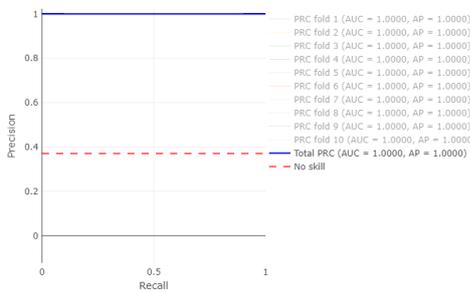


ROC Curve. Positive category: *Fasciola hepatica* adults

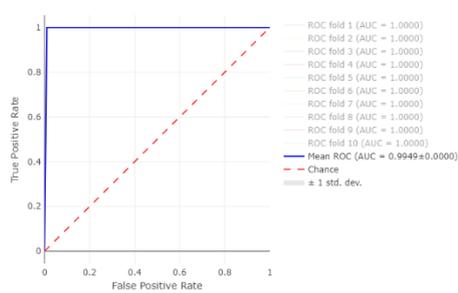


2. SVM

Precision-Recall Curve. Positive category: *Fasciola hepatica* adults

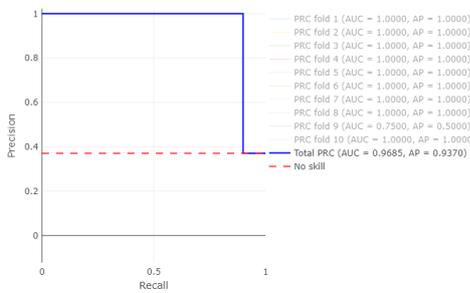


ROC Curve. Positive category: *Fasciola hepatica* adults

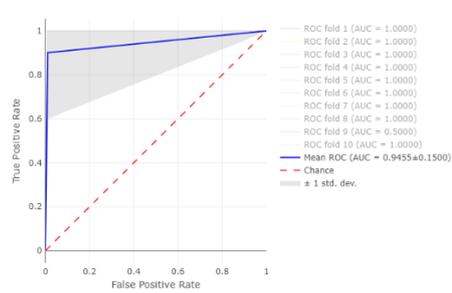


3. KNN

Precision-Recall Curve. Positive category: *Fasciola hepatica* adults

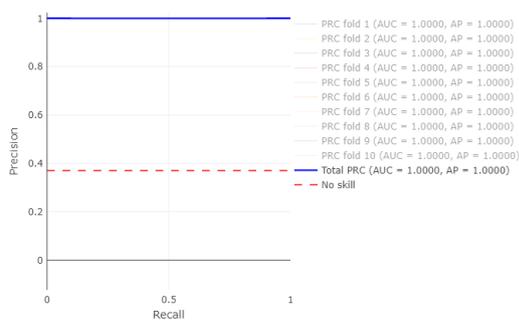


ROC Curve. Positive category: *Fasciola hepatica* adults



4. RF

Precision-Recall Curve. Positive category: *Fasciola hepatica* adults



ROC Curve. Positive category: *Fasciola hepatica* adults

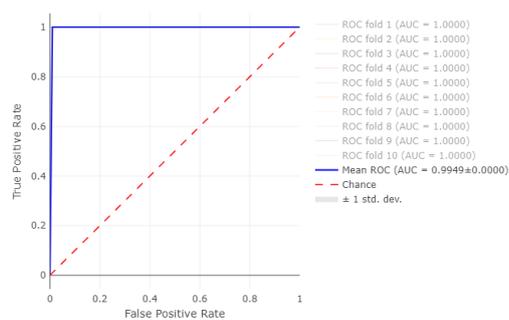


Abbildung 32: AUROC und ARCP – *Fasciola spp.*

AUROC und ARCP – *Schistosoma* spp.

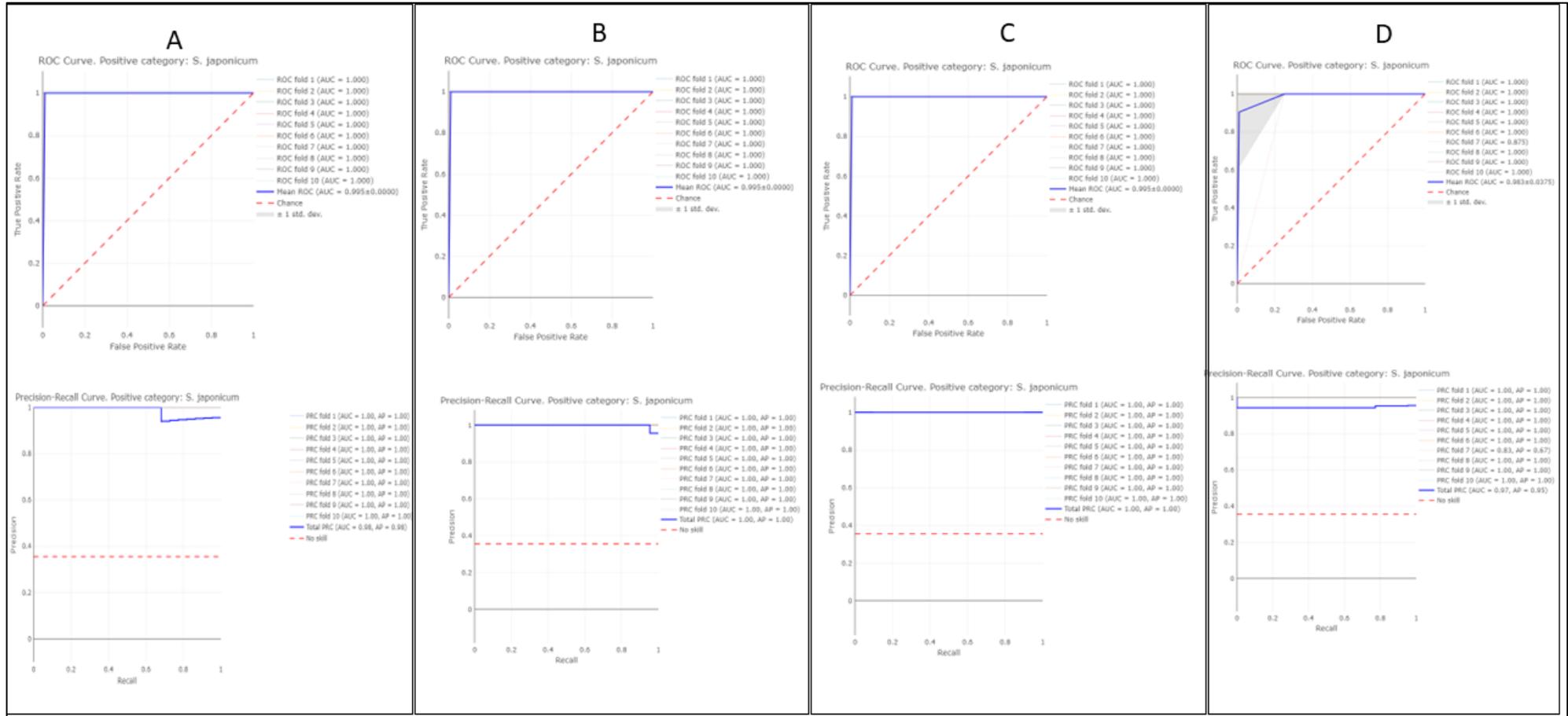


Abbildung 33: AUROC und ARCP – *Schistosoma* spp.: A - PLS-DA, B – SVM, C – KNN, D - RF

14. Erklärung an Eides statt

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise unentgeltlich geholfen: Issa Sy (Auswertung und Interpretation der Ergebnisse), Dr. Dr. Sören Becker (Interpretation der Ergebnisse).

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberaterinnen/Promotionsberater oder anderer Personen) in Anspruch genommen. Außer den Angegebenen hat niemand von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder in ähnlicher Form in einem anderen Verfahren zur Erlangung des Doktorgrades einer anderen Prüfungsbehörde vorgelegt. Ich versichere an Eides statt, dass ich nach bestem Wissen die Wahrheit gesagt und nichts verschwiegen habe. Die Bedeutung der eidesstattlichen Erklärung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Erklärung sind mir bekannt.

Ort, Datum

Unterschrift