

Aus dem Institut für Mikrobiologie und Infektionsimmunologie  
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Anwendung aktueller Typisierungsverfahren zur Identifikation  
epidemiologisch verbundener *Giardia duodenalis* Isolate  
Identification of epidemiologically linked *Giardia duodenalis* isolates  
using current typing procedures

zur Erlangung des akademischen Grades  
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät  
Charité – Universitätsmedizin Berlin

von

Andreas Woschke

Datum der Promotion: 30.11.2023

## Inhaltsverzeichnis

<b>Abkürzungsverzeichnis</b> .....	<b>3</b>
<b>Zusammenfassung</b> .....	<b>4</b>
<b>Abstract</b> .....	<b>5</b>
<b>1. Einleitung</b> .....	<b>6</b>
<b>2. Material und Methodik</b> .....	<b>11</b>
<b>2.1 Einverständnis der Ethikkommission</b> .....	<b>11</b>
<b>2.2 Arbeitsablauf</b> .....	<b>11</b>
<b>2.3 Probensammlung</b> .....	<b>11</b>
<b>2.4 DNA-Extraktion</b> .....	<b>12</b>
<b>2.5 Real-time PCR</b> .....	<b>12</b>
<b>2.6 MLST der G. duodenalis Isolate</b> .....	<b>13</b>
<b>2.7 MLST-Schema für Assemblage A</b> .....	<b>14</b>
<b>2.8 Analyse der DNA-Sequenzen</b> .....	<b>15</b>
<b>2.9 Testansatz zur Identifizierung epidemiologisch verbundener G. duodenalis Isolate</b> .....	<b>16</b>
<b>2.10 Analyse mit phylogenetischen Werkzeugen</b> .....	<b>17</b>
<b>2.11 Statistische Analyse</b> .....	<b>17</b>
<b>2.12 Zugangsnummern der DNA-Sequenzen</b> .....	<b>17</b>
<b>3. Ergebnisse</b> .....	<b>18</b>
<b>4. Diskussion</b> .....	<b>20</b>
<b>Literaturverzeichnis</b> .....	<b>26</b>
<b>Eidesstattliche Versicherung</b> .....	<b>33</b>
<b>Anteilerklärung an den erfolgten Publikationen</b> .....	<b>34</b>
<b>Auszug aus der Journal Summary List</b> .....	<b>35</b>
<b>Druckexemplar der Publikation</b> .....	<b>37</b>
<b>Lebenslauf</b> .....	<b>52</b>
<b>Komplette Publikationsliste</b> .....	<b>53</b>
<b>Danksagung</b> .....	<b>54</b>

## Abkürzungsverzeichnis

Die Begriffe sind in alphabetischer Reihenfolge gelistet.

Abkürzung	Begriff
ASH	Allelische Sequenzheterozygotie, engl. <i>allelic sequence heterozygosity</i>
°C	Grad Celsius
BG	Beta-giardin
BP	Basenpaare
CATH	Cathepsin L precursor
CID1	Caffeine-induced death protein 1-like protein
ct	Zyklusschwellwert, engl. <i>cycle threshold</i>
DALY	Verlorene gesunde Lebensjahre, engl. <i>disability adjusted life years</i>
DIS3	Mitotic Control Protein Dis3
DNA	Desoxyribonucleinsäure, engl. <i>deoxyribonucleic acid</i>
dNTP	Desoxyribonukleosidtriphosphate
ELISA	Enzyme-linked Immunosorbent Assay
g	G-Zahl (Beschleunigung der Zentrifuge)
GDH	Glutamate dehydrogenase
HCMP22547	High Cystein Membrane Protein Group 2
HCMP6372	High Cystein Protein
IAC	Interne Amplifikationskontrolle
IUPAC	The International Union of Pure and Applied Chemistry
M	Molar, Mol/l
MgCl <sub>2</sub>	Magnesiumchlorid
Min	Minute
MLST	Multi-Locus-Sequenztypisierungsverfahren, engl. <i>multiple locus sequence typing</i>
n	Anzahl der Chromosomensätze
NEK15411	NEK Kinase
NGS	Sequenzierung der nächsten Generation, engl. <i>next generation sequencing</i>
PCR	Polymerase-Kettenreaktion, engl. <i>polymerase chain reaction</i>

qPCR	Quantitative Echtzeit-Polymerase-Kettenreaktion, engl. <i>real time quantitative polymerase chain reaction</i>
RHP26	DNA repair and recombination protein Rhp26p
RKI	Robert Koch-Institut
SD	Standardabweichungen, engl. <i>standard deviation</i>
SNP	Einzelne Nucleotide betreffenden Polymorphismen, engl. <i>single nucleotide polymorphisms</i>
SSU-rRNA	Small subunit rRNA
TIF	Translation Initiation Factor
TPI	Triosephosphate isomerase
U	Einheiten, engl. units

## Zusammenfassung

**Hintergrund:** Der Parasit *Giardia duodenalis* verursacht Durchfallerkrankungen im Menschen und stellt weltweit ein relevantes Risiko für die öffentliche Gesundheit dar. Geschätzt erkranken jedes Jahr 184 Millionen Menschen an Giardiasis. Vorangegangene Studien haben acht unterschiedliche genetische Gruppen identifiziert, die Assemblage A-H genannt werden. Davon sind A und B sicher humanpathogen. Verbreitete Genotypisierungsverfahren besitzen eine niedrige Auflösung und werden durch allelische Sequenzheterozygotie (ASH) - vor allem bei Assemblage B -, sowie dem Auftreten von gemischten Infektionen beeinträchtigt. Primärziel dieses Pilotprojekts war die Evaluation aktuell zur Verfügung stehender Multi-Locus-Sequenztypisierungsverfahren (MLST) in Bezug auf deren Nutzen für epidemiologische Zwecke wie der molekularen Nachverfolgung von Infektionsketten, Infektionsquellen und Ausbruchsgeschehen. Sekundärziel war die Untersuchung der Transmissionsdynamik von *G. duodenalis* in Deutschland.

**Methodik:** Die menschlichen Stuhlproben für die zwei Datensätze stammten entweder aus einer Klinik für Tropenmedizin oder von vier Diagnostiklaboren. Mittels qPCR wurden diese auf A-, B-, und gemischte A + B Infektionen untersucht (Genorte *TIF*, *CATH*) und durch MLST auf den Genorten *TPI*, *BG*, und *GDH* charakterisiert. Für Assemblage A konnten mithilfe von kürzlich veröffentlichten MLST weitere sechs Genorte (*HCMP22547*, *CID1*, *RHP26*, *HCMP6372*, *DIS3*, *NEK15411*) untersucht werden.

**Ergebnisse:** Der Assemblage-Typ von *G. duodenalis* konnte in 175/202 (86,6 %) Personen identifiziert werden: Die Mehrheit gehörte zu Assemblage B 115/175 (65,7 %), gefolgt von A + B gemischten Infektionen 35/175 (20,0 %) und Assemblage A 25/175 (14,3 %). Mithilfe der beobachteten Rate der gemischten A + B Infektionen ließ sich die Anzahl der multiklonalen Infektionen für Assemblage A und B auf 25 % bzw. 50 % schätzen. Die MLST Analyse erfolgte inklusive ASH mittels einer neu entwickelten reziproken Ausreißeranalyse und konnte für Assemblage B im ersten und zweiten Datensatz jeweils 6/9 (66,7 %) und 4/5 (80,0 %) der Proben von chronisch infizierten Erkrankten richtig zuordnen. Zusätzlich wurde ein Cluster von fünf verschiedenen, mit Assemblage B infizierten, Personen identifiziert, der ein identisches MLST-Profil aufwies. Das kürzlich veröffentlichte MLST Schema für Assemblage A erhöhte die Sensitivität der Zuordnung von Stuhlproben chronisch infizierter Erkrankter. Es konnten 5/6 (83,3 %) Personen richtig identifiziert werden. Zusätzlich wurden 15 neue Genotypen charakterisiert.

**Schlussfolgerung:** Aktuell verfügbare Typisierungsverfahren lassen sich für epidemiologische Fragestellungen gewinnbringend nutzen. Assemblage A und B benötigen dabei verschiedene Protokolle und Auswertungsstrategien. Bei der Interpretation von MLST-Ergebnissen sollten polyklonale Infektionen und ASH berücksichtigt werden. Dieses Paper sensibilisiert für diese Herausforderungen und zeigt einen Lösungsansatz auf.

## Abstract

**Background:** The protozoan parasite *Giardia duodenalis* causes diarrheal disease in humans and represents a relevant public health risk worldwide. An estimated number of 184 million symptomatic cases are recognised globally per year. Previous studies identified eight distinct genetic groups (referred to as assemblage A-H) of which A and B are pathogenic to humans. Combined molecular and epidemiological data are largely lacking, partly due to low discriminatory power of current genotyping techniques, which is impeded by allelic sequence heterozygosity (ASH) mainly in assemblage B, and by occurrence of mixed infections. This pilot project assessed the suitability of current genotyping protocols of *G. duodenalis* assemblages A and B for epidemiological applications, such as molecular tracing of transmission chains, source attribution and outbreak investigations. A secondary objective was to investigate transmission dynamics of *G. duodenalis* in Germany.

**Methods:** Two sets of human stool samples, one from an outpatient tropical medicine clinic and the other from four primary care laboratories, were collected. Specimens were evaluated by real-

time PCRs to detect mixed infections (*TIF*, *CATH* gene loci) and characterised using a common multi locus sequence typing (MLST) scheme of *TPI*, *BG*, *GDH* gene loci. Additionally, six recently published loci for assemblage A (*HCMP22547*, *CID1*, *RHP26*, *HCMP6372*, *DIS3*, *NEK15411*) were investigated. To improve the analysis of common MLST data for assemblage B, a reciprocal outlier analysis was newly developed.

**Results:** The *G. duodenalis* assemblage type could be identified in 175/202 (86.6 %) persons: The majority belonged to assemblage B 115/175 (65.7 %), followed by A + B mixed infections 35/175 (20.0 %) and assemblage A 25/175 (14.3 %). Using the observed A + B mixed infection rate, we estimated that 25 % and 50 % of assemblage A or B cases represented multiple strain infections. Typing results were concordant with that prediction. Including allelic sequence heterozygosity into MLST investigation and using a newly developed reciprocal outlier analysis, we could correctly identify 6/9 (66.7 %) and 4/5 (80.0 %) subsequent samples from chronic assemblage B infected patients in the two sample collections, respectively. Additionally, we detected a cluster of five patients carrying independently assemblage B type *G. duodenalis* of an identical MLST type. Recently developed six gene MLST for assemblage A greatly increased sensitivity and identified 5/6 (83.3 %) subsequent samples from chronic assemblage A infected patients. Furthermore, we discovered 15 novel genotypes.

**Conclusion:** Considering different typing protocols and analytic tools for assemblages A and B, implementation of MLST for epidemiological purposes is possible. Polyclonal infections and ASH should be incorporated in MLST sequence analysis. This study sensitizes regarding these challenges and presents an approach to solution.

## 1. Einleitung

Der Eukaryot *Giardia duodenalis* (*G. duodenalis*, Syn: *G. intestinalis*, Syn: *G. lamblia*) ist ein einzelliger Parasit, der bei einer Vielzahl von Säugetieren wie Menschen, Haus-, Nutz- und Wildtieren gastrointestinale Beschwerden verursachen kann, die als Giardiasis (Syn: Lambliasis) zusammengefasst werden (1-3). *G. duodenalis* wird als relevante Gefahr für die öffentliche Gesundheit eingeschätzt, da der Parasit weltweit eine der häufigsten Ursachen für Durchfallerkrankungen darstellt (4, 5). Laut neueren Schätzungen infizieren sich jährlich mehr als 184 Millionen Menschen symptomatisch (4), was ca. 171.000 verlorene gesunde Lebensjahre (engl. *disability adjusted life years*, *DALYs*) zur Folge hat (6, 7). Neben den symptomatischen Infektionen verlaufen – vor allem in Endemiegebieten - ca. 50 % der Infektionen asymptomatisch

(2, 8, 9). Zusätzlich zu der zuvor beschriebenen Morbidität entsteht auch ein substanzieller ökonomischer Schaden durch Infektionen bei Nutztieren, spezielle Maßnahmen zur Aufbereitung von Wasser, Hospitalisierung und Medikamentenkosten (2, 10-14). Je nach untersuchter Region, den Lebensbedingungen und genutzten Methoden zur Diagnostik von *G. duodenalis* Infektionen unterscheidet sich die Prävalenz beim Menschen deutlich: Schätzungen zur Folge sind 10-40 % der Bevölkerung in „Entwicklungsländern“ mit dem Parasiten infiziert, wohingegen in „Industrienationen“ die Prävalenz zwischen 0,5–7 % schwankt (15-20). Studien, die das Erforschen von Risikofaktoren für eine Infektion mit *G. duodenalis* zum Ziel hatten, haben widersprüchliche Ergebnisse im Hinblick auf die Relevanz von Kontakt zu Tieren (Haus-, Nutz-, und Wildtiere) ergeben (15-18). Eindeutiger waren die Daten jedoch hinsichtlich der Übertragung von Mensch-zu-Mensch: Hier wurde der Kontakt mit kleinen Kindern, die eine Kinder-Tagesbetreuung besuchten, Männer, die Geschlechtsverkehr mit Männern hatten und vor allem Reisen als Risikofaktoren identifiziert (15-18). Studien aus Deutschland, den USA, Kanada und Neuseeland haben gezeigt, dass sich 10-55 % der an Giardiasis Erkrankten auf Reisen infizieren, während bei den anderen 45-90 % eine autochthone Infektion (Infektionsquelle wahrscheinlich nicht im Ausland, kein bekannter Kontakt zu aus dem Ausland zurückgekehrten Erkrankten) angenommen wurde (17-20). Aufgrund der Meldedaten in Deutschland geht man davon aus, dass aktuell ca. 50 % der Infektionen im Ausland erworben werden, während die anderen 50 % als autochthone Fälle gelten (21). Zur Diagnose von *G. duodenalis* stehen verschiedene Verfahren zur Verfügung: Die Spannweite reicht vom Nachweis von Trophozoiten und/oder Zysten des Erregers mittels Mikroskopie frischer Stuhlproben, über Antigennachweise, z.B. mittels „*Enzyme-linked Immunosorbent Assays*“ (ELISAs), bis hin zu verschiedenen PCR-Methoden, die sich allerdings alle hinsichtlich Sensitivität und Spezifität unterscheiden, was bei der Vergleichbarkeit der jeweiligen Daten berücksichtigt werden muss (22). Diese Methoden, die in der Routinediagnostik zur Anwendung kommen, unterscheiden allerdings nicht zwischen den verschiedenen Genotypen (Assemblagen) von *G. duodenalis*. Dadurch ist die molekulare Überwachung des durch *G. duodenalis* hervorgerufenen Infektionsgeschehens deutlich erschwert. Krankheitsausbrüche, potenzielle Infektionsquellen und Infektionsketten werden schlecht und wenn, dann spät erkannt, was die Nachverfolgung und Möglichkeiten der Krankheitsprävention einschränkt (23-25). Gegenwärtig werden acht verschiedenen Assemblagen von *G. duodenalis* unterschieden. Da es keine morphologischen Unterschiede zwischen den Assemblagen gibt, beruht die Einteilung vor allem auf der Amplifizierung mittels PCR und Sequenzanalyse verschiedener genetischer Loci. Ziele sind dabei in mehreren Kopien vorliegende Gene wie die „*small subunit rRNA*“ (*SSU-rRNA*), aber auch nur als einzelne Kopie vorkommende „Haushaltsgene“ wie „*triosephosphate isomerase*“

(*TPI*), „*beta-giardin*“ (*BG*) und „*glutamate dehydrogenase*“ (*GDH*) (23, 26). Die verschiedenen Assemblagen weisen verschiedene Wirtsspezifitäten auf: Assemblagen A und B sind humanpathogen, infizieren allerdings ebenfalls eine große Spannweite an anderen Säugetieren (23). Diesen beiden Assemblagen wird deshalb häufig ein zoonotisches Potential zugeschrieben (27-29). Bei den anderen sechs Assemblagen ist eine Humanpathogenität umstritten, da diese nur selten im Menschen nachweisbar sind und ein relativ enges Wirtsspektrum besitzen: Assemblagen C und D werden vor allem in Hunden (*Canidae*) gefunden, Assemblage E kommt vor allem in Huftieren vor (gelegentlich wird Assemblage E auch in Menschen beschrieben (9)), Assemblage F in Katzen (*Felidae*), Assemblage G in Nagetieren und Assemblage H in Robben (*Pinnipedia*) (23). Da sich die einzelnen Assemblagen nicht nur hinsichtlich ihres Wirtsspektrums unterscheiden, sondern auch beträchtliche genetische Differenzen vorliegen, wurde vorgeschlagen die Assemblagen als separate Spezies zu betrachten (23, 26). In dieser Arbeit wird die weit verbreitete Nomenklatur der „Assemblagen“ verwendet, da die Diskussion um die Phylogenie noch nicht abgeschlossen ist.

Möchte man die Assemblagen genetisch charakterisieren, so stößt man in der Literatur auf eine große Anzahl verschiedener Zielgene und Kombinationen von Zielgenen (Genotypisierungsschemata), die in der Regel alle geeignet sind, die Assemblagen voneinander zu unterscheiden (3, 23, 24, 30-34). Eine standardisierte, allgemein gebräuchliche Kombination von Zielgenen existiert jedoch bislang nicht und auch die gebräuchliche Kombination von *TPI*, *BG* und *GDH* besitzt - je nach Assemblagentyp - nur eine geringe Auflösung, wenn die einzelnen Assemblagen in Sub-Assemblagen oder Subgruppen unterteilt werden sollen (10, 24, 29, 34). Die Subtypisierung von Isolaten wird auch dadurch erschwert, dass sich *G. duodenalis* lediglich ineffizient kultivieren lässt. Daneben wäre auch die Analyse einzelner Zysten ein Verfahren, mit dem die Trennschärfe erhöht werden könnte, da polyklonale Infektionen mit verschiedenen Sub-Assemblagen die Multi-Locus-Sequenztypisierung (engl. *multi locus sequence typing*, MLST) Ergebnisse nicht negativ beeinflussen könnten. Das ist aber praktisch kaum umsetzbar, wenn man schnell Ausbruchsgeschehen oder Transmissionsketten untersuchen möchte (35, 36).

Sämtliche Genotypisierungsverfahren sind somit auf die Nutzung von DNA angewiesen, die direkt aus Stuhlproben oder aus vorher angereicherten Zysten gewonnen wird. Das hat zur Folge, dass Ko-Infektionen mit verschiedenen Assemblagen oder verschiedenen Subgruppen die Analyse der Gensequenzen verkomplizieren (29). Trotz dieser Problematik wird DNA, die direkt aus Stuhlproben gewonnen wird, gewöhnlich als „Isolat“ bezeichnet, da die Kultivierung für die routinemäßige Molekularepidemiologie keine Rolle spielt (29). Da *G. duodenalis* weltweit für zahlreiche Ausbrüche über verschiedene Infektionswege – Trinkwasser, Lebensmittel, direkte



Mensch-zu-Mensch Übertragung und auch zoonotische Tier-zu-Mensch Infektionen – verantwortlich ist, benötigt man zur sicheren Identifikation von Übertragungsketten zuverlässige Genotypisierungsschema (37-39). Wie bereits erwähnt werden dafür zumeist die Genorte *TPI*, *BG* und *GDH* allein oder im Sinne einer MLST genutzt (3, 23, 24, 30-34). Bei der Interpretation der Typisierungsergebnisse sollte man jedoch die Unterschiede zwischen den Assemblagen A und B beachten, da diese wegen einer DNA-Übereinstimmung von nur 77 % und einer Aminosäuren-Übereinstimmung von 78 % erhebliche Unterschiede aufweisen (10, 40). Ein wichtiger Unterschied besteht im Ausmaß der allelischen Sequenzheterozygotie (ASH) die bei verschiedenen Subpopulationen von Assemblage A 0,01-0,04 % beträgt, bei Assemblage B jedoch bis zu ca. 0,5 % (10, 41-43). Die Ursache für ASH liegt primär in der Tetraploidität (4n) des zweikernigen Parasiten begründet: Die vermehrungsfähige Form der Trophozoiten variiert zwischen 4n und 8n, jeweils abhängig vom Zeitpunkt im Zellzyklus, wohingegen die mit dem Stuhl ausgeschiedene vierkernige Zyste einen Ploidiegrad von 16n aufweist (35). Für *G. duodenalis* wurde lange Zeit angenommen, dass der Parasit sich ausschließlich asexuell reproduziert (44, 45). Dabei müssten jedoch in beiden Zellkernen völlig unabhängig voneinander verschiedenste Mutationen akkumulieren, weshalb man höhere ASH Raten als die ca. 0,5 %, die bei Assemblage B festgestellt wurden, erwarten würde (10). Es wird deshalb davon ausgegangen, dass *G. duodenalis* sich nicht ausschließlich asexuell fortpflanzt und auch andere genetischen Rekombinationsstrategien wie z.B. Formen der kryptischen sexuelle Reproduktion zur Anwendung kommen (35, 44-46). Die Ursache für die großen Unterschiede im Ausmaß der ASH zwischen den Assemblagen A und B ist unklar. Die Assemblagen A und B werden nach einer von Caccio et al. (40, 47) vorgeschlagenen Nomenklatur weiter in Subassemblagen unterteilt. Diese Einteilung basiert auf den Genotypen, die durch die MLST der Genorte *TPI*, *BG* und *GDH* ermittelt werden (40, 47). Demnach lässt sich die Assemblage A in die Subassemblagen AI, AII und AIII unterteilen, für die verschiedene Wirtsspezifitäten und Risikofaktoren postuliert werden (24, 34, 40, 47). Subassemblage AI wird als panglobaler, klonaler Erregerstamm beschrieben, der sowohl in Menschen als auch in vielen anderen Säugetieren vorkommt (48). Subassemblage AII ist genetisch diverser, wird allerdings fast ausschließlich im Menschen nachgewiesen, wohingegen Subassemblage AIII vor allem Huftiere und nur selten Menschen infiziert (47-49). Die MLST ist in der Lage, die verschiedenen Subtypen von Assemblage A zu unterscheiden; sie reicht jedoch nicht aus, um in Ausbruchssituationen Verwandtschaftsbeziehungen zwischen den untersuchten Isolaten festzustellen und Transmissionsketten sicher zu identifizieren (25). Deshalb wurde im Jahr 2018 ein neues MLST-Schema entwickelt, welches auf sechs neuen Genmarkern aufbaut, die über vier der fünf Chromosomen von *G. duodenalis* verteilt sind (25). Diese Genorte wurden

wegen ihrer hohen Variabilität zwischen den Subassemblagen AI und AII ausgesucht und waren wegen ihrer hohen Trennschärfe in der Lage, einen Ausbruch von *G. duodenalis* Subassemblage AII in einem schwedischen Krankenhaus nachzuweisen, der mittels MLST von nur *TPI*, *BG* und *GDH* unentdeckt geblieben wäre (25). Das neue MLST-Schema scheint also in der Lage zu sein, epidemiologische Verlinkungen zwischen verschiedenen Isolaten nachzuweisen. Es wurde jedoch bisher noch nicht von anderen Arbeitsgruppen verifiziert, ob diese erhöhte Trennschärfe auch in anderen Gegebenheiten reproduzierbar ist.

Für Assemblage B ist bisher das MLST-Schema am weitesten verbreitet, welches *TPI*, *BG* und *GDH* zum Ziel hat. Die Sequenzvariabilität ist hier so groß, dass theoretisch eine verlässliche Subklassifikation von Assemblage B möglich sein sollte, jedoch durch das Vorkommen von multiplen Infektionen bzw. aufgrund des hohen ASH-Anteils in der praktischen Umsetzung bisher keine adäquate Auflösung zur Unterscheidung von Isolaten erbrachte. (26, 47). ASH verursacht doppelte Linien in den Chromatogrammen der auszuwertenden Sequenzen. Diese, einzelne Nukleotide betreffenden Polymorphismen (engl. *single nucleotide polymorphisms*, SNP), beeinträchtigen die weitere Analyse der Sequenzen in z. B. phylogenetischen Programmen. Sequenzen mit ASH werden deshalb häufig aus Analysen ausgeschlossen (35, 50, 51). Da ASH in Assemblage B eher die Regel als die Ausnahme darstellt, ist in dieser Studie nach Möglichkeiten gesucht worden ASH in die Sequenzauswertung zu integrieren.

Primärziel dieser Studie war deshalb herauszufinden, inwiefern sich derzeit verfügbare MLST-Methoden für epidemiologische Anwendungen wie das molekulare Nachverfolgen von Transmissionsketten, Ausbruchsanalysen und die Identifikation von Infektionsquellen eignen. Dabei lag der Fokus bei Assemblage A auf der Verifizierung des neuen Typisierungsschemas, wohingegen bei Assemblage B die ansonsten wenig beachtete ASH in die Analyse integriert wurde. Als Sekundärziel sollte die Transmissionsdynamik in Deutschland untersucht und herausgefunden werden, ob Unterschiede zwischen autochthonen und Reise-assoziierten *G. duodenalis* Assemblage A und B Infektionen bestehen.

## 2. Material und Methodik

### 2.1 Einverständnis der Ethikkommission

Das Studiendesign wurde von der Ethikkommission der Charité – Universitätsmedizin Berlin Campus Benjamin Franklin genehmigt (EA4/171/19).

### 2.2 Arbeitsablauf

Der Arbeitsablauf der verschiedenen PCR-basierten Untersuchungen ist in der Abbildung S1 der publizierten Studie dargestellt (52). Für alle PCRs konnten bereits publizierte Protokolle genutzt werden. Alle Daten zu den einzelnen Genorten, den Primern, Zugangsnummern (engl. *accession numbers*) und zu den Referenzsequenzen finden sich in den Tabellen S1-3 der Originalpublikation (52).

### 2.3 Probensammlung

Folgende Labore haben anonymisierte Stuhlproben von symptomatischen, positiv auf *G. duodenalis* getesteten Erkrankten zur weiteren Genotypisierung an das RKI gesandt: Labor 28 (Berlin), Labor Enders (Stuttgart), Labor LADR (Geesthacht), das Labor des Instituts für Medizinische Mikrobiologie, Immunologie und Parasitologie der Universitätsklinik Bonn und das Labor der tropenmedizinischen Ambulanz des Universitätsklinikums Düsseldorf.

Die erste Probensammlung wurde zwischen 2012 und 2019 in der Tropenmedizinischen Ambulanz des Universitätsklinikums Düsseldorf durchgeführt. Der Datensatz bestand aus 112 Stuhlproben von 64 verschiedenen Erkrankten, die sich fast ausschließlich mit Reise-assoziierten Infektionen in der Klinik vorgestellt hatten. 24 chronisch infizierte Personen stellten mehr als eine Stuhlprobe zur Verfügung. Diese Fälle wurden als „longitudinale Fälle“ bezeichnet. Die Proben der longitudinalen Fälle sollten als Testmöglichkeit dienen, um herauszufinden, ob man eine hier anzunehmende epidemiologische Verbindung mittels MLST nachweisen kann. Die zweite Probensammlung wurde von 2017 bis 2019 in vier verschiedenen medizinischen Laboren durchgeführt. Diese Labore waren nicht an ein Krankenhaus angegliedert, sodass Stuhlproben von einem breiteren Patientenspektrum zu erwarten waren. Dadurch sollten nicht nur Reise-assoziierte Infektionen, sondern auch autochthone Infektionen erfasst werden. Die zweite Probensammlung bestand aus 171 Stuhlproben von 138 verschiedenen Erkrankten, darunter 23 chronisch infizierte, longitudinale Fälle. Da ein Teil dieser Proben von autochthonen Infektionen stammten, konnte das Sekundärziel untersucht werden. Hierzu wurden die Genotypisierungsergebnisse jeweils dem

Meldefall zugeordnet und auf Unterschiede in Bezug auf Geschlecht, Alter und Reiseanamnese (ja/nein) der Erkrankten untersucht. Diese Analyse wurde durch die Fachgruppe 35, Infektionsepidemiologie gastrointestinaler Infektionen, Zoonosen und tropischer Infektionen des RKI durchgeführt.

## 2.4 DNA-Extraktion

Am RKI wurden die Reststuhlproben bis zur schnellstmöglichen Weiterverarbeitung bei 4 °C gelagert. Die maximale Lagerungsdauer von zwei Wochen wurde nicht überschritten. Zysten wurden mittels Saccharose-Gradient-Flotation aus den Stuhlproben angereichert. Dabei wurden die Proben mit sterilem Wasser homogenisiert und durch vier Lagen Gaze hindurch gefiltert. Anschließend wurden die Proben bei 900 g bei Raumtemperatur für fünf Minuten zentrifugiert (Varifuge 3.OR des Herstellers Heraeus). Dieser Waschschrift wurde so oft wiederholt, bis der Überstand klar war. Das Pellet wurde resuspendiert und darunter eine Lage mit 1 M eiskalter Saccharose pipettiert. Anschließend wurden die Proben erneut bei 900 g und auf 4 °C gekühlt, ohne Bremse, zentrifugiert. Die Waschschriffe mit und ohne Saccharose wurden je zwei Mal wiederholt und das Pellet anschließend in 1 ml nukleasefreiem Wasser resuspendiert. Die Anzahl der Zysten wurde mittels einer Neubauer Zählkammer (C-chip, VWR international) bestimmt. Dies war wichtig, da sich herausstellte, dass die DNA-Gewinnung mittels Saccharose-Gradient bei einer Zystenanzahl unter  $10^5$  ineffizient war. Die Extraktion der DNA erfolgte für Proben mit  $>10^5$  Zysten mittels Maxwell 16 FFPE Plus LEV DNA Purification Kit und dem Maxwell 16 Instrument (Promega Corporation, Wisconsin, USA) wie vom Hersteller beschrieben. Bei Proben mit  $<10^5$  Zysten wurde die DNA direkt aus der Stuhlprobe mittels QIAamp Fast DNA Stool Mini KIT (Qiagen, Hilden, Germany) nach Protokoll des Herstellers gewonnen. Die DNA wurde auf 100 µl mit nukleasefreiem Wasser gelöst und die Konzentration der DNA mittels Quantus Fluorometer (Promega) spektrophotometrisch bestimmt. Die DNA wurde dann bis zur weiteren Nutzung bei 4 °C gelagert.

## 2.5 Real-time PCR

Zunächst wurde eine real-time PCR (qPCR) durchgeführt, um die Anwesenheit der *G. duodenalis* DNA zu bestätigen und PCR-inhibierende Faktoren auszuschließen. Zielgen zur Bestätigung der Infektion mit *G. duodenalis* war ein 62 Basenpaare (BP) großes Fragment des in mehreren Kopien vorliegenden *ssu-rRNA* Gens (53, 54). Dabei wurden 133,3 nM *Giardia*-80F, 133,3 nM *Giardia*-127R und 66,6 nM der doppelt markierten TaqMan Sonde *Giardia*-105 verwendet. Im gleichen

Ansatz wurde die qPCR zum Inhibitionsausschluss parallel durchgeführt. Dabei wurde als interne Amplifikationskontrolle (IAC) 800 nM dd-IAC-f, 800 nM dd-IAC-r, 100 nM IAC Sonde dd-IAC-Cy5 und 2,5 µl DNA der IAC-Matrize als Kontrolle, wie von Deer et al. (55) beschrieben, verwendet. Das Volumen der beiden PCRs betrug 25 µl, zusätzlich wurden 12,5 µl Maxima Probe qPCR Master Mix 2X (Thermo Scientific), 1-2 µl DNA der zu untersuchenden Stuhlprobe and 2 µl nukleasefreies Wasser zugesetzt. Die Amplifikation erfolgte zunächst durch Inkubation für 10 Min bei 95 °C, gefolgt von 40 Zyklen mit 15 s bei 95 °C, 30 s bei 60 °C und 30 s bei 72 °C. Die qPCRs wurden entweder im AriaMx Real-time PCR System (Agilent) oder im CFX96 Touch Real-Time PCR Detection System (Bio-Rad) durchgeführt und die erhaltenen Daten mittels passender Software (Agilent Aria 1.5 (Agilent) oder CFX Maestro 1.0 (Bio-Rad Laboratories)) ausgewertet. Die PCR wurde als inhibiert betrachtet, wenn der Zyklusschwellenwert (engl. *cycle threshold*, ct) der Probe für die Detektion eines klaren Signals mit dem Farbstoff Cy5  $\geq 2$  Zyklen über dem Signal der verunreinigungsfreien Positivkontrollen lag (55). Im nächsten Schritt wurden alle Proben auf zugrundeliegende Infektionen mit Assemblagen A und B überprüft. Dafür wurden ein 168 BP großes Fragment des *TIF* Gens für die Detektion von Assemblage A und ein 99 BP großes Stück des *CATH* Gens für den Nachweis von Assemblage B verwendet (56). Wie im Protokoll von Van Lith et al. (56) beschrieben, wurden 133,3 nM *TIF* Assemblage A Vorwärtsprimer und 133,3 nM *TIF* Assemblage A Rückwärtsprimer oder 133,3 nM *CATH* Assemblage B Vorwärtsprimer and 133,3 nM *CATH* Assemblage B Rückwärtsprimer eingesetzt. Die beiden Reaktionen wurden in Duplikaten in verschiedenen Probengefäßen mit jeweils einem Volumen von 25 µl durchgeführt, denen zusätzlich zu den Assemblage-spezifischen Primern 12,5 µl Maxima SYBR Green/ROX qPCR Master Mix (2X) (Thermo Scientific), 1-2 µl DNA der untersuchten Probe und 8,5 µl nukleasefreies Wasser hinzugesetzt wurden. Die Amplifikation bestand aus einer Inkubation für 5 Min bei 95 °C, gefolgt von 45 Zyklen mit 12 s bei 95 °C, 12 s bei 55 °C und 12 s bei 72 °C (56). Dabei wurden am Ende von jedem Zyklus ct-Wert und Schmelzpunkt erfasst. Die qPCR wurde mit einem CFX96 Touch Real-Time PCR Detection System durchgeführt und die Daten mit der dazugehörigen Software (CFX Maestro 1.1, Bio-Rad) ausgewertet. Proben wurden als positiv betrachtet, wenn die Schmelzpunkte für Assemblage A zwischen 80,5 und 82,0 °C und für Assemblage B zwischen 75,5 und 78,0 °C lagen.

## 2.6 MLST der *G. duodenalis* Isolate

Die Vervielfältigung mittels PCR und Sequenzanalyse der Genloci *TPI*, *BG* und *GDH* sind die am weitesten verbreiteten Typisierungsmethoden. Die Amplifikation fand als Nested-PCR (zu

Deutsch „geschachtelte PCR“) statt und wurde nach standardisierten Protokollen durchgeführt: *TPI* (3), *BG* (31) und *GDH* (30). Die separaten PCR-Ansätze mit einem Volumen von jeweils 50 µl enthielten 200 nM von jedem der *TPI* Primer, 300 nM von jedem der *BG* Primer und 500 nM von jedem der *GDH* Primer, 2,5 U DNA-Polymerase (DreamTaq, Thermo Scientific), 5 µl 10X Puffer (DreamTaq, Thermo Scientific), 200 µM der verschiedenen dNTP (Carl Roth), bei *TPI* noch zusätzlich 1 mM MgCl<sub>2</sub> (Bioline) und je nach verwendeten Primern 38,5 µl, 39,5 µl oder 37,5 µl nukleasefreies Wasser. Die PCRs wurden wie zuvor beschrieben (3, 30, 31) in T-Gradient Thermocycler-Geräten (Biometra) durchgeführt. In jedem PCR-Durchlauf wurde *G. duodenalis* DNA (aus der RKI-eigenen Biobank) als Positivkontrolle und nukleasefreies Wasser als Negativkontrolle mitgeführt. Anschließend wurden 8 µl des PCR-Produkts mittels Gelelektrophorese in einem 1,35 % Agarose Gel (Biozym Scientific) analysiert. Diese Produkte wurden mit Midori Green Advance (NPPON Genetics Europe) gefärbt und unter einem Transilluminator (Serva) mit Blaulicht sichtbar gemacht. Um die Größe des Produkts abschätzen zu können, wurde eine DNA-Leiter mit jeweils 100 BP großen DNA-Fragmenten (Fermentas) verwendet. Proben, die die passende Größe in der Gelelektrophorese aufwiesen, wurden nach Protokoll des Herstellers mit Exo Sap-IT PCR Product Cleanup (Thermo Scientific) aufgereinigt. Anschließend wurde nochmals eine Sequenzierungsreaktion mit den zweiten Primern der jeweiligen Genorte und dem BigDye Terminator v3.1 Cycle sequencing kit (Applied Biosystems) durchgeführt (3, 30, 31). Die Amplifikation erfolgte durch Inkubation für 2 Min bei 96 °C, gefolgt von 25 Zyklen mit 10 s bei 96 °C, 5 s bei – je nach Primer – 55 bis 60 °C und 4 Min bei 60 °C. Nach dieser PCR folgte die Sanger-Sequenzierung in der RKI-eigenen Sequenzierungseinrichtung. Alle Proben wurden bidirektional sequenziert. Für die Auswertung wurden die Primer vom DNA-Strang entfernt, sodass für *TPI*, *BG* und *GDH* nur noch jeweils 490 BP, 475 BP und 393 BP übrigblieben. Die Sequenzen der Proben wurden dann in der Reihenfolge *TPI-BG-GDH* verknüpft, sodass ein insgesamt 1358 BP großes DNA-Fragment für die weitere Analyse entstand.

## 2.7 MLST-Schema für Assemblage A

Stuhlproben, die als Assemblage A oder gemischte A + B Infektionen charakterisiert worden waren, wurden zusätzlich, wie 2018 von Ankarklev et al. (25) publiziert, auf den Genorten *HCMP22547*, *CID1*, *RHP26*, *HCMP6372*, *DIS3* und *NEK15411* (25) mittels Nested-PCR untersucht. Der Reaktionsansatz (50 µl) enthielt 400 nM von jedem der vier Primer, 2,5 U DNA-Polymerase (DreamTaq, Thermo Scientific), 10X Puffer (DreamTaq, Thermo Scientific), 200 µM

dNTP (Carl Roth) und 37,5 µl nukleasefreies Wasser (25). Die Amplifikation wurde wie zuvor beschrieben durchgeführt (25). Positiv- und Negativkontrollen wurden mitgeführt. In der Gelelektrophorese positive Proben, die in ihrer Größe dem Zielprodukt entsprachen, wurden mittels Sequenzierungsreaktion amplifiziert. Die Genfragmente wurden in der Reihenfolge *HCMP22547*, *CID1*, *RHP26*, *HCMP6372*, *DIS3* und *NEK15411* zu einer 3414 BP langen DNA-Sequenz zusammengefügt.

## 2.8 Analyse der DNA-Sequenzen

Für die Bestimmung der Konsensussequenzen wurden die Ergebnisse beider Sequenzierreaktionen jedes PCR-Produktes manuell analysiert. Im Gegensatz zu vorherigen Studien wurde dabei die ASH gezielt berücksichtigt, d. h., die Chromatogramme der beiden Sequenzierungen wurden zunächst positionsgerecht übereinandergelegt und visuell durchgesehen, um „unklare“ Positionen mit doppelten Signalen für verschiedene Basenpaare, falls erforderlich, zu annotieren. Dieses Verfahren war jedoch nicht nur zeitaufwendig, sondern auch fehleranfällig. Zur Reduktion der Fehlerwahrscheinlichkeit wurde daher standardisiert vorgegangen: Um die heterozygoten Positionen schneller zu markieren, wurde die „heterozygotes plugin“ Erweiterung von Geneious 11.1.5 benutzt. Da vom Programm allerdings nicht automatisch alle uneindeutigen Positionen richtig annotiert wurden, wurden zwei Kriterien definiert, um nachvollziehbar heterozygote Positionen zu bewerten: Das Primärkriterium war, dass der kleinere Gipfel im Chromatogramm mindestens 25 % der Gipfelhöhe des größeren Peaks erreichen musste, um als uneindeutige Position zu gelten. Das zweite Kriterium war, dass das zweigipflige – oder seltener mehrgipflige – Chromatogramm sowohl im Vorwärts-, als auch im Rückwärtsstrang vorhanden war. Erst wenn diese beiden Kriterien erfüllt waren, wurde die Position im Patientenstrang als heterozygot entsprechend der Klassifikation der IUPAC gekennzeichnet (57). Anschließend wurden die Patientenstränge der einzelnen Genfragmente mithilfe des „multiple alignment“ Programms von Geneious 11.1.5 mit den passenden Referenzsträngen verknüpft. Dazu wurden für die Loci *TPI*, *BG* und *GDH* die Referenzen von Fahmy et al. (58) verwendet, während für das erweiterte Schema der Assemblage A die von Ankarklev et al. (25) publizierten Sequenzen genutzt werden konnten. Die einzelnen Referenzstränge für *TPI*, *BG* und *GDH* wurden in der GenBank (59) anhand der Zugangsnummern (engl. accession numbers) identifiziert: Verwendet wurden Referenzsequenzen für Subassemblage AI (Zugangsnummern L02120, X85958, M84604), AII (Zugangsnummern U57897, AY072723, AY178737), AIII (Zugangsnummern EU781002, EU769206, EU769223), BIII (Zugangsnummern AY228628, AY072726, AF069059) und BIV (Zugangsnummern L02116,

AY072725, AY178738). Die Referenzsequenzen für das erweiterte Typisierungsschema nach Ankarklev et al. (25) enthielten alle von diesen beschriebenen Genotypen: Für *HCMP22547* Referenzstränge mit den Zugangsnummern MG520233 bis MG520242, für *CID1* die Zugangsnummern MG520215-MG520224, für *RHP26* die Zugangsnummern MG520254-MG520262, für *HCMP6372* die Zugangsnummern MG520225-MG520232, für *DIS3* die Zugangsnummern MG520263-MG520270 und für *NEK15411* die Zugangsnummern MG520243-MG520253. Anschließend wurden die verschiedenen Patientenstränge der einzelnen Genfragmente mit korrekter ASH-Annotation zusammengefügt. In den wenigen Fällen, in denen Personen mehrere Stuhlproben an einem einzigen Tag abgegeben hatte, wurde die Sequenz mit der höchsten Qualität der Sequenzierung für die verkettete Endsequenz verwendet. Die Endsequenzen der verschiedenen Proben (entweder 1358 oder 3414 BP) wurden miteinander nach Assemblagen geordnet mittels Geneious 11.1.5 Software zusammengefügt, eine Datei dieser Zusammenstellung exportiert und mithilfe des Programms Prism (GraphPad software, LLC) als Distanzmatrix der unterschiedlichen SNP visualisiert.

## 2.9 Testansatz zur Identifizierung epidemiologisch verbundener *G. duodenalis* Isolate

Wie in der Publikation dieser Arbeit (52) beschrieben, wurde eine Ausreißeranalyse verwendet, um Isolate als zusammenhängend zu definieren. Dabei wurde der Mittelwert der SNP-Unterschiede zu allen anderen Isolaten im jeweiligen Datensatz minus zwei Standardabweichungen berechnet. Zwei Isolate wurden als epidemiologisch verbunden betrachtet, wenn sie reziprok unter dem Grenzwert des jeweils anderen Isolates waren. Diese Methodik ließ sich allerdings nur für Assemblage B verwenden, da die genetischen Unterschiede auf den häufig verwendeten Genloci *TPI*, *BG* und *GDH* bei Assemblage A zu gering sind, um Ausreißer identifizieren zu können. Bei dem neuen Typisierungsschema von Assemblage A waren die Unterschiede zwischen den einzelnen Isolaten dagegen so groß, dass eine Ausreißeranalyse zur Identifikation verwandter Isolate nicht erforderlich war. Ein Beispiel für die beschriebene Ausreißeranalyse ist in der zusätzlichen Abbildung 2 der Publikation dargestellt (52). Sogenannte „longitudinale Fälle“ wurden untersucht, um mittels Ausreißeranalyse die Hypothese zu überprüfen, ob sich epidemiologisch verbundene Isolate - unter Berücksichtigung der ASH - in weniger SNPs unterscheiden als Proben ohne epidemiologischen Zusammenhang. Longitudinale Fälle waren Proben, die von derselben Person zu unterschiedlichen Zeitpunkten gewonnen wurden. Die Proben zeichneten sich dadurch aus, dass sie jeweils von einer Person mit einer



chronischen oder rezidivierenden *G. duodenalis* Infektion stammten. Da in der Mehrheit davon ausgegangen werden kann, dass die Personen über die Zeit überwiegend mit demselben Isolat infiziert waren (Deutschland ist ein Niedriginzidenzgebiet, weitere Infektionen in kurzer Folge mit anderen Isolaten sind daher unwahrscheinlich), wurden diese Proben als Testfall für epidemiologisch verbundene Isolate benutzt.

## 2.10 Analyse mit phylogenetischen Werkzeugen

Die verketteten Sequenzdaten der jeweiligen MLST-Schemata wurde mit verschiedenen phylogenetischen Analysewerkzeugen gängiger Programme untersucht (60). Darunter waren in Geneious 11.1.5 implementierte Phylogenieanalysetools, PhyML (61) und bayesianische Analysen (62). Die Analysen, die den Phylogramme in Abbildungen 2-4 zugrunde liegen (52), wurden mit der MEGA 7.0.26 Software (63) und der *Neighbor-Joining* Methode durchgeführt. Weitere Angaben zur verwendeten Phylogenie finden sich in der Publikation (52).

## 2.11 Statistische Analyse

Um signifikante Unterschiede hinsichtlich Alter, Geschlecht und Typisierungsergebnissen zwischen der Gruppe der autochthonen und der reiseassoziierten Infektionen zu identifizieren, wurden die Daten mit dem Programm Graph Pad Prism 8.4.0 analysiert. Dabei kamen vor allem der exakte Test nach Fisher und der Chi-Quadrat-Test zur Anwendung. P-Werte < 0,05 wurden als signifikant betrachtet (64). Die ergänzende Abbildung 3 wurde mithilfe der „Open Source“ Software QGIS und freien Vektordaten der Website [naturalearthdata.com](http://naturalearthdata.com) erstellt. Für die Schätzung der Anzahl der multiklonalen Infektionen kam ein geometrisches Folge-Modell mit der Gleichung  $\sum_{k=0}^{\infty} (P(\text{monoAss}))^{k+1}$  zum Einsatz (52).  $P(\text{monoAss})$  stellt dabei die Wahrscheinlichkeit für eine Monoinfektion dar, die durch die Anzahl der in dieser Studie beobachteten gemischten Infektionen für Assemblage B (B plus gemischte A + B Infektionen) und Assemblage A (A plus gemischte A + B Infektionen) geschätzt wurde (52).

## 2.12 Zugangsnummern der DNA-Sequenzen

Sämtliche in dieser Studie erzeugten DNA-Sequenzen wurden in die öffentliche Datenbank *Genbank* (59) hochgeladen (Zugangsnummern (*accession number*) MT878631 - MT879098).

### 3. Ergebnisse

Bei dieser Dissertation handelt es sich um eine Publikationspromotion (52).

In Bezug auf die eingangs beschriebenen Fragestellungen lassen sich die Ergebnisse von folgt zusammenfassen:

Primärziel war die Evaluation aktuell zur Verfügung stehender MLST-Verfahren in Bezug auf deren Nutzen zur Identifizierung von epidemiologisch potenziell zusammenhängenden humanen *G. duodenalis* Infektionen. Das derzeit am häufigsten verwendete MLST-Schema mit den Genloci *TPI*, *BG* und *GDH* eignete sich sowohl bei Assemblage A als auch B nur sehr eingeschränkt für epidemiologische Anwendungen (vergleiche Abbildungen (52)).

Das hatte bei Assemblage A und B verschiedene Gründe. Assemblage A kann mit diesem MLST Schema lediglich in eine geringe Anzahl von Sequenz-Untergruppen eingeteilt werden. Das gilt insbesondere für die beim Menschen am häufigsten vorkommende Assemblage AII, sodass die Variabilität zur sicheren Nachverfolgung von z. B. Infektionsketten nicht genügt. Im Gegensatz zu Assemblage B zeigt Assemblage A lediglich geringe ASH in den beschriebenen Sequenzabschnitten. Assemblage B Isolate hingegen zeichnen sich durch eine relativ höhere genetische Diversität in den Sequenzen aus, zeigen jedoch auch einen höheren Anteil von ASH-Positionen. ASH stellt für die Unterscheidung einer multiklonalen Infektion von einem heterozygoten Isolat eine große Herausforderung dar. Da *Giardia* Isolate nicht effizient kultiviert werden können, ist man im Wesentlichen auf die Analyse von Gesamt-DNA aus Stuhlproben angewiesen, in denen Infektionen mit multiplen *G. duodenalis* Genotypen vorhanden sein können (siehe unten) oder wie dargestellt heterozygote Parasiten eines Genotyps. Für Assemblage A Proben konnte die geringe Sequenzvariation in den Gensequenzen des gewöhnlichen MLST (*GDH*, *TPI* und *BG*) bestätigt werden: Longitudinalen Fälle waren nicht sicher zu identifizieren. Das vor kurzem publizierte Typisierungsschema von Ankarklev et al. (25) - basierend auf sechs unterschiedlichen Genloci - zeigte eine deutlich bessere Auflösung und identifizierte fünf von sechs longitudinalen Fällen. Es wurden sowohl Genotypen gefunden, die in Ankarklev et al. (25) 2018 publiziert wurden, als auch 15 neue Genotypen, die zur weiteren Aufklärung der Populationsstruktur von Assemblage A beitragen.

Bei Assemblage B konnte durch repetitive PCR- und Sequenzanalysen von DNA-Proben, die von klonalen, axenischen in vitro-Kulturen stammten, gezeigt werden, dass die Wiederfindung von ASH Positionen eines Isolates gut reproduzierbar ist. Multiklonale Infektionen mit verschiedenen Assemblage B Subtypen, wie sie in Stuhlproben vorkommen können, hemmten jedoch eine

Identifikation von epidemiologisch verbundenen Isolaten.

Um die Wahrscheinlichkeit von zugrundeliegenden Mehrfachinfektionen gleicher Assemblagen zu schätzen, wurden die in der Studie ermittelten Proportionen von Assemblage A, B und A + B Infektionen herangezogen und mittels geometrischer Folgeberechnung analysiert. Multiklonale Infektionen mit Assemblage B wurden darauf basierend auf ca. 50 % geschätzt und multiklonale Infektionen mit Assemblage A auf ca. 25 %. Um ein standardisiertes Bewertungssystem zur Beurteilung von als zusammenhängend zu definierenden Isolaten einzuführen, wurde eine reziproke Ausreißeranalyse für jedes Isolat im Vergleich mit jedem anderen Isolat der Untersuchungsgruppe durchgeführt. Hierzu wurde der Mittelwert der SNP-Unterschiede zwischen jedem Isolat im Datensatz minus zwei Standardabweichungen berechnet und als Grenzwert für die Bewertung „epidemiologisch verbunden ja oder nein“ definiert. Im Sinne eines zweiseitigen Testes wurden zwei Isolate lediglich dann als „wahrscheinlich epidemiologisch verbunden“ beurteilt, wenn die Ausreißertests für jedes Isolat das jeweils andere Isolat identifizierten. Dadurch konnten im ersten Datensatz sechs von neun und im zweiten Datensatz vier von fünf longitudinalen Infektionen identifiziert werden. Zusätzlich clusterten stets einige Isolate um die Referenzsequenzen von BIII und BIV. Im zweiten Datensatz wurde ein Cluster aus fünf im MLST identischen Isolaten detektiert: Ob dieser Cluster eine häufigere Subgruppe oder eine lokale Transmission einer wahrscheinlich aus Ost-Afrika importierten Infektion darstellte, ließ sich nicht abschließend klären. Mögliche Gründe hierfür werden in der Diskussion angeführt. Zusammenfassend lässt sich sagen, dass sich aktuelle Typisierungsverfahren in Kombination mit der vorgestellten reziproken Ausreißeranalyse wahrscheinlich für epidemiologische Fragestellungen wie die Identifikation von Infektionsketten, Infektionsquellen und Ausbruchsanalysen nutzen lassen. Multiklonale Infektionen erschweren jedoch die Zuordnung, sodass sich eine Verbindung von zwei Isolaten nicht sicher ausschließen lässt. Sekundärziel war die Untersuchung der Transmissionsdynamik von *G. duodenalis* in Deutschland. Unseres Wissens nach, war dies die erste Studie aus Deutschland, die neben Reise-assoziierten Infektionen auch gezielt autochthone Giardiasis-Fälle untersucht hat. Autochthone und Reise-assoziierte *Giardia*-Infektionen sind in Deutschland zu jeweils etwa 50 % am Infektionsgeschehen beteiligt. Die Frage, ob es sich bei den „autochthonen Infektionen“ tatsächlich um Infektionen durch bislang unbekannte Infektionsquellen in Deutschland handelt oder diese unentdeckte sekundäre Transmission widerspiegeln, konnte in der vorliegenden Arbeit nicht abschließend beantwortet werden. In beiden Datensätzen gab es keine signifikanten Unterschiede zwischen den beiden Gruppen hinsichtlich Assemblage Typ, Geschlecht oder das Alter der Betroffenen.

## 4. Diskussion

In der vorliegenden Arbeit ließ sich zeigen, dass aktuelle MLST-Schemata in der Lage sind, epidemiologisch verbundene *G. duodenalis* Isolate - am Beispiel der longitudinalen Fälle - mit vertretbarer Sensitivität und Spezifität zu identifizieren (52). Dafür sind jedoch je nach Assemblage unterschiedliche Typisierungsschemata nötig. Des Weiteren konnte der Wert des neuen erweiterten MLST-Schemas für Assemblage A bestätigt und 15 neue Genotypen zu den von Ankarklev et. al (25) beschriebenen hinzugefügt werden (52). Anhand der identifizierten Fälle von gemischten Assemblage A + B Infektionen konnte außerdem die Anzahl der multiklonalen Infektionen einer Assemblage abgeschätzt und die Auswirkungen auf die Erkennung epidemiologisch verbundener Isolate demonstriert werden (52). Je nach Assemblage sollten verschiedene Strategien für die MLST und die Auswertung der Ergebnisse angewendet werden (24, 25, 34, 65). Unseres Wissens waren wir die erste Arbeitsgruppe, die das von Ankarklev et al. (25) neu entwickelte MLST-Schema für Assemblage A angewendet hat. Auch wenn die Genotypisierung von sechs, anstelle der gewöhnlich genutzten drei Genorten einen erhöhten Arbeitsaufwand darstellt, hat sich dieser Mehraufwand in dieser Studie gelohnt. Epidemiologisch verbundene Isolate ließen sich bis auf einen Fall problemlos identifizieren. Bei den Fallnummern 468-01/-02 unterschieden sich die beiden Proben durch ASH auf einem Fragment von *NEK15411* in elf BP. ASH kommt in Assemblage A weniger häufig vor als in Assemblage B, und könnte, wenn sie doch vorkommt ein Indikator für eine gemischte Assemblage A + A Ko-Infektion sein. Die höchste Anzahl an ASH bei dem neuen MLST-Schema wurde in einer Probe gefunden, die zunächst auf den häufig verwendeten Genen *TPI*, *BG* und *GDH* als Assemblage B und in der qPCR als A + B Mischinfektion klassifiziert wurde. Andere Proben mit ähnlichen Typisierungsergebnisse in qPCR und häufig genutztem MLST zeigten keine ASH. Dies verdeutlicht noch einmal die Schwierigkeit Mischinfektionen – auf Assemblage oder Subassemblage Level - zu identifizieren. Die erschwerte Detektion der epidemiologisch verbundenen Proben könnte also an einer Infektion mit verschiedenen Assemblage A Isolatoren liegen. Die Tatsache, dass die Proben der Fälle 469-01 und 468-01 völlig identisch waren, ließ aufgrund der sehr guten Trennschärfe des neunten MLST eine unbekannt epidemiologische Verbindung vermuten, auch wenn beide Proben aus unterschiedlichen Städten eingesandt worden waren. Erwähnenswert ist, dass die Verkettung der Genfragmente *CID1*, *RHP26* und *HCMP6372* ähnliche Ergebnisse erbrachte wie das volle sechs Loci umfassende MLST-Schema und alle longitudinalen Fälle erfolgreich identifiziert wurden. Weitere Studien sind wünschenswert, um zu

untersuchen, ob die drei genannten Gene als ressourcenschonende Variante des neunten MLST-Schemas genutzt werden könnten.

Um eine Probe einer potenziellen Mischinfektionen mit Assemblage A + B zuzuordnen, wurden sowohl die Assemblage-spezifische qPCR als auch die Ergebnisse der PCR und Sequenzierung der Markergene *TPI*, *GDH* und *BG* herangezogen. Die Ergebnisse zeigten, dass mittels qPCR in 13 von 137 Proben Assemblage A + B Mischinfektionen entdeckt wurden, die nicht durch MLST identifiziert werden konnten, wohingegen es umgekehrt in 8 Fällen nicht möglich war, die apparente A + B Mischinfektion, die in der MLST Analyse angezeigt wurde, durch die qPCR zu bestätigen. Ursächlich dafür sind wahrscheinlich Unterschiede in den methodischen Ansätzen: Die konventionelle PCR amplifiziert nur den am häufigsten vorkommenden Genotyp, wohingegen die Detektion der Mischinfektionen in der qPCR über die Analyse von Schmelzkurven funktioniert. In der qPCR kann so besser die Assemblage nachgewiesen werden, die in geringerer Quantität vorhanden ist (56). Die Ergebnisse dieser Arbeit legen nahe, verschiedene Methoden zu nutzen, um die Anzahl an gemischten Infektionen nicht zu unterschätzen. Dazu raten bereits auch andere Forschungsgruppen (36, 60, 66, 67). Einige Arbeitsgruppen (48, 68, 69) haben genetische Unterschiede innerhalb der einzelnen Assemblagen berichtet, die auf dem geographischen Ursprung der Proben basierten. Zahlreiche andere Studien (60, 71, 72) konnten diese Unterschiede jedoch nicht feststellen. Tatsächlich zeigt diese Arbeit, dass die niedrige Trennschärfe des verbreiteten MLST-Schemas, welches *TPI*, *BG* und *GDH* nutzt, die Identifikation von genetischen Unterschieden innerhalb der Assemblagen enorm beeinträchtigt (35, 50, 68). Ohne eindeutige Detektion von Subassemblagen und ohne Berücksichtigung der in dieser Arbeit dargestellten Auswirkungen multiklonaler Infektionen auf die Typisierungsergebnisse erscheint deshalb auch die häufig geführte Diskussion, welche Symptome bei Infektionen mit Assemblage A bzw. B häufiger auftreten, schwierig (16, 23, 65).

Es wurde zuvor vorgeschlagen einzelne Genloci wie *GDH* von der Auswertung auszuschließen, da *GDH* sich bei der Analyse von Assemblage B Sequenzen inkongruent zu den Typisierungsergebnissen von *TPI* und *BG* verhalten und so die Subgruppenanalyse noch weiter erschweren würden (68). In dieser Studie verbesserte dieses Vorgehen allerdings in keinem Fall die Identifikation der longitudinalen Fälle. Es verursachte dagegen eine große Anzahl falsch positiver, epidemiologischer Verbindungen, da sich die Isolate am von der Interpretation ausgeschlossenen Genort deutlich unterschieden. Da trotz aller Mängel Genotypisierungsverfahren immer häufiger und bei immer mehr Spezies ihre Anwendung finden (23-25, 35, 65), ist die Erforschung besser diskriminierender genetischer Marker sowie die Entwicklung neuer Hilfsmittel für die Interpretation der Genotypisierungsergebnisse

wünschenswert. In dieser Studie wurde deshalb eine reziproke Ausreißeranalyse durchgeführt und gezeigt, dass dadurch in den vorliegenden Datensätze sechs von neun bzw. vier von fünf longitudinalen Fällen korrekt zugeordnet werden konnten (52). Proben ohne bekannte epidemiologische Verbindung, die mittels Ausreißeranalyse ebenfalls als genetisch verbunden eingestuft wurden, wurden als falsch positiv bewertet. Die errechnete Sensitivität lag somit bei ca. 67-80 %, die Spezifität betrug etwa 80 %. Die Sensitivität wurde durch die Anzahl der multiklonalen Infektionen reduziert, die in dieser Arbeit für Assemblage B ca. 50 % betrug (52). Wie in der ersten Abbildung dargestellt, konnten bei mehrfacher PCR-Untersuchung derselben Stuhlprobe verschiedene Genotypen mit bis zu 17 BP Unterschied identifiziert werden (52). Da der Mittelwert der Unterschiede zwischen allen 115 in dieser Studie untersuchten Assemblage B Isolaten 19,7 BP betrug, ließ sich in diesen Fällen auch trotz reziproker Ausreißeranalyse keine epidemiologische Verbindung nachweisen. In einigen Fällen von multiklonalen Infektionen waren die BP-Unterschiede allerdings deutlich geringer, sodass sich je nach BP-Abstand und -Beschaffenheit der anderen Proben des Datensatzes mittels reziproker Ausreißeranalyse trotzdem epidemiologische Verbindungen vermuten ließen.

Neben den longitudinalen Fällen wurde ein aus fünf Personen bestehender Cluster identifiziert, in dem alle Isolate, auf allen drei Genen völlig identisch waren und keinerlei ASH aufwiesen. Anhand der Fallangaben ließ sich ermitteln, dass sich der Indexpatient wahrscheinlich in Ostafrika infiziert hatte. Die folgenden vier Personen haben sich in den zehn Monaten danach angesteckt, die Infektionen wurden jedoch von den Gesundheitsämtern als autochthone Infektionen eingestuft. Alle Personen konnten dem Berliner Raum zugeordnet werden. Die Sequenzanalyse ergab des Weiteren, dass die Sequenz des Clusters Ähnlichkeiten mit einem weiteren Cluster hatte, in dem auch der Referenzstrang von BIV enthalten war. Obwohl sich die sechs Personen aus diesem „BIV-Cluster“ – laut Gesundheitsamt - in verschiedenen Ländern infiziert hatten, könnte die große Ähnlichkeit für eine u. U. global existierende Subgruppe BIV-ähnlicher *G. duodenalis*-Isolate sprechen. Dies wird auch dadurch gestützt, dass sich im ersten Datensatz einige Isolate in der relativen Nähe von BIII- und BIV-Referenzsträngen befanden. Obwohl eine strenge Einteilung der Assemblage B in die Subassemblagen BIII und BIV umstritten ist (65, 68, 73), scheint es sich doch um Stämme zu handeln, die weltweit stärker verbreitet sind als andere und deshalb auch mit mehr Isolaten verbunden scheinen. Die fünf Personen mit im MLST identischen Isolaten könnten deshalb eine nicht entdeckte Transmissionskette darstellen oder mit einem global oder lokal gehäuft auftretenden Genotyp infiziert worden sein. Die einhundertprozentige Übereinstimmung aller fünf Proben legt jedoch einen lokalen und unentdeckt gebliebenen Ausbruch nahe.

Die Ergebnisse dieser Forschungsarbeit zeigen, dass eine Würdigung der lokalen

Populationsstruktur eminent ist, um Genotypisierungsergebnisse richtig einzuordnen. Lokal gehäuft auftretende Subgruppen könnten zum Beispiel die Grenzwerte bei der Berechnung der reziproken Ausreißeranalyse massiv beeinflussen, während man falsch positive Verbindungen erhalten würde, wenn man die zu untersuchenden Isolate ohne lokale Bezüge mit zufälligen Sequenzen aus der Genbank vergleicht. Die Subgruppenanalyse ließe sich eventuell mit einem neuen, 2021 publizierten, MLST-Schema für Assemblage B weiter verbessern (74). Folgende Studien müssen herausarbeiten, ob sich die gezeigten Effekte reproduzieren lassen und ob die vorgeschlagenen Genfragmente sich auch in der molekularen Routineüberwachung effektiv anwenden lassen. Neben der vorgeschlagenen Ausreißeranalyse mit reziprokem Grenzwert wurde auch versucht, die Lösungsstrategien anderer Forschungsbereiche aufzugreifen. Beispielhaft hierfür ist der Fokus auf die Analyse der SNP-Positionen, an denen sich 5-10 % aller Proben unterscheiden („majority variants“). SNP an anderer Stelle wurden aus der Analyse ausgeschlossen. Alternativ wurde versucht, durch Weglassen genau dieser häufigen SNP-Positionen und Konzentration auf die Auswertung der selteneren SNP die Detektionsrate der longitudinalen Fälle zu verbessern. Diese Strategien waren zuvor bei Studien zu Mykobakterien erfolgreich (75-77). In dieser Studie jedoch, erbrachte keiner dieser beiden Ansätze vielversprechende Ergebnisse. Im Vergleich zu anderen Studien konnte in dieser Arbeit bei einem vergleichsweise großen Anteil der Proben die Assemblage bestimmt werden (78-80). Die Unfähigkeit, bei allen vorher positiv getesteten Stuhlproben die Assemblage zu ermitteln, liegt wahrscheinlich darin begründet, dass für die Amplifikation der zumeist als eine Kopie vorkommenden Genfragmente eine hohe Sensitivität benötigt wird, die in Proben mit geringem DNA-Gehalt trotz mehrfacher Nested-PCR nicht erreicht werden kann (66, 81). Wie auch von anderen Forschungsgruppen beschrieben, beeinträchtigen gemischte Assemblage A + B Infektionen sowie multiklonale Infektionen einer Assemblage die Reproduzierbarkeit der Typisierungsergebnisse (51, 82). Bei der Interpretation von Genotypisierungsergebnissen ist zu beachten, dass PCR-basierte Methoden dazu neigen, das am häufigsten vorkommende Isolat mit der größten DNA-Menge zu amplifizieren (1, 65, 67, 83). Befinden sich in einer Stuhlprobe verschiedene Assemblagen oder Subassemblagen, wäre die Chance deshalb hoch, bei mehreren PCRs mit Material aus einer Probe verschiedene Assemblagen oder Subassemblagen zu detektieren. Diese Theorie wird durch die Beobachtung von „wechselnden Assemblagen“ (engl. *assemblage swapping*) unterstützt. Bei diesen „wechselnden Assemblagen“ werden durch Genotypisierung von Stuhlproben von aufeinanderfolgenden Tagen verschiedene, meist „eindeutige“ Assemblagen generiert. Dieses Phänomen konnte in dieser Studie bei zwölf Personen (4,9 %) beobachtet werden, was etwa der Hälfte der von Sprong et al. beschriebenen Fälle entspricht (29).

Meist wurden die Parasiten initial als Assemblage B klassifiziert worden. Häufig wurde in einer weiteren Stuhlprobe eine gemischte Infektion mit Assemblage A + B oder eine Assemblage A „Monoinfektion“ festgestellt. Auch wenn zwölf Personen eine kleine Fallzahl darstellt, ist dieses Phänomen von axenischen Kulturen bekannt, die initial als Assemblage B genotypisiert wurden, während später die Assemblage B nicht mehr nachweisbar, sondern durch Assemblage A ersetzt worden war (84). Eine mögliche Erklärung hierfür wäre, dass die initial in größerer Menge vorhandene DNA der Assemblage B durch die PCR vermehrt und bei der Genotypisierung detektiert wurde, im Verlauf allerdings Assemblage B von der schneller wachsenden Assemblage A verdrängt wurde (21, 70). Voraussetzung dafür wäre eine simultane Ko-Infektion. Alternativ könnten sich die Personen auch erst im Verlauf mit einer zweiten Assemblage infiziert haben (29). Die erste Hypothese wird von der Tatsache gestützt, dass in diesen Datensätzen in elf von zwölf Fällen der „Wechsel der Assemblage“ zwischen der ersten und der zweiten Probe geschah. Alternativ müsste das bedeuten, dass sich die Personen sehr zeitnah nach der ersten Infektion mit einer anderen Assemblage infizieren hätten, was bei der in Deutschland eher niedrigen Inzidenz der Giardiasis allerdings eher unwahrscheinlich erscheint (86). In der Mehrzahl der Fälle wurde zunächst die langsamer reproduzierende Assemblage B vor der schneller wachsenden Assemblage A detektiert. Dies könnte an einer höheren initialen infektiösen Dosis durch die häufigere Assemblage B liegen. Möglich wäre aber auch eine unterschiedliche Empfindlichkeit der Assemblagen gegenüber der verbreitet eingesetzten Therapie mit Metronidazol (87-89). Nutzt man die weltweit häufig angewendete Sanger-Sequenzierung zur Genotypisierung von *G. duodenalis*, sollte man sich auch der methodischen Limitationen bewusst sein. Für HIV ist beispielsweise belegt, dass im Gegensatz zum NGS (engl. *next generation sequencing*) durch die Sanger-Sequenzierung nur SNP-Positionen erfasst werden, die in mehr als 20-35 % der amplifizierten DNA-Stränge auftreten (90, 91). Ob diese Ergebnisse auf tetraploide Organismen wie *G. duodenalis* übertragbar sind, ist unklar. Gerade deshalb ist die Untersuchung der identifizierten SNP-Positionen von großer Wichtigkeit. Leider besteht in großen Datenbanken eine beträchtliche Verzerrung hin zu Sequenzen, die weniger oder keine ASH enthalten: Tatsächlich fand sich in 79,1 % der 738 untersuchten Assemblage B TPI Sequenzen aus der Genbank keine ASH. Auf dem gleichen Genabschnitt fand sich dagegen in den Proben dieser Studie in 59,1 % ASH und nur in 40,1 % keine ASH. Drei Faktoren werden meist dafür verantwortlich gemacht, dass die Trennschärfe der MSLT bei Assemblage B nicht zur Identifikation von Subgruppen ausreicht: i) aufgrund der zwei Zellkerne und möglicher sexueller oder parasexueller Rekombinationsstrategien entstandene ASH (44, 45, 92, 93), ii) Fehler in der PCR oder in der Sequenzierung (94) und iii) gemischte Infektionen mit verschiedenen Assemblagen oder – seltener



diskutiert – mit verschiedenen Subassemblagen (10, 29, 66, 67). In dieser Arbeit wurde gezeigt, dass die manuelle Annotation der DNA-Sequenzen zwar zeitaufwändig ist, aber zu reproduzierbaren ASH-Ergebnissen führt (52). ASH trägt also sicherlich dazu bei, dass sich epidemiologisch verbundene Isolate und Subgruppen schlechter identifizieren lassen, ist aber nicht allein dafür verantwortlich. Fehler in der PCR, in der Sequenzierung und auch Artefakte, wie die für andere Spezies beschriebene, durch die PCR erzeugte Mutagenese (95, 96), können sicherlich in Einzelfällen auftreten. Diese Einzelfälle werden jedoch wahrscheinlich nicht die Erkennung von Subgruppen verhindern, da Typisierungsergebnisse und ASH in dieser Studie zu gut reproduzierbar waren (52). Zwei Forschungsarbeiten haben sich mit der Subklonierung von Assemblage B Isolaten mit ASH beschäftigt und dabei für das Genfragment *GDH* bis zu drei und für den Locus *BG* bis zu sieben verschiedene Sequenzen aus einer Stuhlprobe (häufig als „ein Isolat“ bezeichnet) identifiziert (82, 94). Diese (82, 94) und auch die vorliegende Arbeit (52) zeigen somit polyklonale Infektionen als plausibelste Erklärung für die erschwerte Subklassifikation von Assemblage B in die Subassemblagen BIII und BIV (65, 68, 73) und damit auch die erschwerte Identifikation von epidemiologisch verbundenen Isolaten. Während sich vorangehende Studien in Deutschland hauptsächlich auf Reise-assoziierte Infektionen konzentriert haben, umfasst diese Arbeit auch Daten zu autochthonen *G. duodenalis* Infektionen (97-99). In dieser Studie hatten ca. 50 % der Infektionen einen Zusammenhang zu Auslandsaufenthalten, wohingegen die anderen 50 % als autochthon klassifiziert wurden, was den für Deutschland gemeldeten Daten entspricht (100). Global berichten die meisten Studien, davon, dass Assemblage B überwiegt (23, 24, 34, 65, 97). Aufgrund der Tatsache, dass Assemblage B mit etwas über 65 % ebenfalls deutlich häufiger als Assemblage A mit knapp 15 % und gemischte A + B Infektionen mit ca. 20 % vorkam, konnte dieser Trend für Deutschland bestätigt werden. In Krankheitsausbrüchen ist allerdings eine andere Assemblage-Verteilung möglich (65). Auch wenn in Einzelfällen (101) DNA der wirtsspezifischen *G. duodenalis* Assemblagen C-H in Menschen nachgewiesen wurden, war dies in der vorliegenden Studie nicht der Fall. Es wurde vorgeschlagen, dass Infektionen mit Subassemblage AI mit einer möglichen zoonotischen Transmission und Infektionen mit Subassemblage AII und Assemblage B eher mit anthroponotischer Transmission verbunden seien (16, 29, 65). Das würde bedeuten, dass eine zoonotische Transmission in Deutschland, wie auch in anderen „Industrienationen“ eine untergeordnete Rolle spielt (18, 24, 34, 102). Auch wenn es keine signifikanten Unterschiede bezüglich Altersverteilung, Geschlechterverteilung oder Assemblage zwischen den Gruppen der autochthonen und Reise-assoziierten Infektionen gab, zeigte sich doch ein Trend hin zu einem prozentual größeren Anteil an Assemblage A Infektionen in den vermutlich autochthonen erworbenen Infektionen. Für

Unterschiede zwischen den beiden Gruppen spricht auch eine Forschungsarbeit, die die Risikofaktoren für Giardiasis in Deutschland untersuchte. Dort gab es signifikante Unterschiede zwischen den beiden Gruppen, so korrelierte z.B. Tierhaltung mit autochthonen Infektionen (17). Jedoch wurden in der genannten Arbeit nicht die zugrundeliegenden Assemblagen untersucht (17). Im Gegensatz zu diesem Datensatz gibt es bei den auf ganz Deutschland bezogenen Daten eine zweigipflige Altersverteilung mit einem Peak bei Kindern zwischen ein und neun Jahren und bei Erwachsenen zwischen 25 und 49 Jahren. Außerdem sind mehr Männer als Frauen betroffen (86). Die Frage, ob es sich bei „autochthonen Infektionen“ tatsächlich um Infektionen durch eine bislang unbekannte Infektionsquelle in Deutschland handelt oder diese eher eine von den Gesundheitsämtern unentdeckte, sekundäre Transmission von im Ausland erworbenen Isolaten widerspiegeln, konnte in der vorliegenden Arbeit nicht abschließend beantwortet werden. Hierzu wäre eine prospektive epidemiologische Studie mit entsprechender Datenerhebung, inklusive Genotypisierung der Parasiten nötig. Zusammenfassend lässt sich sagen, dass in dieser Dissertation die Stärken und Schwächen der aktuell verfügbaren Genotypisierungsverfahren für *G. duodenalis* herausgearbeitet wurden. Das Problem von ASH und polyklonalen Infektionen wurde dargestellt und es wurde aufgezeigt, wie man trotz dieser Herausforderungen epidemiologisch verbundene Isolate identifizieren kann. Zukünftige Fortschritte in der Methodik der Sequenzgenerierung, z. B. verbesserte MLST-Schemata, und/oder Neuerungen im Bereich der Ergebnisinterpretation werden zu einer verbesserten Genotypisierung von *G. duodenalis* führen und dadurch zur Assoziation von Assemblagen zu Symptomen, akuten oder chronischen Infektionen, Langzeitfolgen einer Infektion und der Interaktion mit anderen Pathogenen beitragen.

## Literaturverzeichnis

1. Currie SL, Stephenson N, Palmer AS, Jones BL, Hawkins G, Alexander CL. Under-reporting giardiasis: time to consider the public health implications. *Epidemiol Infect.* 2017;145(14):3007-11.
2. Ankarklev J, Jerlstrom-Hultqvist J, Ringqvist E, Troell K, Svard SG. Behind the smile: cell biology and disease mechanisms of *Giardia* species. *Nat Rev Microbiol.* 2010;8(6):413-22.
3. Sulaiman IM, Fayer R, Bern C, Gilman RH, Trout JM, Schantz PM, Das P, Lal AA, Xiao L. Triosephosphate isomerase gene characterization and potential zoonotic transmission of *Giardia duodenalis*. *Emerg Infect Dis.* 2003;9(11):1444-52.
4. Torgerson PR, Devleeschauwer B, Praet N, Speybroeck N, Willingham AL, Kasuga F, Rokni MB, Zhou XN, Fevre EM, Sripa B, Gargouri N, Furst T, Budke CM, Carabin H, Kirk MD, Angulo FJ, Havelaar A, de Silva N. World Health Organization Estimates of the Global and Regional Disease Burden of 11 Foodborne Parasitic Diseases, 2010: A Data Synthesis. *PLoS Med.* 2015;12(12):e1001920.

5. Kirk MD, Pires SM, Black RE, Caipo M, Crump JA, Devleeschauwer B, Dopfer D, Fazil A, Fischer-Walker CL, Hald T, Hall AJ, Keddy KH, Lake RJ, Lanata CF, Torgerson PR, Havelaar AH, Angulo FJ. World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. *PLoS Med.* 2015;12(12):e1001921.
6. Pires SM, Fischer-Walker CL, Lanata CF, Devleeschauwer B, Hall AJ, Kirk MD, Duarte ASR, Black RE, Angulo FJ. Aetiology-Specific Estimates of the Global and Regional Incidence and Mortality of Diarrhoeal Diseases Commonly Transmitted through Food. *PLOS ONE.* 2015;10(12):e0142927.
7. Pires SM, Fischer-Walker CL, Lanata CF, Devleeschauwer B, Hall AJ, Kirk MD, Duarte AS, Black RE, Angulo FJ. Aetiology-Specific Estimates of the Global and Regional Incidence and Mortality of Diarrhoeal Diseases Commonly Transmitted through Food. *PLoS One.* 2015;10(12):e0142927.
8. Buret AG, Cotton J. Pathophysiological Processes and Clinical Manifestations of Giardiasis. In: Luján HD, Svärd S, editors. *Giardia: A Model Organism.* Vienna: Springer Vienna; 2011:301-18.
9. Helmy YA, Klotz C, Wilking H, Krucken J, Nockler K, Von Samson-Himmelstjerna G, Zessin KH, Aebischer T. Epidemiology of *Giardia duodenalis* infection in ruminant livestock and children in the Ismailia province of Egypt: insights by genetic characterization. *Parasites & vectors.* 2014;7:321.
10. Franzen O, Jerlstrom-Hultqvist J, Castro E, Sherwood E, Ankarklev J, Reiner DS, Palm D, Andersson JO, Andersson B, Svärd SG. Draft genome sequencing of *giardia intestinalis* assemblage B isolate GS: is human giardiasis caused by two different species? *PLoS Pathog.* 2009;5(8):e1000560.
11. Savioli L, Smith H, Thompson A. *Giardia* and *Cryptosporidium* join the 'Neglected Diseases Initiative'. *Trends Parasitol.* 2006;22(5):203-8.
12. Sprong H, Caccio SM, van der Giessen JW, network Z, partners. Identification of zoonotic genotypes of *Giardia duodenalis*. *PLoS neglected tropical diseases.* 2009;3(12):e558.
13. Roxström-Lindquist K, Palm D, Reiner D, Ringqvist E, Svärd SG. *Giardia* immunity – an update. *Trends in Parasitology.* 2006;22(1):26-31.
14. Ryan U, Hijjawi N, Feng Y, Xiao L. *Giardia*: an under-reported foodborne parasite. *Int J Parasitol.* 2019;49(1):1-11.
15. Hunter PR, Thompson RC. The zoonotic transmission of *Giardia* and *Cryptosporidium*. *International journal for parasitology.* 2005;35(11-12):1181-90.
16. Pijnacker R, Mughini-Gras L, Heusinkveld M, Roelfsema J, van Pelt W, Kortbeek T. Different risk factors for infection with *Giardia lamblia* assemblages A and B in children attending day-care centres. *European journal of clinical microbiology & infectious diseases* : official publication of the European Society of Clinical Microbiology. 2016;35(12):2005-13.
17. Espelage W, an der Heiden M, Stark K, Alpers K. Characteristics and risk factors for symptomatic *Giardia lamblia* infections in Germany. *BMC public health.* 2010;10:41.
18. Reses HE, Gargano JW, Liang JL, Cronquist A, Smith K, Collier SA, Roy SL, Vanden Eng J, Bogard A, Lee B, Hlavsa MC, Rosenberg ES, Fullerton KE, Beach MJ, Yoder JS. Risk factors for sporadic *Giardia* infection in the USA: a case-control study in Colorado and Minnesota. *Epidemiol Infect.* 2018;146(9):1071-8.
19. Swirski AL, Pearl DL, Peregrine AS, Pintar K. A comparison of exposure to risk factors for giardiasis in non-travellers, domestic travellers and international travellers in a Canadian community, 2006-2012. *Epidemiol Infect.* 2016;144(5):980-99.

20. Hoque ME, Hope VT, Kjellström T, Scragg R, Lay-Yee R. Risk of giardiasis in Aucklanders: a case-control study. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases*. 2002;6(3):191-7.
21. Robert Koch-Institut. *Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2019*, 20th ed. Berlin, RKI; 2020. p. 85-88.
22. Ignatius R, Gahutu JB, Klotz C, Musemakweri A, Aebischer T, Mockenhaupt FP. Detection of *Giardia duodenalis* assemblage A and B isolates by immunochromatography in stool samples from Rwandan children. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. 2014;20(10):O783-5.
23. Caccio SM, Lalle M, Svard SG. Host specificity in the *Giardia duodenalis* species complex. *Infect Genet Evol*. 2018;66:335-45.
24. Feng Y, Xiao L. Zoonotic potential and molecular epidemiology of *Giardia* species and giardiasis. *Clin Microbiol Rev*. 2011;24(1):110-40.
25. Ankarklev J, Lebbad M, Einarsson E, Franzen O, Ahola H, Troell K, Svard SG. A novel high-resolution multilocus sequence typing of *Giardia intestinalis* Assemblage A isolates reveals zoonotic transmission, clonal outbreaks and recombination. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases*. 2018;60:7-16.
26. Monis PT, Caccio SM, Thompson RC. Variation in *Giardia*: towards a taxonomic revision of the genus. *Trends Parasitol*. 2009;25(2):93-100.
27. Caccio SM, Thompson RC, McLauchlin J, Smith HV. Unravelling *Cryptosporidium* and *Giardia* epidemiology. *Trends Parasitol*. 2005;21(9):430-7.
28. Ryan U, Zahedi A. Molecular epidemiology of giardiasis from a veterinary perspective. *Advances in parasitology*. 2019;106:209-54.
29. Sprong H, Caccio SM, van der Giessen JW. Identification of zoonotic genotypes of *Giardia duodenalis*. *PLoS Negl Trop Dis*. 2009;3(12):e558.
30. Read CM, Monis PT, Thompson RC. Discrimination of all genotypes of *Giardia duodenalis* at the glutamate dehydrogenase locus using PCR-RFLP. *Infect Genet Evol*. 2004;4(2):125-30.
31. Lalle M, Pozio E, Capelli G, Bruschi F, Crotti D, Caccio SM. Genetic heterogeneity at the beta-giardin locus among human and animal isolates of *Giardia duodenalis* and identification of potentially zoonotic subgenotypes. *International journal for parasitology*. 2005;35(2):207-13.
32. Bertrand I, Albertini L, Schwartzbrod J. Comparison of two target genes for detection and genotyping of *Giardia lamblia* in human feces by PCR and PCR-restriction fragment length polymorphism. *Journal of clinical microbiology*. 2005;43(12):5940-4.
33. Heyworth MF. *Giardia duodenalis* genetic assemblages and hosts. *Parasite (Paris, France)*. 2016;23:13.
34. Cai W, Ryan U, Xiao L, Feng Y. Zoonotic giardiasis: an update. *Parasitol Res* 120. 2021:4199–4218.
35. Ankarklev J, Svärd S, Lebbad M. Allelic sequence heterozygosity in single *Giardia* parasites. *BMC microbiology*. 2012: 65.
36. Almeida A, Pozio E, Caccio SM. Genotyping of *Giardia duodenalis* cysts by new real-time PCR assays for detection of mixed infections in human samples. *Applied and environmental microbiology*. 2010;76(6):1895-901.
37. Adam EA, Yoder JS, Gould LH, Hlavsa MC, Gargano JW. Giardiasis outbreaks in the United States, 1971-2011. *Epidemiol Infect*. 2016;144(13):2790-801.
38. Collier SA, Deng L, Adam EA, Benedict KM, Beshearse EM, Blackstock AJ, Bruce BB, Derado G, Edens C, Fullerton KE, Gargano JW, Geissler AL, Hall AJ, Havelaar AH, Hill

- VR, Hoekstra RM, Reddy SC, Scallan E, Stokes EK, Yoder JS, Beach MJ. Estimate of Burden and Direct Healthcare Cost of Infectious Waterborne Disease in the United States. *Emerg Infect Dis.* 2021;27(1):140-9.
39. Yoder JS, Gargano JW, Wallace RM, Beach MJ. Giardiasis surveillance--United States, 2009-2010. *Morbidity and mortality weekly report Surveillance summaries (Washington, DC : 2002).* 2012;61(5):13-23.
40. Ryan U, Caccio SM. Zoonotic potential of Giardia. *International journal for parasitology.* 2013;43(12):943-56.
41. Jerlstrom-Hultqvist J, Ankarklev J, Svard SG. Is human giardiasis caused by two different Giardia species? *Gut microbes.* 2010;1(6):379-82.
42. Adam RD, Dahlstrom EW, Martens CA, Bruno DP, Barbian KD, Ricklefs SM, Hernandez MM, Narla NP, Patel RB, Porcella SF, Nash TE. Genome sequencing of Giardia lamblia genotypes A2 and B isolates (DH and GS) and comparative analysis with the genomes of genotypes A1 and E (WB and Pig). *Genome Biol Evol.* 2013;5(12):2498-511.
43. Xu F, Jex A, Svard SG. A chromosome-scale reference genome for Giardia intestinalis WB. *Sci Data.* 2020;7(1):38.
44. Birky CW. Sex: Is Giardia Doing It in the Dark? *Current Biology.* 2005;15(2):56-58.
45. Birky CW, Jr. Giardia sex? Yes, but how and how much? *Trends Parasitol.* 2010;26(2):70-74.
46. Ramesh MA, Malik S-B, Logsdon JM. A Phylogenomic Inventory of Meiotic Genes: Evidence for Sex in Giardia and an Early Eukaryotic Origin of Meiosis. *Current Biology.* 2005;15(2):185-91.
47. Caccio SM, Beck R, Lalle M, Marinculic A, Pozio E. Multilocus genotyping of Giardia duodenalis reveals striking differences between assemblages A and B. *International journal for parasitology.* 2008;38(13):1523-31.
48. Tsui CK, Miller R, Uyaguari-Diaz M, Tang P, Chauve C, Hsiao W, Isaac-Renton J, Prystajeky N. Beaver Fever: Whole-Genome Characterization of Waterborne Outbreak and Sporadic Isolates To Study the Zoonotic Transmission of Giardiasis. *mSphere.* 2018;3(2):e00090-18.
49. Coronato Nunes B, Pavan MG, Jaeger LH, Monteiro KJ, Xavier SC, Monteiro FA, Boia MN, Carvalho-Costa FA. Spatial and Molecular Epidemiology of Giardia intestinalis Deep in the Amazon, Brazil. *PLoS One.* 2016;11(7):e0158805.
50. Lecova L, Tumova P, Nohynkova E. Clone-based haplotyping of Giardia intestinalis assemblage B human isolates. *Parasitol Res.* 2019;118(1):355-61.
51. Mizuno T, Matey EJ, Bi X, Songok EM, Ichimura H, Tokoro M. Extremely diversified haplotypes observed among assemblage B population of Giardia intestinalis in Kenya. *Parasitol Int.* 2019:102038.
52. Woschke A, Faber M, Stark K, Holtfreter M, Mockenhaupt F, Richter J, Regnath T, Sobottka I, Reiter-Owona I, Diefenbach A, Gosten-Heinrich P, Friesen J, Ignatius R, Aebischer T, Klotz C. Suitability of current typing procedures to identify epidemiologically linked human Giardia duodenalis isolates. *PLoS Negl Trop Dis.* 2021;15(3):e0009277.
53. Verweij JJ, Blange RA, Templeton K, Schinkel J, Brienen EA, van Rooyen MA, van Lieshout L, Polderman AM. Simultaneous detection of Entamoeba histolytica, Giardia lamblia, and Cryptosporidium parvum in fecal samples by using multiplex real-time PCR. *Journal of clinical microbiology.* 2004;42(3):1220-3.
54. Verweij JJ, Schinkel J, Laeijendecker D, van Rooyen MA, van Lieshout L, Polderman AM. Real-time PCR for the detection of Giardia lamblia. *Molecular and cellular probes.* 2003;17(5):223-5.

55. Deer DM, Lampel KA, Gonzalez-Escalona N. A versatile internal control for use as DNA in real-time PCR and as RNA in real-time reverse transcription PCR assays. *Letters in applied microbiology*. 2010;50(4):366-72.
56. Van Lith L, Soba B, Vizcaino VV, Svard S, Sprong H, Tosini F, Pozio E, Caccio SM. A real-time assemblage-specific PCR assay for the detection of *Giardia duodenalis* assemblages A, B and E in fecal samples. *Veterinary parasitology*. 2015;211(1-2):28-34.
57. Johnson AD. An extended IUPAC nomenclature code for polymorphic nucleic acids. *Bioinformatics (Oxford, England)*. 2010;26(10):1386-9.
58. Fahmy HM, El-Serougi AO, El Deeb HK, Hussein HM, Abou-Seri HM, Klotz C, Aebischer T, El Sayed Khalifa Mohamed K. *Giardia duodenalis* assemblages in Egyptian children with diarrhea. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology*. 2015;34(8):1573-81.
59. GenBank ®, NIH genetic sequence database; <https://www.ncbi.nlm.nih.gov/genbank/>. Accessed January 06, 2022.
60. Brynildsrud O, Tysnes KR, Robertson LJ, Debenham JJ. *Giardia duodenalis* in primates: Classification and host specificity based on phylogenetic analysis of sequence data. *Zoonoses and public health*. 2018;65(6):637-47.
61. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*. 2003;52(5):696-704.
62. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol*. 2012;29(8):1969-73.
63. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol*. 2016;33(7):1870-4.
64. Bender R, Lange S. [What is the p-value?]. *Dtsch Med Wochenschr*. 2007;132 Suppl 1:e15-6.
65. Ryan U, Caccio SM. Zoonotic potential of *Giardia*. *International journal for parasitology*. 2013;43(12-13):943-56.
66. Caccio SM, Ryan U. Molecular epidemiology of giardiasis. *Mol Biochem Parasitol*. 2008;160(2):75-80.
67. Levecke B, Geldhof P, Claerebout E, Dorny P, Vercammen F, Caccio SM, Vercruyssen J, Geurden T. Molecular characterisation of *Giardiaduodenalis* in captive non-human primates reveals mixed assemblage A and B infections and novel polymorphisms. *International journal for parasitology*. 2009;39(14):1595-601.
68. Wielinga C, Ryan U, Andrew Thompson RC, Monis P. Multi-locus analysis of *Giardia duodenalis* intra-Assemblage B substitution patterns in cloned culture isolates suggests sub-Assemblage B analyses will require multi-locus genotyping with conserved and variable genes. *International journal for parasitology*. 2011;41(5):495-503.
69. Karim MR, Wang R, Yu F, Li T, Dong H, Li D, Zhang L, Li J, Jian F, Zhang S, Rume FI, Ning C, Xiao L. Multi-locus analysis of *Giardia duodenalis* from nonhuman primates kept in zoos in China: Geographical segregation and host-adaptation of assemblage B isolates. *Infection, Genetics and Evolution*. 2015;30:82-8.
70. Teodorovic S, Braverman JM, Elmendorf HG. Unusually low levels of genetic variation among *Giardia lamblia* isolates. *Eukaryot Cell*. 2007;6(8):1421-30.
71. Garcia-R JC, French N, Pita A, Velathanthiri N, Shrestha R, Hayman D. Local and global genetic diversity of protozoan parasites: Spatial distribution of *Cryptosporidium* and *Giardia* genotypes. *PLoS neglected tropical diseases*. 2017;11(7):e0005736-e.
72. Wielinga CM, Thompson RC. Comparative evaluation of *Giardia duodenalis* sequence data. *Parasitology*. 2007;134(12):1795-821.

73. Seabolt MH, Konstantinidis KT, Roellig DM. Hidden Diversity within Common Protozoan Parasites as Revealed by a Novel Genotyping Scheme. *Applied and environmental microbiology*. 2021;87(6): e02275-20.
74. Merker M, Kohl TA, Niemann S, Supply P. The Evolution of Strain Typing in the *Mycobacterium tuberculosis* Complex. In: Gagneux S, editor. *Strain Variation in the Mycobacterium tuberculosis Complex: Its Role in Biology, Epidemiology and Control*. Cham: Springer International Publishing; 2017. p. 43-78.
75. Meehan CJ, Moris P, Kohl TA, Pecerska J, Akter S, Merker M, Utpatel C, Beckert P, Gehre F, Lempens P, Stadler T, Kaswa MK, Kuhnert D, Niemann S, de Jong BC. The relationship between transmission time and clustering methods in *Mycobacterium tuberculosis* epidemiology. *EBioMedicine*. 2018;37:410-6.
76. Meehan CJ, Goig GA, Kohl TA, Verboven L, Dippenaar A, Ezewudo M, Farhat MR, Guthrie JL, Laukens K, Miotto P, Ofori-Anyinam B, Dreyer V, Supply P, Suresh A, Utpatel C, van Soolingen D, Zhou Y, Ashton PM, Brites D, Cabibbe AM, de Jong BC, de Vos M, Menardo F, Gagneux S, Gao Q, Heupink TH, Liu Q, Loiseau C, Rigouts L, Rodwell TC, Tagliani E, Walker TM, Warren RM, Zhao Y, Zignol M, Schito M, Gardy J, Cirillo DM, Niemann S, Comas I, Van Rie A. Whole genome sequencing of *Mycobacterium tuberculosis*: current standards and open issues. *Nat Rev Microbiol*. 2019;17(9):533-45.
77. Vermeulen ET, Ashworth DL, Eldridge MD, Power ML. Investigation into potential transmission sources of *Giardia duodenalis* in a threatened marsupial (*Petrogale penicillata*). *Infect Genet Evol*. 2015;33:277-80.
78. Rafiei A, Baghlaninezhad R, Koster PC, Bailo B, Hernandez de Mingo M, Carmena D, Panabad E, Beiromvand M. Multilocus genotyping of *Giardia duodenalis* in Southwestern Iran. A community survey. *PLoS One*. 2020;15(2):e0228317.
79. Capewell P, Krumrie S, Katzer F, Alexander CL, Weir W. Molecular Epidemiology of *Giardia* Infections in the Genomic Era. *Trends Parasitol*. 2021;37(2):142-53.
80. Thompson RCA, Ash A. Molecular epidemiology of *Giardia* and *Cryptosporidium* infections. *Infect Genet Evol*. 2016;40:315-23.
81. Hussein AIA, Yamaguchi T, Nakamoto K, Iseki M, Tokoro M. Multiple-subgenotype infections of *Giardia intestinalis* detected in Palestinian clinical cases using a subcloning approach. *Parasitology International*. 2009;58(3):258-62.
82. Geurden T, Geldhof P, Levecke B, Martens C, Berkvens D, Casaert S, Vercruyse J, Claerebout E. Mixed *Giardia duodenalis* assemblage A and E infections in calves. *International journal for parasitology*. 2008;38(2):259-64.
83. Bénére E, Geurden T, Robertson L, Van Assche T, Cos P, Maes L. Infectivity of *Giardia duodenalis* Assemblages A and E for the gerbil and axenisation of duodenal trophozoites. *Parasitology International*. 2010;59(4):634-7.
84. Ankarklev J, Franzen O, Peirasmaki D, Jerlstrom-Hultqvist J, Lebbad M, Andersson J, Andersson B, Svard SG. Comparative genomic analyses of freshly isolated *Giardia intestinalis* assemblage A isolates. *BMC Genomics*. 2015;16:697.
85. Benere E, T VANA, Cos P, Maes L. Variation in growth and drug susceptibility among *Giardia duodenalis* assemblages A, B and E in axenic in vitro culture and in the gerbil model. *Parasitology*. 2011;138(11):1354-61.
86. Arguello-Garcia R, Leitsch D, Skinner-Adams T, Ortega-Pierres MG. Drug resistance in *Giardia*: Mechanisms and alternative treatments for Giardiasis. *Advances in parasitology*. 2020;107:201-82.
87. Fantinatti M, Lopes-Oliveira LAP, Cascais-Figueroa T, Austriaco-Teixeira P, Verissimo E, Bello AR, Da-Cruz AM. Recirculation of *Giardia lamblia* Assemblage A After

- Metronidazole Treatment in an Area With Assemblages A, B, and E Sympatric Circulation. *Frontiers in Microbiology*. 2020;11:571104.
88. Mohamed S, Penaranda G, Gonzalez D, Camus C, Khiri H, Boulme R, Sayada C, Philibert P, Olive D, Halfon P. Comparison of ultra-deep versus Sanger sequencing detection of minority mutations on the HIV-1 drug resistance interpretations after virological failure. *AIDS*. 2014;28(9):1315-24.
  89. Palmer S, Kearney M, Maldarelli F, Halvas EK, Bixby CJ, Bazmi H, Rock D, Falloon J, Davey RT, Jr., Dewar RL, Metcalf JA, Hammer S, Mellors JW, Coffin JM. Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. *Journal of clinical microbiology*. 2005;43(1):406-13.
  90. Lasek-Nesselquist E, Welch DM, Thompson RC, Steuart RF, Sogin ML. Genetic exchange within and between assemblages of *Giardia duodenalis*. *The Journal of eukaryotic microbiology*. 2009;56(6):504-18.
  91. Ramesh MA, Malik SB, Logsdon JM, Jr. A phylogenomic inventory of meiotic genes; evidence for sex in *Giardia* and an early eukaryotic origin of meiosis. *Current biology : CB*. 2005;15(2):185-91.
  92. Lebbad M, Ankarklev J, Tellez A, Leiva B, Andersson JO, Svard S. Dominance of *Giardia* assemblage B in Leon, Nicaragua. *Acta tropica*. 2008;106(1):44-53.
  93. Beser J, Hagblom P, Fernandez V. Frequent in vitro recombination in internal transcribed spacers 1 and 2 during genotyping of *Pneumocystis jirovecii*. *Journal of clinical microbiology*. 2007;45(3):881-6.
  94. Tanabe K, Sakihama N, Färnert A, Rooth I, Björkman A, Walliker D, Ranford-Cartwright L. In vitro recombination during PCR of *Plasmodium falciparum* DNA: a potential pitfall in molecular population genetic analysis. *Molecular & Biochemical Parasitology*. 2002;122:211-6.
  95. Broglia A, Weitzel T, Harms G, Caccio SM, Nockler K. Molecular typing of *Giardia duodenalis* isolates from German travellers. *Parasitol Res*. 2013;112(10):3449-56.
  96. Sagebiel D, Weitzel T, Stark K, Leitmeyer K. Giardiasis in kindergartens: prevalence study in Berlin, Germany, 2006. *Parasitol Res*. 2009;105(3):681-7.
  97. Jelinek T, Loscher T. Epidemiology of giardiasis in German travelers. *Journal of travel medicine*. 2000;7(2):70-3.
  98. RKI. Epidemiologisches Bulletin des RKI. 2019;48.  
[https://edoc.rki.de/bitstream/handle/176904/6752/48\\_19.pdf?sequence=1&isAllowed=y](https://edoc.rki.de/bitstream/handle/176904/6752/48_19.pdf?sequence=1&isAllowed=y)  
 Accessed January 06, 2022.
  99. Soliman RH, Fuentes I, Rubio JM. Identification of a novel Assemblage B subgenotype and a zoonotic Assemblage C in human isolates of *Giardia intestinalis* in Egypt. *Parasitol Int*. 2011;60(4):507-11.
  100. Rehbein S, Klotz C, Ignatius R, Muller E, Aebischer A, Kohn B. *Giardia duodenalis* in small animals and their owners in Germany: A pilot study. *Zoonoses and public health*. 2019;66(1):117-24.



## Eidesstattliche Versicherung

„Ich, Andreas Woschke, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: „Anwendung aktueller Typisierungsverfahren zur Identifikation epidemiologisch verbundener Giardia duodenalis Isolate“ oder auf Englisch „Identification of epidemiologically linked Giardia duodenalis isolates using current typing procedures“ selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe. Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren/innen beruhen, sind als solche in korrekter Zitierung kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) werden von mir verantwortet.

Ich versichere ferner, dass ich die in Zusammenarbeit mit anderen Personen generierten Daten, Datenauswertungen und Schlussfolgerungen korrekt gekennzeichnet und meinen eigenen Beitrag sowie die Beiträge anderer Personen korrekt kenntlich gemacht habe (siehe Anteilserklärung). Texte oder Textteile, die gemeinsam mit anderen erstellt oder verwendet wurden, habe ich korrekt kenntlich gemacht.

Meine Anteile an etwaigen Publikationen zu dieser Dissertation entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Erstbetreuer/in, angegeben sind. Für sämtliche im Rahmen der Dissertation entstandenen Publikationen wurden die Richtlinien des ICMJE (International Committee of Medical Journal Editors; [www.icmje.org](http://www.icmje.org)) zur Autorenschaft eingehalten. Ich erkläre ferner, dass ich mich zur Einhaltung der Satzung der Charité – Universitätsmedizin Berlin zur Sicherung Guter Wissenschaftlicher Praxis verpflichte. Weiterhin versichere ich, dass ich diese Dissertation weder in gleicher noch in ähnlicher Form bereits an einer anderen Fakultät eingereicht habe.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§§156, 161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

Datum

Unterschrift

## **Anteilerklärung an den erfolgten Publikationen**

**Konzept:** Mirko Faber, Klaus Stark, Ralf Ignatius, Toni Aebischer, Christian Klotz

**Datenkuration:** Andreas Woschke, Mirko Faber, Christian Klotz.

**Datenanalyse:** Andreas Woschke, Mirko Faber, Christian Klotz.

**Untersuchungen:** Andreas Woschke, Mirko Faber, Martha Holtfreter, Frank Mockenhaupt, Joachim Richter, Thomas Regnath, Ingo Sobottka, Ingrid Reiter-Owona, Andreas Diefenbach, Petra Gosten-Heinrich, Johannes Friesen, Christian Klotz.

**Methodik:** Andreas Woschke, Petra Gosten-Heinrich, Christian Klotz.

**Ressourcen:** Martha Holtfreter, Frank Mockenhaupt, Joachim Richter, Thomas Regnath, Ingo Sobottka, Ingrid Reiter-Owona, Andreas Diefenbach, Johannes Friesen, Ralf Ignatius.

**Supervision:** Ralf Ignatius, Toni Aebischer, Christian Klotz. Validation: Christian Klotz.

**Visualisierung:** Andreas Woschke, Mirko Faber, Christian Klotz.

**Erstellen des Manuskripts:** Andreas Woschke, Toni Aebischer, Christian Klotz.

**Überarbeitung des Manuskripts:** Andreas Woschke, Mirko Faber, Klaus Stark, Martha Holtfreter, Frank Mockenhaupt, Joachim Richter, Thomas Regnath, Ingo Sobottka, Ingrid Reiter-Owona, Johannes Friesen, Ralf Ignatius, Toni Aebischer, Christian Klotz.

Andreas Woschke hatte folgenden Anteil an der folgenden Publikation: Woschke A, Faber M, Stark K, Holtfreter M, Mockenhaupt F, Richter J, Regnath T, Sobottka I, Reiter-Owona I, Diefenbach A, Gosten-Heinrich P, Friesen J, Ignatius R, Aebischer T, Klotz C. Suitability of current typing procedures to identify epidemiologically linked human *Giardia duodenalis* isolates. PLoS Negl Trop Dis. 2021;15(3):e0009277.

Kommunikation mit Kliniken und Laboren in Fragen des Probenversandes. Durchführung der Laborarbeiten nach Einarbeitung: Durchführung fast aller PCRs für das ursprüngliche MLST-Schema, sämtlicher PCRs für die Genloci, die von Ankarklev et al. 2018 beschrieben wurden und sämtlicher qPCRs. Durchführung der Datenanalyse in diversen Programmen wie z.B. Microsoft Excel, Geneious 11.1.5 und MEGA 7.0.26, Einpflegen der Daten in die RKI-interne Datenbank und in die Datenbank der NCBI. Erstellen der Distanzmatrices, Errechnen der individuellen Cut-off Werte für die einzelnen Isolate, statistische Auswertung mit Graph Pad Prism 8.4.0. Verfassen des Manuskripts, Einpflegen von Korrekturen, Erstellen aller Graphiken (außer Figure 3) und Tabellen.

---

Unterschrift, Datum und Stempel des erstbetreuenden Hochschullehrers

---

Unterschrift des Doktoranden/der Doktorandin

## Auszug aus der Journal Summary List

Journal Data Filtered By: **Selected JCR Year: 2020** Selected Editions: SCIE, SSCI

Selected Categories: **“TROPICAL MEDICINE”** Selected Category Scheme: WoS

**Gesamtanzahl: 23 Journale**

Rank	Full Journal Title	Total Cites	Journal ImpactFactor	Eigenfactor Score
1	Infectious Diseases of Poverty	3,434	4.520	0.006920
2	PLoS Neglected Tropical Diseases	35,321	4.411	0.060440
3	Parasites & Vectors	19,634	3.876	0.031090
4	ACTA TROPICA	11,474	3.112	0.012720
5	MALARIA JOURNAL	15,906	2.979	0.023870
6	Pathogens and Global Health	1,504	2.894	0.002530
7	JOURNAL OF VENOMOUS ANIMALS AND TOXINS INCLUDING TROPICAL DISEASES	1,353	2.831	0.001090
8	MEMORIAS DO INSTITUTO OSWALDO CRUZ	7,753	2.743	0.004530
9	TROPICAL MEDICINE & INTERNATIONAL HEALTH	9,155	2.622	0.008390
10	Journal of Tropical Medicine	794	2.488	0.001080
11	AMERICAN JOURNAL OF TROPICAL MEDICINE AND HYGIENE	25,730	2.345	0.025370
12	TRANSACTIONS OF THE ROYAL SOCIETY OF TROPICAL	9,029	2.184	0.004670

	MEDICINE AND HYGIENE			
13	REVISTA DO INSTITUTO DE MEDICINA TROPICAL DE SAO PAULO	2,667	1.846	0.002170
14	JOURNAL OF VECTOR BORNE DISEASES	1,026	1.688	0.001220
15	Revista da Sociedade Brasileira de Medicina Tropical	3,772	1.581	0.003320
16	Asian Pacific Journal of Tropical Biomedicine	4,699	1.545	0.003020
17	Asian Pacific Journal of Tropical Medicine	3,942	1.226	0.003110
18	JOURNAL OF TROPICAL PEDIATRICS	2,268	1.165	0.001830
19	Biomedica	1,278	0.935	0.001280
20	TROPICAL DOCTOR	1,086	0.731	0.000810
21	TROPICAL BIOMEDICINE	1,313	0.623	0.000880
22	LEPROSY REVIEW	989	0.537	0.000610
23	SOUTHEAST ASIAN JOURNAL OF TROPICAL MEDICINE AND PUBLIC HEALTH	3,143	0.267	0.001060

Copyright © 2021 Clarivate Analytics

## RESEARCH ARTICLE

Suitability of current typing procedures to identify epidemiologically linked human *Giardia duodenalis* isolates

Andreas Woschke<sup>1,2</sup>, Mirko Faber<sup>3</sup>, Klaus Stark<sup>3</sup>, Martha Holtfreter<sup>4</sup>, Frank Mockenhaupt<sup>5</sup>, Joachim Richter<sup>4,5</sup>, Thomas Regnath<sup>6</sup>, Ingo Sobottka<sup>7</sup>, Ingrid Reiter-Owona<sup>8</sup>, Andreas Diefenbach<sup>9,10</sup>, Petra Gosten-Heinrich<sup>1</sup>, Johannes Friesen<sup>10</sup>, Ralf Ignatius<sup>2,10</sup>, Toni Aebischer<sup>1</sup>, Christian Klotz<sup>1\*</sup>

**1** Department of Infectious Diseases, Unit for Mycotic and Parasitic Agents and Mycobacteria, Robert Koch Institute, Berlin, Germany, **2** Laboratory of Innate Immunity, Institute of Microbiology, Infectious Diseases and Immunology, Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Campus Benjamin Franklin, Berlin, Germany, **3** Department for Infectious Disease Epidemiology, Gastrointestinal Infections, Zoonoses and Tropical Infections Unit, Robert Koch Institute, Berlin, Germany, **4** Department of Gastroenterology, Hepatology and Infectiology, University Hospital Düsseldorf, Düsseldorf, Germany, **5** Institute of Tropical Medicine and International Health, Charité University Medicine and Berlin Institute of Health, Corporate member of Free University Berlin and Humboldt University Berlin, Berlin, Germany, **6** Laboratory Enders and Partners, Stuttgart, Germany, **7** LADR GmbH, Medizinisches Versorgungszentrum, Geesthacht, Germany, **8** Institute of Medical Microbiology, Immunology and Parasitology (IMMIP), University Clinic Bonn, Germany, **9** Department of Microbiology and Hygiene, Labor Berlin, Charité - Vivantes GmbH, Berlin, Germany, **10** MVZ Labor 28, Berlin, Germany

\* [klotzc@rki.de](mailto:klotzc@rki.de)



## OPEN ACCESS

**Citation:** Woschke A, Faber M, Stark K, Holtfreter M, Mockenhaupt F, Richter J, et al. (2021) Suitability of current typing procedures to identify epidemiologically linked human *Giardia duodenalis* isolates. PLoS Negl Trop Dis 15(3): e0009277. <https://doi.org/10.1371/journal.pntd.0009277>

**Editor:** Renata Rosito Tonelli, Universidade Federal de São Paulo, BRAZIL

**Received:** January 5, 2021

**Accepted:** February 26, 2021

**Published:** March 25, 2021

**Copyright:** © 2021 Woschke et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its [Supporting Information](#) files.

**Funding:** The author(s) received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

## Background

*Giardia duodenalis* is a leading cause of gastroenteritis worldwide. Humans are mainly infected by two different subtypes, i.e., assemblage A and B. Genotyping is hampered by allelic sequence heterozygosity (ASH) mainly in assemblage B, and by occurrence of mixed infections. Here we assessed the suitability of current genotyping protocols of *G. duodenalis* for epidemiological applications such as molecular tracing of transmission chains.

## Methodology/Principal findings

Two *G. duodenalis* isolate collections, from an outpatient tropical medicine clinic and from several primary care laboratories, were characterized by assemblage-specific qPCR (*TIF*, *CATH* gene loci) and a common multi locus sequence typing (MLST; *TPI*, *BG*, *GDH* gene loci). Assemblage A isolates were further typed at additional loci (*HCMP22547*, *CID1*, *RHP26*, *HCMP6372*, *DIS3*, *NEK15411*).

Of 175/202 (86.6%) patients the *G. duodenalis* assemblage could be identified: Assemblages A 25/175 (14.3%), B 115/175 (65.7%) and A+B mixed 35/175 (20.0%). By incorporating allelic sequence heterozygosity in the analysis, the three marker MLST correctly identified 6/9 (66.7%) and 4/5 (80.0%) consecutive samples from chronic assemblage B infections in the two collections, respectively, and identified a cluster of five independent patients carrying assemblage B parasites of identical MLST type. Extended MLST for



assemblage A altogether identified 5/6 (83,3%) consecutive samples from chronic assemblage A infections and 15 novel genotypes. Based on the observed A+B mixed infections it is estimated that only 75% and 50% of assemblage A or B only cases represent single strain infections, respectively. We demonstrate that typing results are consistent with this prediction.

### Conclusions/Significance

Typing of assemblage A and B isolates with resolution for epidemiological applications is possible but requires separate genotyping protocols. The high frequency of multiple infections and their impact on typing results are findings with immediate consequences for result interpretation in this field.

### Author summary

*Giardia duodenalis* is a leading cause of gastroenteritis worldwide. Humans are mainly infected by the two different genetic subtypes, assemblage A and B. Molecular typing tools for epidemiological applications such as tracking transmission, attribution to a source and outbreak investigations have been developed and are highly desirable. However, to what degree the tetraploid genome with allelic sequence heterogeneity (ASH), and the frequent occurrence of mixed, assemblage A and B infections hamper performance is unclear. Here, we assessed the suitability of current genotyping protocols for deciphering the molecular epidemiology of *G. duodenalis*. Against a common reporting bias, we incorporated ASH in the analysis and we show that typing with resolution for epidemiological applications is possible for both, assemblage A and B isolates, but requires separate protocols. We also demonstrate how the high frequency of multiple infections overall impacts on typing results, which has immediate consequences for result interpretation in this field.

### Introduction

The protozoan parasite *Giardia duodenalis* (syn. *G. intestinalis*, *G. lamblia*) is a leading cause of diarrheal disease and a relevant public health problem [1]. Laboratory diagnostics of *G. duodenalis* are well established, but these methods are inadequate for parasite genotyping, thus molecular epidemiological tracing of transmission chains, source attribution and outbreak investigations are hindered [1–3]. Current diagnostics cannot discriminate the eight genetically distinguishable assemblages within the *G. duodenalis* species complex, with assemblages A and B being the major two of these causing human disease [1]. In rare cases, assemblage E that is usually found in ruminants has also been associated with human disease [1]. Importantly, due to the lack of adequate *in vitro* culture systems genotyping analyses are mainly performed on parasite cysts in faecal samples by bulk DNA preparation of enriched cysts or of total stool samples. Such DNA samples are often referred to as isolates [4]. Interpretation of genotyping results is highly complex for several reasons: *Giardia* parasites are binucleated and possess tetraploid genomes with different grades of allelic sequence heterozygosity (ASH). In assemblage B isolates, ASH is typically greater than in assemblage A isolates. ASH per bp is estimated at 0.01–0.04% for assemblage A [5,6] and is nearly tenfold higher (at 0.5%) for assemblage B [5,7,8]. Therefore, different strategies for molecular typing of assemblage A and

B are required to achieve the same power of resolution [3,9]. Moreover, when typing DNA of an isolate as defined above, distinguishing monotypic infections with a strain expressing high ASH and infections with multiple strains exhibiting low ASH becomes a formidable task. Single cyst analysis to circumvent the pitfalls due to multiple infection is possible [10] but not practically feasible, in particular not for multi-locus sequence typing (MLST) approaches. Thus, a reliable tool to characterize isolates for outbreak and source attribution will rely on DNA-based procedures on cysts directly derived from stool samples [11]. The aim of this study was to evaluate and qualify genotyping protocols that could be used for epidemiological applications such as molecular tracing of transmission chains, source attribution and outbreak investigations.

## Methods

### Ethics statement

This study was approved by the ethical review committee of the Charité-University Medicine Berlin (EA4/171/19). Under the Protection against Infection Act, Giardiasis is a notifiable disease in Germany and the study conducted in accordance to the Protection against Infection Act's §13. In this context, informed consent was not required.

### Sample collection

The first set of samples was collected 2012–2019 at one outpatient tropical medicine clinic predominantly treating returning travellers. The collection comprised of 112 samples from 64 patients, including 24 chronically infected patients who provided more than one stool sample (longitudinal cases). Longitudinal cases served as a test-set to evaluate the power of typing procedures to identify epidemiologically linked cases.

The second set of samples ( $n = 171$  from 138 patients, including 23 longitudinal cases) was collected 2017 through 2019 at non-hospital associated primary care laboratories. It represented a broader patient cohort, including persons without travel history. This collection was used to further evaluate the typing procedure and to investigate possible molecular differences between autochthonous and travel-associated cases.

Both sets of samples derived from patients who had been tested for *Giardia* because of gastrointestinal symptoms.

### Molecular and sequence analysis

The workflow of the molecular analysis is depicted in Supplementary S1 Fig. All procedures were performed according to published protocols and information of primer sequences and accession numbers of reference sequences are provided in supplementary tables (S1, S2 and S3 Tables).

**DNA extraction.** *Giardia* positive stool samples were sent from primary diagnostic laboratories to the Robert Koch-Institute, were stored at 4°C and processed within two weeks. *Giardia* cysts were enriched from stool samples by sucrose gradient flotation as previously described [12] and DNA extraction was performed using the Maxwell 16 FFPE Plus LEV DNA Purification Kit (Promega Corporation, Wisconsin, USA) following the manufacturer's protocol. DNA from samples with lower cyst counts ( $\leq 10^5$  cysts/g faeces) was directly extracted from stool using the QIAamp Fast DNA Stool Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. DNA of axenically cultured trophozoites was derived from an in-house biobank (isolate 347-02, 350-01 and GS (ATCC 50581)). The principle workflow to establish axenic cultures is described elsewhere [12].



**Real-time PCR assays to assess assemblage type.** All samples were screened for the presence of assemblage A and B infections by real-time PCR to assess assemblage type. Therefore, a 168-bp amplicon at the *TIF* locus for detection assemblage A and a 99-bp fragment of the *CATH* gene for identification of assemblage B was detected according to a previously described protocol [13]. Reactions were performed in duplicates using Maxima SYBR Green/ROX qPCR Master Mix (Thermo Scientific) and 1–2  $\mu$ l target DNA. The PCR with *TIF* and *CATH* specific primer were done separately, respectively. Detection and data analysis were done using CFX96 Touch Real-Time PCR Detection System and CFX Maestro 1.1 software (Bio-Rad). Samples were considered positive when melting peaks were at 80.5–82.0°C for assemblage A and 75.5–78.0°C for assemblage B, respectively [13].

**MLST at *TPI*, *BG* and *GDH* loci (common typing scheme)** were performed by nested-PCR approaches and Sanger sequencing according to published protocols [14–16]. All reactions were carried out using DreamTaq DNA polymerase system (Thermo Scientific) and 1–2  $\mu$ l of sample DNA for the primary PCR and 1–2.5  $\mu$ l of the first PCR product for the nested PCR. *G. duodenalis* positive DNA and nuclease-free water were used as positive and negative controls in each run. The primer sequences were omitted from the sequences for the final sequence analyses and resulted in a *TPI* fragment of 490 bp, a *BG* fragment of 475 bp and *GDH* fragment of 393 bp. Sequences were concatenated (*TPI-BG-GDH*) for analysis resulting in a 1358 bp fragment (see also sequence analysis below).

**MLST of assemblage A isolates** were performed by nested-PCR and Sanger sequencing according a previously published protocol at the gene loci *HCMP22547*, *CID1*, *RHP26*, *HCMP6372*, *DIS3* and *NEK15411* [3]. Final fragment size for analysis was 555 bp for *HCMP22547*, 534 bp for *CID1*, 513 bp for *RHP26*, 564 bp for *HCMP6372*, 615 bp for *DIS3* and 633 bp for *NEK15411*. PCR was performed using the DreamTaq DNA polymerase system (Thermo Scientific) and 1–2  $\mu$ l of sample DNA for the primary PCR and 1–2.5  $\mu$ l of the first PCR product for the nested PCR. *G. duodenalis* positive DNA and nuclease-free water were used as positive and negative controls in each run. Sequences were concatenated (order: *HCMP22547*, *CID1*, *RHP26*, *HCMP6372*, *DIS3* and *NEK15411*) for analysis resulting in a 3414 bp fragment (see also sequence analysis below).

**DNA sequencing and sequence analysis.** Amplified DNA was purified with Exo Sap-IT PCR Product Cleanup (Thermo Scientific) according to the manufacture's protocol. All PCR amplicons were sequenced by Sanger sequencing in both directions in an in-house facility of the RKI. Sequences were analysed using the implemented analysis tools in the software package Geneious 11.1.5 (Biomatters). Briefly, chromatograms of raw reads were all inspected visually and annotated manually if necessary to adequately annotate heterozygous positions. Ambiguous positions were evaluated using the "heterozygotes plugin" of the software tool and based on the height of the double peaks. At least 25% peak high of the lower peak was defined as cut-off for a heterozygote position and was only counted if present in both sequencing directions. Multiple alignments with reference strains were created to determine assemblages and sub-assemblages of *G. duodenalis* isolates. Depending on MLST scheme, respective sequences of one sample were concatenated and aligned. Distance matrix was retrieved from the alignment and implemented in GraphPad Prim (GraphPad software, LLC) for calculation and visualization. Additional information about reference sequences is provided in S2 and S3 Tables.

**Definition of related isolates.** To identify potentially linked isolates we determined the mean number and standard deviation of SNP differences of each isolate to all other isolates of the respective dataset. Two isolates were determined as potentially related if they reciprocally reached the cut-off of mean minus two standard deviations, respectively (see example in S2 Fig).



**Phylogenetic analysis.** Phylogenetic analyses were performed using MEGA 7.0.26 software [12]. Sequence data of *TPI*, *BG* and *GDH* or *HCMP22547*, *CID1*, *RHP26*, *HCMP6372*, *DIS3* and *NEK15411* genes were concatenated and Neighbor-Joining method based on the maximum composite likelihood model with pairwise deletions was applied. Bootstrap analysis using 1000 replicates was performed to review the trees reliability and values above 50% were reported. The trees were depicted unrooted.

**Statistical analysis.** Results of molecular typing were linked to the respective entry among laboratory-confirmed giardiasis cases notified to the RKI according to the infection protection act. This was done to obtain the likely country of infection according to the travel history of the patient and to identify/verify whether the developed method is sensitive and specific enough to detect consecutive samples of one patient. Statistical analysis was performed using Graph Pad Prism 8.4.0 software. Fisher's exact test was applied and P values  $\leq 0.05$  were considered significant. Maps were created using the Free and Open Source Software QGIS and free vector data available at [naturalearthdata.com](http://naturalearthdata.com). Estimating the likelihood of infection with single, double, and multiple strains was done using geometric progression model. In case of detection of an assemblage A or B genotype only, this can be modelled on a geometric progression of the term  $\sum_{k=0}^{\infty} (P(\text{monoAss}))^{k+1}$  where P(monoAss) refers to the probability of single strain infection and for which the observed frequencies with which assemblage B (B plus mixed B/A) or A (A plus B/A) were detected provided estimates of the resulting sum.

**Nucleotide sequence accession numbers.** Nucleotide sequences generated in this study have been deposited into the GenBank database (accession numbers MT878631–MT879098).

## Results

### Assignment of assemblage A versus B infections as a way to estimate likelihood of mono- and multi-strain infections

In 175 of 202 (86.6%) patients, the assemblage type could be determined (S1 Fig). Patients were predominantly infected with assemblage B ( $n = 115/175$ , 65.7%, Table 1). Co-infections with A and B were seen in 35/175 patients (20.0%). There was no statistically significant difference between patients infected with assemblages A, B or mixed infection regarding age, sex or importation status (autochthonous vs. travel associated) (chi-square:  $p > 0.3$ ). Among patients

**Table 1.** Number (%) of patients with successful assemblage typing results by detected assemblage (A, B and A+B mixed infection), age group, sex and likely place of infection ( $n = 175$ ).

	Assemblage A <i>n</i> (%)	Assemblage B <i>n</i> (%)	Ass. A/B mixed infection <i>n</i> (%)	Total <i>n</i> (%)
<b>Age group (years)</b>				
0–19	5 (20%)	18 (16%)	4 (11%)	27 (15%)
20–39	8 (32%)	44 (38%)	13 (37%)	65 (37%)
40–59	5 (20%)	38 (33%)	15 (43%)	58 (33%)
60–99	7 (28%)	14 (12%)	3 (9%)	24 (14%)
Unknown	0 (0%)	1 (1%)	0 (0%)	1 (1%)
<b>Sex</b>				
male	13 (52%)	69 (60%)	15 (43%)	97 (55%)
<b>Likely place of infection</b>				
Germany	12 (48%)	39 (34%)	11 (31%)	62 (35%)
Other country	10 (40%)	70 (61%)	22 (63%)	102 (58%)
Unknown	3 (12%)	6 (5%)	2 (6%)	11 (6%)
<b>Total</b>	25 (100%)	115 (100%)	35 (100%)	175 (100%)

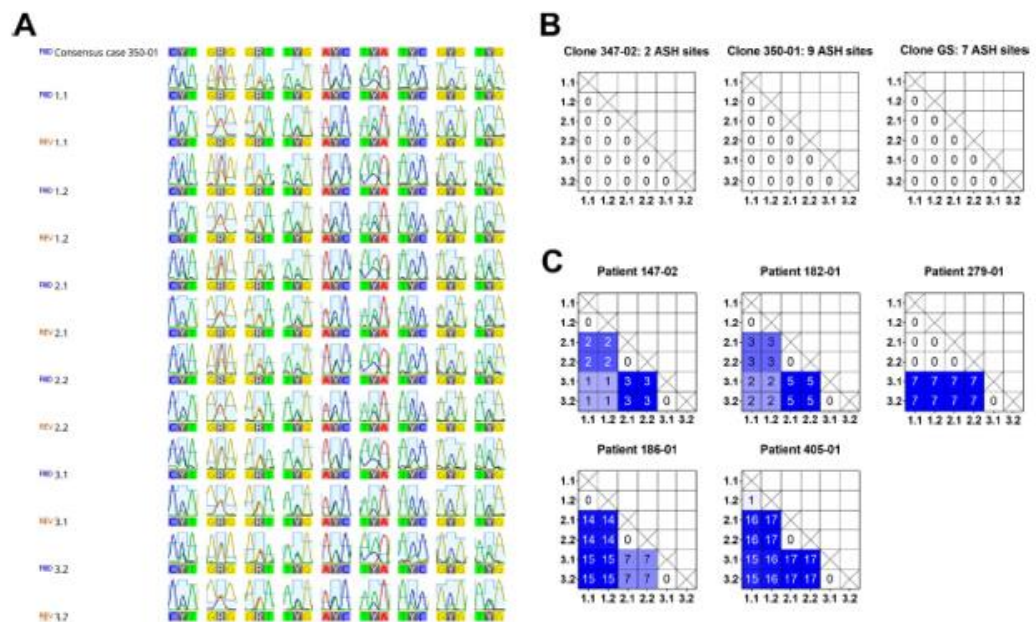
<https://doi.org/10.1371/journal.pntd.0009277.t001>

who likely acquired the infection abroad (n = 102, mean age: 37.8 years; 54.9% women), the specific country of infection was known for n = 93 (S3 Fig).

The data allowed estimating the likelihood of infection with single, double, and multiple strains. In case of detection of an assemblage A or B genotype only, this was modelled using a geometric progression term. Accordingly, approximately 75% assemblage A type infections and 50% B type infections, respectively, were likely to be single strain infections.

### Evaluation of current MLST genotyping schemes

To adapt a common protocol detecting *TPI*, *BG*, *GDH* [4], we typed cloned assemblage B strains (i.e., trophozoites from *in vitro* cultures) and selected recent assemblage B isolates (i.e., cysts enriched from stool samples). ASH in the 1358 bp long sequence covered by the three typing gene-specific PCRs was required for the validation set of templates (Fig 1). Evaluation consisted of comparison of three independent fragments generated from the same template DNA, in which every fragment was sequenced twice (Fig 1). Repeated typing of cloned parasite strains reproduced to 100% the MLST type, which included ASH residues indicating that our workflow was robust (Fig 1A and 1B). Repeated typing of assemblage B isolates surprisingly produced different and distinct degrees of sequence variation between independently



**Fig 1. Reproducibility of ASH calling in assemblage B parasites using a common three locus typing scheme at *TPI*, *BG*, *GDH*.** Each sample was typed three times in independent PCR reactions and the sequence of each PCR product was determined twice in both directions by Sanger sequencing. Results of concatenated sequences (*TPI-BG-GDH*, 1358 bp) are presented as distance matrix showing the number of pairwise differences in single nucleotide positions. Labeling depicts PCR repeats (first digit) and sequencing repeat (second digit). (A) Example depicting the chromatogram of all nine ASH sites of isolate 350-01 in multiple PCRs and after bidirectional sequencing showing the reliable identification of ASH sites in assemblage B. (B) Distance matrices of multiple PCRs showing pairwise differences in single nucleotide positions within three clonal assemblage B isolates from *in vitro* culture. Note, all ASH sites were correctly identified. (C) Distance matrices of multiple PCRs showing pairwise differences in single nucleotide positions within five assemblage B isolates derived from stool samples. Note, repeated PCRs resulted in varying apparent differences in single nucleotide positions within the same isolate.

<https://doi.org/10.1371/journal.pntd.0009277.g001>



generated typing fragments while sequencing/resequencing pairs of the respective fragments was except for one occasion always identical (Fig 1C).

Typing of assemblage A isolates, for which ASH is low and MLST using *TPI*, *BG*, *GDH* has very low discriminatory power, was performed to further assess the principle robustness of the protocol (Fig 2A). Very few SNPs were observed in the assemblage A isolates and differences were mostly shared between strains and not reflecting ASH sites. MLST typing of assemblage A isolates was then extended to include the typing genes recently described by Ankarklev et al [3] (Fig 2B). These genes were highly polymorphic in the population while exhibiting low number of ASH sites [3].

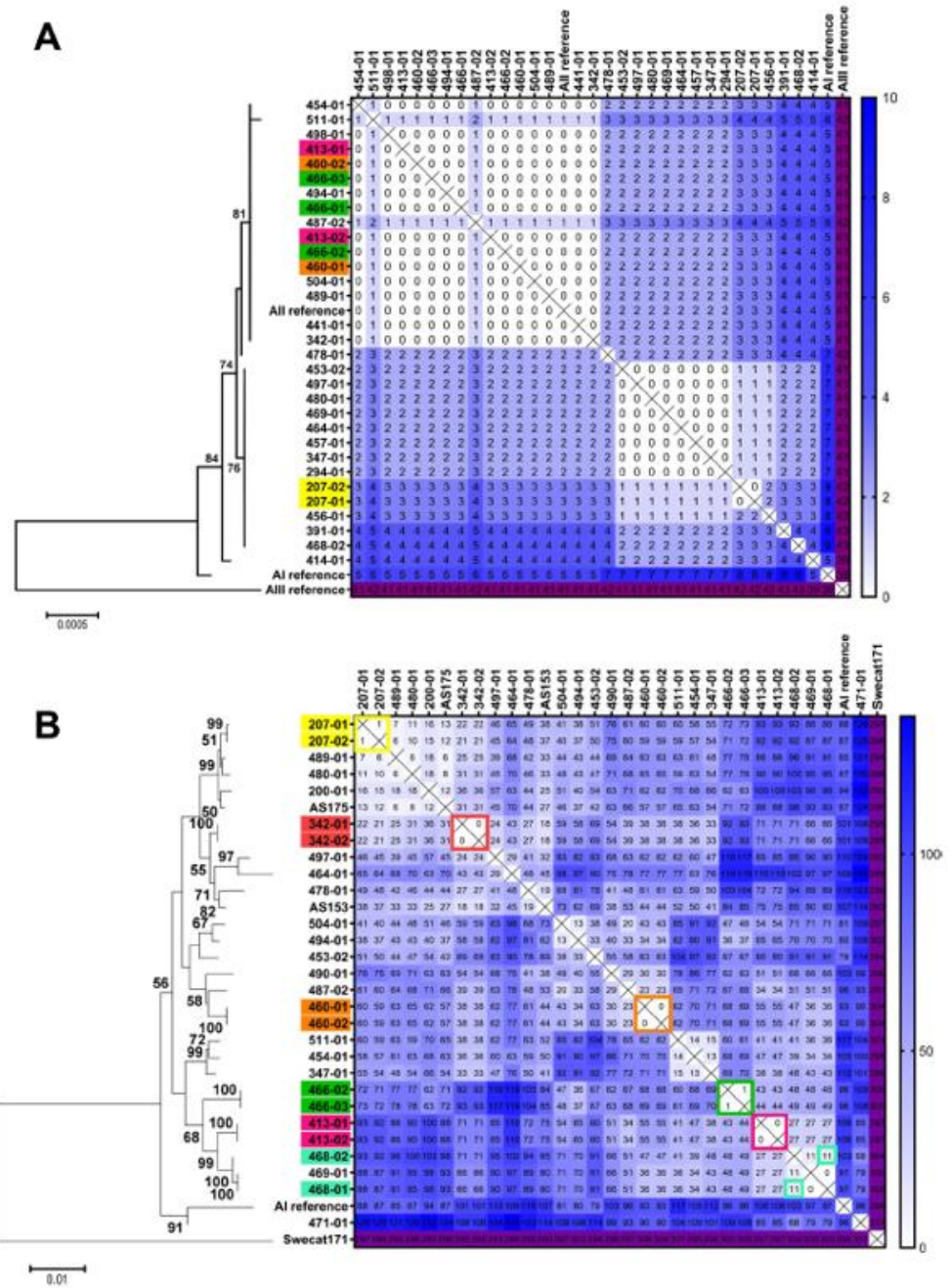
The MLST typing data of the latter scheme (Fig 2B and S1 Data) revealed 15 novel MLST types not reported before while three were identical to the previously described MLST type 18 (2 isolates), 30 (2 isolates) and 42 (1 isolate), respectively [3]. Moreover, in five of the 22 typed isolates, ASH was observed (S1 Data). Thus, five of 22 samples could be due to mixed infections, a number close to the estimated theoretical value of 25% for assemblage A infections (see above). Of note, isolates were clearly distinct from each other with the notable exception of isolates sampled longitudinally from patients with chronic disease (Fig 2B), confirming the data of the initial study by Ankarklev et al. showing the suitability of this typing scheme to identify epidemiologically linked cases of assemblage A infections [3].

### Genotyping of assemblage B by a common MLST scheme to identify epidemiological links between infections

We evaluated the potential of the common three loci MLST typing scheme available for assemblage B isolates [1,2,10,18] for outbreak or transmission chain analysis. This has been hampered by the high degree of ASH in assemblage B, a real feature that for sequence interpretation has usually not been included [2,4,18] and results in a bias in data repositories.

We first used a set of isolates from the first sample collection, which comprised of 49 completely typed assemblage B isolates sampled from 34 patients, including longitudinal samples of nine chronically infected cases. The latter were used as a test set for epidemiologically linked samples. The distance matrix of SNPs that were different in pairwise comparisons is shown in Fig 3. Mostly, these values were normally distributed around an average of 19.7 base sites depending on the individual isolate (see S2 Fig). Notably, pairwise testing for outliers that had fewer than the average minus 2SD differences correctly identified six of the nine longitudinal cases (Fig 3). For the three other isolate pairs typing produced a pattern consistent with multiple strain infections which as indicated before were estimated to be present in every other case of assemblage B infections. Thus, the typing algorithm with outlier detection did positively identify underlying epidemiological links with a sensitivity of 6 of 9 (67%) cases. In contrast, only three pairs of non-linked cases were falsely identified in the outlier testing, estimating a specificity of about 80% of this analysis. However, assessment of specificity of the analysis is more complex since population structure will impact which subset of genotypes has to be considered to determine relevant average distances for outlier analysis. For example, a number of isolates clustered with BIII or BIV reference sequence types and showed lower SNP to strains in these clusters (Fig 3). Hence, outlier detection has to be performed taking into account these related genotypes and, as a consequence, significantly lower SNPs have to be observed to infer epidemiological links.

Next, we applied the approach to samples of consecutive giardiasis cases of assemblage B mainly from Berlin, Germany, in order to test, whether such an MLST approach could identify repeated sampling of chronic cases as in the first sample set and occurrence of transmission chains. Similar to the results of the first sample set, we were able to identify a high proportion

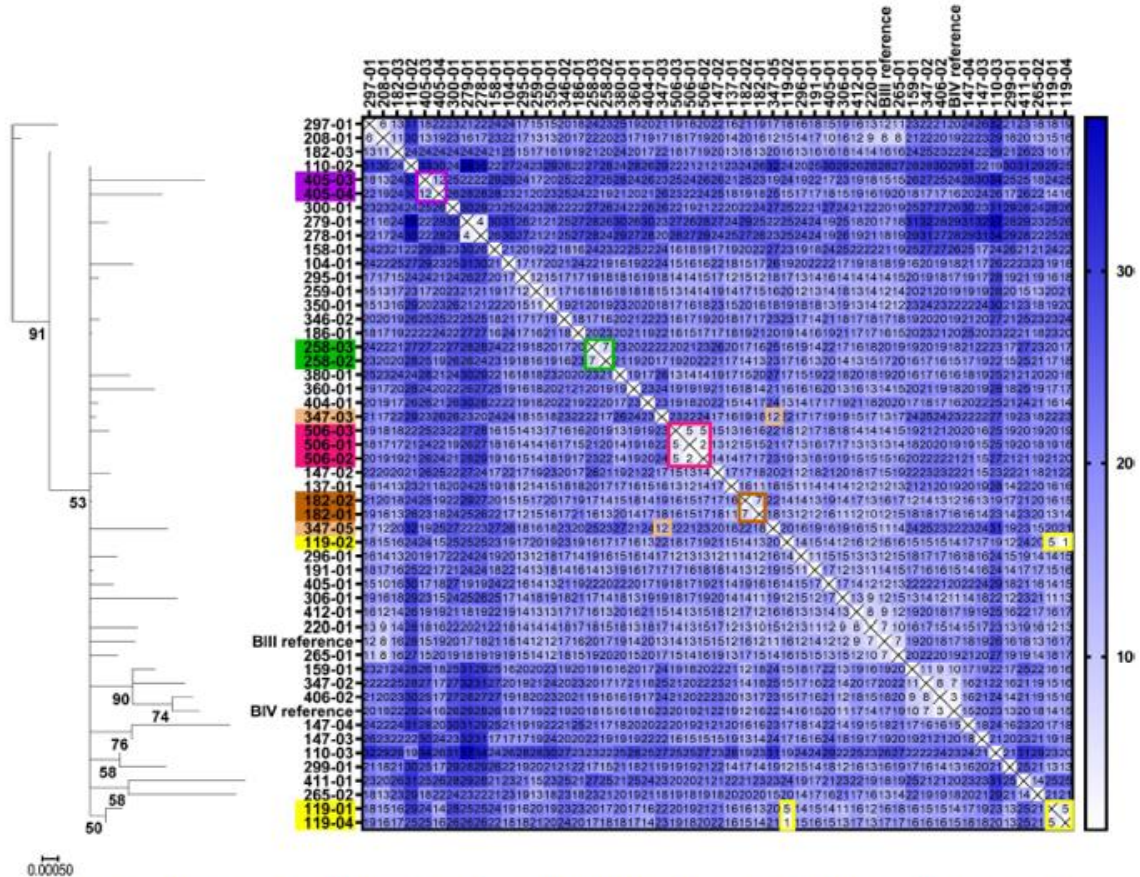




**Fig 2. Distance matrices of pairwise comparison of *G. duodenalis* assemblage A isolates.** (A) Results of concatenated sequences of common typing scheme (*TPI-BG-GDH*, 1358 bp) are presented as distance matrix showing the number of pairwise differences in single nucleotide positions. Phylogram of neighbour-joining analysis (only bootstrap values above 50 are shown) is included for illustration. (B) Sequence comparison of concatenated sequences of assemblage A specific typing scheme (*HCMP22547-CID1-RHP26-HCMP6372-DJS3-NEK15411*, 3414 bp) is presented as distance matrix showing the number of pairwise differences in single nucleotide positions. Phylogram of neighbour-joining analysis (only bootstrap values above 50 are shown) is included for illustration. Note, isolates are not identical with (A) due to varying typing efficiency of isolates. Colored case labels highlight longitudinal samples from chronically infected patients. Scale bars denote substitutions per site.

<https://doi.org/10.1371/journal.pntd.0009277.g002>

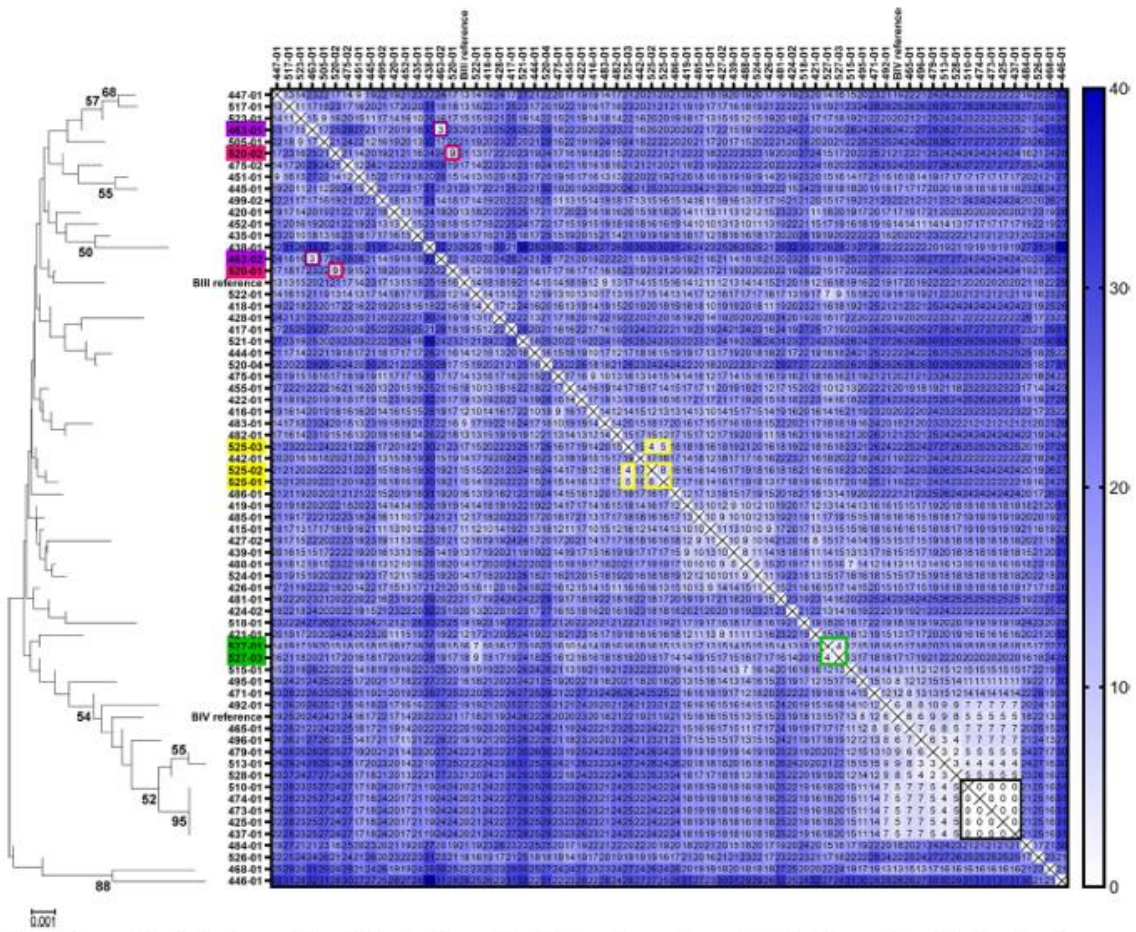
(four out of five; 80%) of longitudinal—i.e., epidemiologically linked—cases (Fig 4). In addition, a cluster of five cases was exceptional because MLST sequences were all identical and did not show any sites of ASH within the typed fragments. The potential index case of this cluster



**Fig 3. Distance matrix of pairwise comparison of *G. duodenalis* assemblage B isolates from a sample collection including known longitudinal samples from chronically infected patients.** Comparison of concatenated sequences of a common typing scheme (*TPI-BG-GDH*, 1358 bp) for assemblage B isolates of collection 1 are presented as distance matrix showing the number of pairwise differences in single nucleotide positions. Phylogram of neighbour-joining analysis (only bootstrap values above 50 are shown) is included for illustration. Longitudinal samples from chronically infected patients were highlighted in the same colour if they met the outlier detection cut-off (mean minus 2SD). Note, six of the nine longitudinal cases were correctly identified. Three pairs of unlinked isolates also met the cut-off (cases 208–01 and 297–01, case 278–01 and 279–01, case 347–02 and 406–02). Scale bar denotes substitutions per site.

<https://doi.org/10.1371/journal.pntd.0009277.g003>





**Fig 4. Distance matrix of pairwise comparison of *G. duodenalis* assemblage B isolates of consecutive samples to identify repeated sampling from chronic cases and possible occurrence of transmission chains.** Comparison of concatenated sequences of a common typing scheme (*TPI-BG-GDH*, 1358 bp) for assemblage B isolates of collection 2 are presented as distance matrix showing the number of pairwise differences in single nucleotide positions. Phylogram of neighbour-joining analysis (only bootstrap values above 50 are shown) is included for illustration. Samples from chronically infected patients were highlighted in the same colour if they met the outlier detection cut-off (mean minus 2SD). Note, five samples revealed identical nucleotide composition suggesting a before unnoticed epidemiological link (black frame). Note furthermore the subsets of isolates around BIII- and BIV-references possessing lower number of variations in single nucleotide positions. Scale bar denotes substitutions per site.

<https://doi.org/10.1371/journal.pntd.0009277.g004>

sampled first was a returning traveller who likely contracted infection in East Africa while the four consecutive cases sampled over the next 10 months were notified as autochthonous cases without a recent travel history. Notably, the isolates with this unique MLST type were very similar to seven other MLST types and highly related to reference BIV sequence type (Fig 4). This suggests that within BIV-like *G. duodenalis* there exists a highly related subpopulation. The latter “BIV-cluster” contained cases that acquired infection in various countries and may indicate the existence of a globally distributed common or dominant parasite genotype.

## Discussion

Using two instructive collections of giardiasis isolates from patients in Germany with and without relevant travel history, we show that MLST analysis can reveal epidemiological links with reasonable sensitivity and specificity, as illustrated by the identification of longitudinal samples from single patients. Our study confirms the usefulness of an extended MLST scheme to differentiate assemblage A strains and reports additional genotypes not described before. Importantly and based on assemblage typing, estimates of concurrent infections could be derived and their impact on the reproducibility of the MLST results demonstrated. Overall, samples were characterized by a dominance of assemblage B infections while no differences in age, sex or importation status (travel vs. no travel history) of infection were detectable between assemblages A, B or mixed infection case groups.

Overrepresentation of assemblage B in human infections has been reported by most studies, worldwide and also in Germany [1,2,9,19]. Other proportions of assemblage distribution may occur in specific settings such as during outbreaks or zoonotic transmission [9]. For example, a study performed in children in the Netherlands suggested that assemblage B infections are associated with anthroponotic infections and assemblage A with zoonotic transmission [20]. This is in line with a previous observation in Germany showing association of autochthonous infections with owning pets [21]. Although not statistically significant we also detected a slightly higher proportion of assemblage A in autochthonous vs. travel associated infections. Both assemblages are pathogenic in humans but there is overall inconsistent data whether or not they mediate different symptoms and whether infection sources differ [1,9,20].

Assemblage typing in the present study was successful in 86.6% of the cases. Failure to type all *Giardia* positive samples has been commonly described [22] and can likely be explained by the very high sensitivity required to amplify single copy gene targets that are mostly used for typing and that may not be reached in samples with low parasite load. Genotyping is applied more and more frequently to characterize *Giardia* populations in different hosts and from different geographic regions [1–3,9,10]. As shown here and by others [23,24] mixed infections, reduce typing reproducibility, which needs to be considered when interpreting typing results. Notably, standard Sanger sequencing, which is mostly used for *Giardia* typing procedures, is only able to reliably identify minor sequence variants with frequencies of 20% or higher [25]. Moreover, there is generally little appreciation of the pitfalls of interpreting the validity of the typing results. This is in particular true with respect to the issue of ASH in the tetraploid *Giardia* organisms. Typing entries in sequence databases are heavily biased towards sequences that do not contain polymorphic sites although the clinically predominant assemblage B parasites show extensive ASH. For example, analysis of 738 TPI sequences of assemblage B in GenBank revealed a proportion of 20.9% sequences with ambiguous nucleotide sites. In contrast, the respective proportion in our dataset was 59.1%. However, a useful typing methodology has to reflect ASH and we here show that workflows for reliable identification of ASH are feasible.

It has been highlighted that transmission chain identification needs different genotyping strategies depending on the underlying assemblage type [2,3,9]. For human assemblage A infections, we confirm that the newly developed MLST scheme based on six genome markers is reliable and useful to identify epidemiologically linked isolates [3] which supports its application in future studies.

We report MLST clusters related to the reference sub-assemblage types BIII and BIV. Our analysis illustrates that awareness of this population structure is critical for correctly interpreting typing data if the goal is to identify transmission events. Thresholds that establish epidemiological links based on number of SNPs, as proposed here, will critically depend on the relation of an isolate to relevant sub-populations, i.e. its kin.



Improvements of current typing approaches as shown here are useful and required for epidemiological investigations and studies that aim at correlating *Giardia* genotypes with infection outcome, such as clinical manifestations, or relevance for co-infections with other pathogens. These aspects cannot be adequately addressed currently. Of note, we focused on human samples only but future studies are needed to test whether the presented typing schemes are also suitable to identify possible zoonotic transmission from animal to humans.

The public health impact of typing, e.g., on outbreak strain definition, is critically dependent on the ability to communicate results unambiguously which requires a nomenclature. Such nomenclatures are established for bacteria (<https://pubmlst.org/>). For other protozoa like *Cryptosporidium* spp. typing nomenclature is already advanced, partly because these organisms are haploid [26]. An appropriate nomenclature for the tetraploid *G. duodenalis* may only be possible through defining allele types for sets of highly polymorphic marker genes. Full genotypes, i.e., calling respective sets of alleles, may then be possible via novel sequencing independent, CRISPR/Cas-mediated typing approaches [27]. Harnessing the power of such new approaches for genotyping *Giardia* again will require in-depth insight into the parasites' population structure.

Whole genome sequencing at low cost may allow this insight as it is revolutionizing molecular surveillance of microbial pathogens, outbreak detection and source identification (see reviews for listeria [28] and salmonella [29,30]). The lack of a useful *G. duodenalis* reference of its pan- and accessory genome representing the population is delaying progress in this field [6]. The recent proof-of-concept for successful whole genome sequencing of *Giardia* cysts derived from faecal samples promises to change this [31,32].

In conclusion, the present study demonstrates the power and limitations of current genotyping approaches for *G. duodenalis*: Potential for identification of epidemiological links between *G. duodenalis* infections and limitations that are inherent attributes of the tetraploid nature of the parasite and of the frequency of concurrent infections with distinct *G. duodenalis* strains.

## Supporting information

**S1 Fig. Workflow of the molecular analysis and summary of typing results.**

(PDF)

**S2 Fig. Graphical illustration of cut-off value for exemplary samples.**

(PDF)

**S3 Fig. *G. duodenalis* assemblage type by country of infection.**

(PDF)

**S1 Table. Primer sequences used in the study.**

(PDF)

**S2 Table. Accession number of references used for analysis of the common MLST scheme.**

(PDF)

**S3 Table. Accession numbers of references used for analysis of MLST results for the assemblage A typing scheme.**

(PDF)

**S1 Data. Data table depicting overall typing results.**

(XLSX)



### Author Contributions

**Conceptualization:** Mirko Faber, Klaus Stark, Ralf Ignatius, Toni Aebischer, Christian Klotz.

**Data curation:** Andreas Woschke, Mirko Faber, Christian Klotz.

**Formal analysis:** Andreas Woschke, Mirko Faber, Christian Klotz.

**Investigation:** Andreas Woschke, Mirko Faber, Martha Holtfreter, Frank Mockenhaupt, Joachim Richter, Thomas Regnath, Ingo Sobottka, Ingrid Reiter-Owona, Andreas Diefenbach, Petra Gosten-Heinrich, Johannes Friesen, Christian Klotz.

**Methodology:** Andreas Woschke, Petra Gosten-Heinrich, Christian Klotz.

**Resources:** Martha Holtfreter, Frank Mockenhaupt, Joachim Richter, Thomas Regnath, Ingo Sobottka, Ingrid Reiter-Owona, Andreas Diefenbach, Johannes Friesen, Ralf Ignatius.

**Supervision:** Ralf Ignatius, Toni Aebischer, Christian Klotz.

**Validation:** Christian Klotz.

**Visualization:** Andreas Woschke, Mirko Faber, Christian Klotz.

**Writing – original draft:** Andreas Woschke, Toni Aebischer, Christian Klotz.

**Writing – review & editing:** Andreas Woschke, Mirko Faber, Klaus Stark, Martha Holtfreter, Frank Mockenhaupt, Joachim Richter, Thomas Regnath, Ingo Sobottka, Ingrid Reiter-Owona, Johannes Friesen, Ralf Ignatius, Toni Aebischer, Christian Klotz.

### References

1. Caccio SM, Lalle M, Svard SG. Host specificity in the *Giardia duodenalis* species complex. *Infect Genet Evol.* 2018; 66:335–45. <https://doi.org/10.1016/j.meegid.2017.12.001> PMID: 29225147.
2. Feng Y, Xiao L. Zoonotic potential and molecular epidemiology of *Giardia* species and giardiasis. *Clin Microbiol Rev.* 2011; 24(1):110–40. Epub 2011/01/15. <https://doi.org/10.1128/CMR.00033-10> PMID: 21233509; PubMed Central PMCID: PMC3021202.
3. Ankarklev J, Lebbad M, Einarsson E, Franzen O, Ahola H, Troell K, et al. A novel high-resolution multilocus sequence typing of *Giardia intestinalis* Assemblage A isolates reveals zoonotic transmission, clonal outbreaks and recombination. *Infect Genet Evol.* 2018; 60:7–16. <https://doi.org/10.1016/j.meegid.2018.02.012> PMID: 29438742.
4. Sprong H, Caccio SM, van der Giessen JW, network Z, partners. Identification of zoonotic genotypes of *Giardia duodenalis*. *PLoS Negl Trop Dis.* 2009; 3(12):e558. Epub 2009/12/04. <https://doi.org/10.1371/journal.pntd.0000558> PMID: 19956662; PubMed Central PMCID: PMC2777335.
5. Adam RD, Dahlstrom EW, Martens CA, Bruno DP, Barbian KD, Ricklefs SM, et al. Genome sequencing of *Giardia lamblia* Genotypes A2 and B isolates (DH and GS) and comparative analysis with the genomes of Genotypes A1 and E (WB and Pig). *Genome Biol Evol.* 2013; 5(12):2498–511. Epub 2013/12/07. <https://doi.org/10.1093/gbe/ev1197> PMID: 24307482.
6. Xu F, Jex A, Svard SG. A chromosome-scale reference genome for *Giardia intestinalis* WB. *Sci Data.* 2020; 7(1):38. <https://doi.org/10.1038/s41597-020-0377-y> PMID: 32019935; PubMed Central PMCID: PMC7000408.
7. Franzen O, Jerlstrom-Hultqvist J, Castro E, Sherwood E, Ankarklev J, Reiner DS, et al. Draft genome sequencing of *giardia intestinalis* assemblage B isolate GS: is human giardiasis caused by two different species? *PLoS Pathog.* 2009; 5(8):e1000560. Epub 2009/08/22. <https://doi.org/10.1371/journal.ppat.1000560> PMID: 19696920; PubMed Central PMCID: PMC2723961.
8. Jerlstrom-Hultqvist J, Franzen O, Ankarklev J, Xu F, Nohynkova E, Andersson JO, et al. Genome analysis and comparative genomics of a *Giardia intestinalis* assemblage E isolate. *BMC Genomics.* 2010; 11:543. Epub 2010/10/12. <https://doi.org/10.1186/1471-2164-11-543> PMID: 20929575; PubMed Central PMCID: PMC3091692.
9. Ryan U, Caccio SM. Zoonotic potential of *Giardia*. *Int J Parasitol.* 2013; 43:943–56. Epub 2013/07/17. <https://doi.org/10.1016/j.ijpara.2013.06.001> PMID: 23856595.

10. Anarklev J, Svard SG, Lebbad M. Allelic sequence heterozygosity in single *Giardia* parasites. *BMC Microbiol.* 2012; 12(1):65. Epub 2012/05/05. <https://doi.org/10.1186/1471-2180-12-65> PMID: 22554281.
11. Chalmers RM, Robertson LJ, Domy P, Jordan S, Karssin A, Katzer F, et al. Parasite detection in food: Current status and future needs for validation. *Trends in Food Science & Technology.* 2020; 99:337–50. <https://doi.org/10.1016/j.tifs.2020.03.011> WOS:000526719400027.
12. Hahn J, Seeber F, Kolodziej H, Ignatius R, Laue M, Aebischer T, et al. High Sensitivity of *Giardia* duodenalis to Tetrahydrolipstatin (Orlistat) In Vitro. *PLoS One.* 2013; 8(8):e71597. Epub 2013/08/27. <https://doi.org/10.1371/journal.pone.0071597> PMID: 23977083; PubMed Central PMCID: PMC3747212.
13. Van Lith L, Soba B, Vizcaino VV, Svard S, Sprong H, Tosini F, et al. A real-time assemblage-specific PCR assay for the detection of *Giardia duodenalis* assemblages A, B and E in fecal samples. *Vet Parasitol.* 2015; 211(1–2):28–34. <https://doi.org/10.1016/j.vetpar.2015.04.017> PMID: 25935292.
14. Lalle M, Pozio E, Capelli G, Bruschi F, Crotti D, Caccio SM. Genetic heterogeneity at the beta-giardin locus among human and animal isolates of *Giardia duodenalis* and identification of potentially zoonotic subgenotypes. *Int J Parasitol.* 2005; 35(2):207–13. Epub 2005/02/16. <https://doi.org/10.1016/j.ijpara.2004.10.022> PMID: 15710441.
15. Read CM, Monis PT, Thompson RC. Discrimination of all genotypes of *Giardia duodenalis* at the glutamate dehydrogenase locus using PCR-RFLP. *Infect Genet Evol.* 2004; 4(2):125–30. Epub 2004/05/26. <https://doi.org/10.1016/j.meegid.2004.02.001> PMID: 15157630.
16. Sulaiman IM, Fayer R, Bern C, Gilman RH, Trout JM, Schantz PM, et al. Triosephosphate isomerase gene characterization and potential zoonotic transmission of *Giardia duodenalis*. *Emerg Infect Dis.* 2003; 9(11):1444–52. Epub 2004/01/14. <https://doi.org/10.3201/eid0911.030084> PMID: 14718089; PubMed Central PMCID: PMC3035538.
17. Kumar S, Stecher G, Li M, Niyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol.* 2018; 35(6):1547–9. <https://doi.org/10.1093/molbev/msy096> PMID: 29722887; PubMed Central PMCID: PMC5967553.
18. Caccio SM, Beck R, Lalle M, Marinculic A, Pozio E. Multilocus genotyping of *Giardia duodenalis* reveals striking differences between assemblages A and B. *Int J Parasitol.* 2008; 38(13):1523–31. Epub 2008/06/24. <https://doi.org/10.1016/j.ijpara.2008.04.008> PMID: 18571176.
19. Broglia A, Weitzel T, Harms G, Caccio SM, Nockler K. Molecular typing of *Giardia duodenalis* isolates from German travellers. *Parasitol Res.* 2013; 112(10):3449–56. <https://doi.org/10.1007/s00436-013-3524-y> PMID: 23892479.
20. Pijnacker R, Mughini-Gras L, Heusinkveld M, Roelfsema J, van Pelt W, Kortbeek T. Different risk factors for infection with *Giardia lamblia* assemblages A and B in children attending day-care centres. *Eur J Clin Microbiol Infect Dis.* 2016; 35(12):2005–13. <https://doi.org/10.1007/s10096-016-2753-2> PMID: 27599710.
21. Espelage W, an der Heiden M, Stark K, Alpers K. Characteristics and risk factors for symptomatic *Giardia lamblia* infections in Germany. *BMC Public Health.* 2010; 10:41. Epub 2010/01/29. <https://doi.org/10.1186/1471-2458-10-41> PMID: 20105338; PubMed Central PMCID: PMC2824735.
22. Thompson RC, Ash A. Molecular epidemiology of *Giardia* and *Cryptosporidium* infections. *Infect Genet Evol.* 2016:315–23. <https://doi.org/10.1016/j.meegid.2015.09.028> PMID: 26458528.
23. Hussein AIA, Yamaguchi T, Nakamoto K, Iseki M, Tokoro M. Multiple-subgenotype infections of *Giardia intestinalis* detected in Palestinian clinical cases using a subcloning approach. *Parasitology International.* 2009; 58(3):258–62. <https://doi.org/10.1016/j.parint.2009.04.002> WOS:000268917800011. PMID: 19361570.
24. Mizuno T, Matey EJ, Bi XQ, Songok EM, Ichimura H, Tokoro M. Extremely diversified haplotypes observed among assemblage B population of *Giardia intestinalis* in Kenya. *Parasitology International.* 2020; 75. <https://doi.org/10.1016/j.parint.2019.102038> WOS:000518493100023. PMID: 31837398.
25. Mohamed S, Penaranda G, Gonzalez D, Camus C, Khiri H, Boulime R, et al. Comparison of ultra-deep versus Sanger sequencing detection of minority mutations on the HIV-1 drug resistance interpretations after virological failure. *AIDS.* 2014; 28(9):1315–24. <https://doi.org/10.1097/QAD.0000000000000267> PMID: 24698843.
26. Chalmers RM, Robinson G, Elwin K, Elson R. Analysis of the *Cryptosporidium* spp. and gp60 subtypes linked to human outbreaks of cryptosporidiosis in England and Wales, 2009 to 2017. *Parasit Vectors.* 2019; 12(1):95. <https://doi.org/10.1186/s13071-019-3354-6> PMID: 30867023; PubMed Central PMCID: PMC6417012.
27. Ackeman CM, Myhrvold C, Thakku SG, Freije CA, Metsky HC, Yang DK, et al. Massively multiplexed nucleic acid detection with Cas13. *Nature.* 2020; 582(7811):277–82. <https://doi.org/10.1038/s41586-020-2279-8> PMID: 32349121; PubMed Central PMCID: PMC7332423.

28. Datta AR, Burall LS. Serotype to genotype: The changing landscape of listeriosis outbreak investigations. *Food Microbiol.* 2018; 75:18–27. <https://doi.org/10.1016/j.fm.2017.06.013> PMID: 30056958.
29. Kruy SL, van Cuyck H, Koeck JL. Multilocus variable number tandem repeat analysis for *Salmonella enterica* subspecies. *Eur J Clin Microbiol Infect Dis.* 2011; 30(4):465–73. <https://doi.org/10.1007/s10096-010-1110-0> PMID: 21153561.
30. Mughini-Gras L, Franz E, van Pelt W. New paradigms for *Salmonella* source attribution based on microbial subtyping. *Food Microbiol.* 2018; 71:60–7. <https://doi.org/10.1016/j.fm.2017.03.002> PMID: 29366470.
31. Hanevik K, Bakken R, Brattbakk HR, Saghaug CS, Langeland N. Whole genome sequencing of clinical isolates of *Giardia lamblia*. *Clin Microbiol Infect.* 2015; 21(2):192 e1–3. <https://doi.org/10.1016/j.cmi.2014.08.014> PMID: 25596782.
32. Kooyman FNJ, Wagenaar JA, Zomer A. Whole-genome sequencing of dog-specific assemblages C and D of *Giardia duodenalis* from single and pooled cysts indicates host-associated genes. *Microb Genom.* 2019; 5(12). <https://doi.org/10.1099/mgen.0.000302> PMID: 31821130; PubMed Central PMCID: PMC6939161.

## **Lebenslauf**

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

## Komplette Publikationsliste

1. **Woschke A**, Faber M, Stark K, Holtfreter M, Mockenhaupt F, Richter J, Regnath T, Sobottka I, Reiter-Owona I, Diefenbach A, Gosten-Heinrich P, Friesen J, Ignatius R, Aebischer T, Klotz C. Suitability of current typing procedures to identify epidemiologically linked human *Giardia duodenalis* isolates. *PLoS Negl Trop Dis.* 2021;15(3):e0009277.  
*Impact factor* des Journals: 4.411

## **Danksagung**

Zuallererst möchte ich mich bei meinem Doktorvater Prof. Dr. Ralf Ignatius und meinen beiden Betreuern im RKI, Dr. Anton Aebischer und Dr. Christian Klotz, herzlichst bedanken. Zum einen für die Überlassung dieses spannendes Forschungsthemas, vor allem aber für die exzellente Betreuung. Wann immer ich eine Frage hatte oder Rat suchte, nahmen sie sich Zeit für mich und gaben mir die Hilfe zur Selbsthilfe, die ich brauchte. Das ist ganz sicher keine Selbstverständlichkeit. Besonders möchte ich mich auch für das kritische, konstruktive und motivierende Feedback bedanken. Wichtig waren mir auch die Labor-Meetings, die mich stets intellektuell herausforderten und mein wissenschaftliches Denken förderten. Die Zusammenarbeit mit den anderen Wissenschaftler\*innen des RKI war mir eine große Freude. Nicht nur, weil das soziale Miteinander sehr angenehm war, sondern auch, weil ich von jedem Mitglied der Arbeitsgruppen etwas lernen konnte und alle ihr Wissen bereitwillig teilten. Besonders erwähnenswert sind dabei Prof. Dr. Frank Seeber, Dr. Totta Ehret Kasemo, Dr. Benedikt Fabian, Dr. Martin Kraft, Estefania Delgado Betancourt, Sharaeh Salehi und David Holthaus. Mein Dank gilt ebenfalls den Medizinisch- technischen Assistentinnen, allen voran Petra Gosten-Heinrich und Elke Radam, die mir so manchen praktischen „Kniff“ gezeigt haben. Die Arbeitsatmosphäre in der gesamten Arbeitsgruppe war hervorragend, dafür danke ich allen Mitgliedern von FG16. Bedanken möchte ich mich auch bei meinen Ko-Autor\*innen für die gute Zusammenarbeit und die konstruktiven Denkanstöße.

Ich bedanke mich auch bei der Studienstiftung des Deutschen Volkes für das entgegengebrachte Vertrauen in Form eines Stipendiums, dass es mir nicht nur ermöglicht sorgenfrei zu studieren, sondern auch auf zahlreichen Veranstaltungen meinen Horizont beständig zu erweitern.

Zu guter Letzt möchte ich auch meinen Eltern danken, die mir stets den Rücken freigehalten und es mir ermöglicht haben, meinen Traum zu leben und Medizin zu studieren.