Phylogenomics and Diversification of the Schistosomatidae Based on Targeted Sequence Capture of Ultra-Conserved Elements

Erika T. Ebbs 1,*, Eric S. Loker 2, Lijing Bu 2, Sean A. Locke 3, Vasyl V. Tkach 4, Ramesh Devkota 5, Veronica R. Flores 6, Hudson A. Pinto 7 and Sara V. Brant 2

1 Department of Biology, Purchase College, The State University of New York, Purchase, NY 10577, USA
2 Center for Evolutionary and Theoretical Immunology, Department of Biology, Museum of Southwestern Biology Parasite Division, University of New Mexico, Albuquerque, NM 87131, USA; esloker@unm.edu (E.S.L.); lijing@unm.edu (L.B.); sbrant@unm.edu (S.V.B.)
3 Department of Biology, University of Puerto Rico at Mayagüez, Box 9000, Mayagüez 00681-9000, Puerto Rico; sean.locke@upr.edu
4 Grand Forks Department of Biology, University of North Dakota, Grand Forks, ND 58202, USA; vasyl.tkach@und.edu
5 Vance Granville Community College, Henderson, NC 27536, USA; devkorat@vgcc.edu
6 Laboratorio de Parasitología, INIBIOMA (CONICET-Universidad Nacional del Comahue), Quintral 1250, San Carlos de Bariloche 8400, Argentina; veronicaroxanaflores@gmail.com
7 Department of Parasitology, Institute of Biological Science, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil; hudsonalves13@icb.ufmg.br
* Correspondence: erika.ebbs@purchase.edu

Abstract: Schistosomatidae Stiles and Hassall 1898 is a medically significant family of digenetic trematodes (Trematoda: Digenea), members of which infect mammals or birds as definitive hosts and aquatic or amphibious gastropods as intermediate hosts. Currently, there are 17 named genera, for many of which evolutionary interrelationships remain unresolved. The lack of a resolved phylogeny has encumbered our understanding of schistosomatid evolution, specifically patterns of host-use and the role of host-switching in diversification. Here, we used targeted sequence capture of ultra-conserved elements (UCEs) from representatives of 13 of the 17 named genera and 11 undescribed lineages that are presumed to represent either novel genera or species to generate a phylogenomic dataset for the estimation of schistosomatid interrelationships. This study represents the largest phylogenetic effort within the Schistosomatidae in both the number of loci and breadth of taxon sampling. We present a near-comprehensive family-level phylogeny providing resolution to several clades of long-standing uncertainty within Schistosomatidae, including resolution for the placement of the North American mammalian schistosomes, implying a second separate capture of mammalian hosts. Additionally, we present evidence for the placement of Macrobilharzia at the base of the Schistosoma + Bivitellobilharzia radiation. Patterns of definitive and intermediate host use and a strong role for intermediate host-switching are discussed relative to schistosomatid diversification.

Keywords: schistosomatid; blood fluke; UCE; parasite phylogenomics; sequence capture; diversification; host-switching; museum collections; voucher

1. Introduction

An informed understanding of diversification is lacking for most multi-host helminth groups [1,2]. In fact, few such groups have been sampled sufficiently to reconstruct a reliable, well-resolved phylogeny, which is necessary to understand parasite evolution. Elucidating helminth diversification has practical implications for knowledge and the management of human and wildlife helminths, issues relating to host-competency [3], potential for host-switching [4–7], host associations [8–10], character evolution [11], and
cladogenesis [12,13]. Among digenetic trematodes, the Schistosomatidae Stiles and Hassall 1898 is a diverse family with its members infecting birds or mammals (definitive hosts) and aquatic or amphibious gastropods (intermediate hosts). The Schistosomatidae have garnered substantial interest among parasitologists, due to the medical and veterinary importance of its members. Three species within the genus *Schistosoma* Weiland, 1858 are the major etiological agents of human schistosomiasis, one of the world’s most recalcitrant neglected tropical diseases still infecting over 250 million people globally [14,15]. Moreover, non-human schistosomatids are involved in human cercarial dermatitis (swimmer’s itch), a re-emerging zoonotic disease [16,17].

Among digeneans, schistosomatids have unusual characteristics, the most notable being that they are exclusively dioecious (usually with marked sexual dimorphism) parasites of endotherms (birds and mammals), which are collective attributes not shared with any other digenean family [18,19]. Schistosomatids have evolved heterogametic (ZW) sex chromosomes [18], which function in sex determination. Schistosomatids possess other interesting reproductive strategies [20,21] and are among the few digenean lineages in which adults reside outside the alimentary canal of their vertebrate definitive hosts. They also lack a second intermediate host in their life cycles. Decades of sequencing of the DNA from schistosomatids from their definitive and intermediate hosts (e.g., see other papers in this volume) coupled with the use of well-vouchered museum specimens have shed light on the previously under-recognized diversity of schistosomatids, especially for avian-infecting species, including species responsible for human cercarial dermatitis. This effort has provided a foundation to develop a robust evolutionary framework to address schistosomatid diversity and diversification worldwide.

Despite advances, our understanding of schistosomatid interrelationships, patterns of host-use, and character evolution remains incomplete [20,22–24]. Progress has been constrained by a lack of informative morphological characteristics and adequate adult parasite material for description. The inclusion of molecular analyses [25–27] has greatly expanded efforts to explain schistosomatid diversity, particularly for avian-infecting genera, which have been proven to contain cryptic diversity [10,28]. Although these and other studies have made considerable strides, deeper nodes pertaining to the interrelationships between genera remain unresolved and often provide conflicting phylogenetic signals. This is likely because molecular phylogenies have been based on few loci, primarily markers within the ribosomal RNA operon (28S, 18S, ITS1, and 2) and a single mitochondrial gene (cox1).

Currently, within the Schistosomatidae, there are 17 named genera and over 130 named species [21]. Based on the sampling of larval schistosomatids, this is likely an underestimate [8,29–35] of both genera and species. Historically, Schistosomatidae was divided into three subfamilies [24,36], primarily based on the morphology of adult worms. Schistosomatinae Stiles and Hassall 1898 includes *Austrobilharzia*, *Bivitellobilharzia*, *Heterobilharzia*, *Macrobilharzia*, *Ornithobilharzia*, *Schistosomatium*, and *Schistosoma*. Gigantobilharzinae Mehra 1940 comprises *Gigantobilharzia* and *Dendritobilharzia*. Bilharziellinae Price 1929 includes *Bilharziella*, *Trichobilharzia*, *Jilinobilharzia*, *Allobilharzia*, and *Anserobilharzia*. A fourth subfamily, Grippibilharziinae Platt, Blair, Purdie & Melville, 1991, containing a single species *Grippibilharzia amoena*, has been recognized, but sequence-based data places this species within the Spirorchidae [36]. Molecular phylogenies have not supported these subfamily designations [21,25–27,31,34,37–39], and the most appropriate subfamily classifications remain unclear.

As illustrated by the conflicting results from prior studies (Table 1), the continued sequencing of 500–1500 bp markers seems unlikely to resolve deeper schistosomatid relationships. New methodologies are needed to resolve the following pivotal nodes and groups, and thus better characterize diversification, evolution of host-use, host switches and divergence times within Schistosomatidae:  

1. **AO Clade:** Marine avian genera, *Austrobilharzia* and *Ornithobilharzia* (AO clade), are often recovered as well-supported sister genera [36,39,40], and both are considered
globally distributed. In most of the studies over the last 10 years, the AO clade typically acts as a sister group to the remainder of the schistosomatids.

2. **SB Clade:** Species of the Afro-Eurasian mammalian clade (*Schistosoma* + *Bivitellobilharzia*, or the SB clade) are found in tropical and sub-tropical latitudes. *Schistosoma* and *Bivitellobilharzia* are considered probable sister genera, but the use of traditional loci often does not statistically support this grouping. A better understanding of relationships within the SB clade and its placement within Schistosomatidae is essential to understanding the evolution of human schistosomatids.

3. **Macrobilharzia:** *Macrobilharzia* is a monotypic genus with species that infect *Anhinga anhinga* in the Americas and has failed to group consistently with other schistosomatid lineages. Some rRNA phylogenies suggest an (unsupported) affinity with the SB clade, suggesting the possibility of the SB clade having had Afro-Eurasian avian-infecting ancestors. Cercariae (Schistosomatidae-W688) recovered from the freshwater snail *Indoplanorbis exustus* in Nepal [41] represent an otherwise undescribed, but well-supported, sister lineage to *M. macrobilharzia*.

4. **DAS Clade:** The derived avian–schistosomatid (DAS) clade includes the majority of avian-infecting genera (*Anserobilharzia, Allobilharzia, Bilharziella, Dendritobilharzia, Gigantobilharzia, Marinabilharzia, Nausibilharzia, Riverabilharzia, Trichobilharzia*) and several yet-to-be described genera. The monophyly of DAS is consistently supported (Table 1, [39,40]), though its placement as a sister to the SB clade or North American mammalian schistosomatids is unclear. The relationships within the DAS clade are largely unresolved.

5. **HS Clade:** *Heterobilharzia* and *Schistosomatium* (HS clade) are both monotypic genera whose species infect North American mammals and form a well-supported clade in most studies (Table 1). Phylogenies based on oft-used loci have provided weak support for the placement of the HS clade as a sister to the DAS clade (see Table 1). Phylogenetic placement of the HS clade has significant implications for understanding the role of host shifts and biogeography of the Schistosomatidae.

The conflict among these pivotal nodes (Table 1) might be explained by widespread incomplete lineage sorting (ILS) within the Schistosomatidae, as suggested by previous authors [21,27]. Resolving relationships in this context requires increasing the number of characters analyzed, ideally through augmenting the number of independently evolv-
ing loci, as well as more complete taxon sampling [42–44]. Recent analyses of whole mitogenomes for species of Schistosoma [45] have yielded increased resolution, however this approach is unlikely to yield similar results when applied to the Schistosomatidae, as mitochondrial loci evolve in concert. Moreover, among digeneans, whole mitochondrial genomes present evolutionary trajectories that differ strikingly at deeper levels from nuclear genomes [46,47], suggesting a need for caution. Recent developments in phylogenomic methods have proven effective in resolving clades where ILS was suspected [48,49]. One such method uses the targeted sequence capture of ultra-conserved elements (UCEs) [50] to obtain data from hundreds to thousands of independent nuclear loci for phylogenomic analysis. This method requires a panel of probes that target UCE loci and less conserved flanking regions in the group of interest, which are sequenced using next-generation methods, producing alignments of thousands to millions of bases that can resolve divergence points across different time scales [50]. The use of UCEs as phylogenomic markers has been successful at resolving rapid radiations within vertebrates [51–54] as well as relationships within an increasing number of invertebrate groups [55–58]. Within Digenea, the only application of this approach was characterized by a relatively low number of loci (517 UCEs in [47]).

The present study applied a sequence capture approach to generate UCE loci for vouchered schistosomatid specimens collected over a 20+ year period, to generate a more fully resolved phylogeny of Schistosomatidae, to understand patterns of host-switching, character evolution and more broadly, diversification. This more resolved phylogeny will enable the generation of more targeted phylogenetic hypotheses regarding schistosomatid diversification, and can be applied to address questions such as: have specific traits aided schistosomatid diversification? How common is intermediate and definitive host-switching within Schistosomatidae? Did the major schistosomatid clades radiate simultaneously? Additionally, a resolved phylogeny will provide direction for future targeted collection efforts and will help characterize the timing of schistosomatid lineage formation.

2. Results

2.1. UCE Enrichment and Sequencing

Targeted sequence capture of UCE loci was performed on 39 schistosomatid samples (Table 2). Multiplexed sequencing of enriched libraries resulted in an average 6,939,218 (23,988–24,435,956) reads per sample, with an average sequencing depth of 350 × (58–1001 ×). An average of 44,060 (12–442,043) contigs with a mean length of 325 bp was assembled. On average, we recovered ~500, ~1500, ~884, and ~850 UCE loci in R1, R2, R3, and R4, respectively (Table 3). This is a lower recovery rate relative to the published sequence capture datasets of vertebrates [51,52] and the few existing studies on invertebrates [51], which is likely related to the low amounts of starting DNA (Table 3, [59]).

2.2. Phylogenetic Reconstruction

2.2.1. Supermatrix Alignment

Thirteen unique alignments, ranging from being 82–36.4% complete, were analyzed and their likelihood scores were calculated in RAxML (Supplementary File S2, [60]), to determine an appropriate level of data completeness for subsequent analyses. The UCE supermatrix contained taxa and loci, where: (1) ≥70% of taxa were shared for a given locus, and (2) of the remaining loci (nloci = 554), the percentage of nucleotide coverage was ≥70% per locus (Supplementary File S3). The lowest nucleotide coverage per sample, averaged over 554 loci, occurred in W333 (Anserobilharzia) and C1 (Avian Schistosomatidae lineage 3). Nucleotide coverage was high in taxa of key interest such as Macrobilharzia macrobilharzia (81% of 554 loci), Schistosomatidae sp. W688 (84%), Heterobilharzia americana (93%), and Bivitellobilharzia nairi (73%) (Figure 1, Supplementary File S3).
Figure 1. RAxML analyses of concatenated UCE loci. Phylogeny was estimated using the maximum likelihood method using a GTRGAMMA + I model, based on 554 UCE loci. Node circles indicate the bootstrap support value, corresponding to the legend in the top left corner. The tree was edited in FigTree v1.4. The gray open circle denotes the major divergence of the derived schistosomatids. The two gray arrows denote the two major derived nodes leading to two simultaneous and independent radiations, one predominantly in birds the other predominantly in mammals. Spirorchiidae samples fall into two clades concordant with the results of Bullard et al. [61], with a South American species recovered from a freshwater turtle, as sister to Schistosomatidae.

2.2.2. Phylogenies

Across all phylogenies reconstructed (Figures 1 and 2) and methods of phylogenetic inference (ML and BI), the resulting topologies converged to support the monophyly of the AO, SB, M, HS, and DAS clades, and their position within the Schistosomatidae. The AO clade was grouped as a sister to the rest of the Schistosomatidae, suggesting an early diverging or ancestral position. The remaining schistosomatids will be referred to as “derived” taxa, relative to the AO clade. Among the derived taxa two primary clades formed, (1) SB + M form a clade, with moderate (Figure 1) to strong (Figure 2) support values. The M clade contains strong support for Macrobilharzia as a sister genus to Schistosomatidae sp. W688. Within the SB clade, there is strong support for Birbilharzia as a sister to Schistosoma, and within Schistosoma there is strong support for the main species groups [62–64] that was recovered. (2) HS + DAS form a clade with strong support. Bilharziella polonica was found to be basal relative to the DAS clade.
Table 2. The included taxa and relevant accession information. Sample ID is a unique identifier given by the collector. The complete spatio-temporal data can be accessed through the Parasites Division, Museum of Southwestern Biology host and/or parasite records maintained in Arctos. CAN = Canada, USA = United States of America, UA = Ukraine, NP = Nepal, ZA = South Africa, AR = Argentina, BR = Brazil.

<table>
<thead>
<tr>
<th>Taxa</th>
<th>Sample ID</th>
<th>Host</th>
<th>Locality</th>
<th>NCBI</th>
<th>MSB:Para:#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anserobilharzia brantae</td>
<td>W335</td>
<td>Gyraulus parvus</td>
<td>CAN</td>
<td>SRR19593566</td>
<td>14745</td>
</tr>
<tr>
<td>W351</td>
<td>Branta canadensis</td>
<td>CAN</td>
<td>SRR19593565</td>
<td>7984</td>
<td></td>
</tr>
<tr>
<td>W333</td>
<td>Gyraulus parvus</td>
<td>CAN</td>
<td>SRR19593554</td>
<td>14744</td>
<td></td>
</tr>
<tr>
<td>Austrobilharzia variglandis</td>
<td>SAL63.81</td>
<td>Larus delawarensis</td>
<td>CAN</td>
<td>SRR19593543</td>
<td>32451</td>
</tr>
<tr>
<td>SAL63.80</td>
<td>Larus delawarensis</td>
<td>CAN</td>
<td>SRR19593533</td>
<td>29053</td>
<td></td>
</tr>
<tr>
<td>Bilharziella polonica</td>
<td>W930</td>
<td>Anas platyrhynchos</td>
<td>UA</td>
<td>SRR19593532</td>
<td>32667</td>
</tr>
<tr>
<td>Bivitellobilharzia nairi</td>
<td>W465.2</td>
<td>Rhinoceros unicornis</td>
<td>NP</td>
<td>SRR19593531</td>
<td>29075</td>
</tr>
<tr>
<td>Dendritobilharzia pulverulenta</td>
<td>W926</td>
<td>Anas crecca</td>
<td>USA</td>
<td>SRR19593530</td>
<td>29034</td>
</tr>
<tr>
<td>W836</td>
<td>Anas discors</td>
<td>USA</td>
<td>SRR19593529</td>
<td>20795</td>
<td></td>
</tr>
<tr>
<td>Gigantobilharzia huronensis</td>
<td>W414/513</td>
<td>Physa sp.</td>
<td>USA</td>
<td>SRR19593528</td>
<td>18687/25488</td>
</tr>
<tr>
<td>Heterobilharzia americana</td>
<td>W678</td>
<td>Physa sp.</td>
<td>USA</td>
<td>SRR19593564</td>
<td>29074</td>
</tr>
<tr>
<td>W805</td>
<td>Procyn lotor</td>
<td>USA</td>
<td>SRR19593563</td>
<td>19286</td>
<td></td>
</tr>
<tr>
<td>Macrobilharzia macrobilharzia</td>
<td>W931</td>
<td>Anhinga anhinga</td>
<td>USA</td>
<td>SRR19593562</td>
<td>32668</td>
</tr>
<tr>
<td>Schistosomatium douthitti</td>
<td>SAL95_60</td>
<td>Stagnicola sp.</td>
<td>CAN</td>
<td>SRR19593561</td>
<td>2861</td>
</tr>
<tr>
<td>Schistosoma bovis</td>
<td>PM1</td>
<td>Bulinus sp.</td>
<td>KE</td>
<td>SRR19593560</td>
<td>32666</td>
</tr>
<tr>
<td>Schistosoma nasaleae</td>
<td>WS46</td>
<td>Indoplanorbis exustus</td>
<td>NP</td>
<td>SRR19593559</td>
<td>—</td>
</tr>
<tr>
<td>Schistosoma spindale</td>
<td>WS45</td>
<td>Indoplanorbis exustus</td>
<td>NP</td>
<td>SRR19593559</td>
<td>—</td>
</tr>
<tr>
<td>Schistosoma edwardsiense</td>
<td>W957</td>
<td>Biomphalaria sp.</td>
<td>KE</td>
<td>SRR19593557</td>
<td>32609</td>
</tr>
<tr>
<td>Schistosoma sp.</td>
<td>W565</td>
<td>Indoplanorbis sp.</td>
<td>NP</td>
<td>SRR19593556</td>
<td>—</td>
</tr>
<tr>
<td>Ornithobilharzia canaliculata</td>
<td>W393</td>
<td>Larus occidentalis</td>
<td>USA</td>
<td>SRR19593555</td>
<td>18542</td>
</tr>
<tr>
<td>Trichobilharzia querquedulae</td>
<td>W664</td>
<td>Spatula smithii</td>
<td>ZA</td>
<td>SRR19593553</td>
<td>19000</td>
</tr>
<tr>
<td>Trichobilharzia stagnicola</td>
<td>W233</td>
<td>Stagnicola emarginata</td>
<td>USA</td>
<td>SRR19593552</td>
<td>19509</td>
</tr>
<tr>
<td>Trichobilharzia cf. szidati</td>
<td>W620A</td>
<td>Lymnaea stagnalis</td>
<td>USA</td>
<td>SRR19593551</td>
<td>259</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. A</td>
<td>W613</td>
<td>Melanitta deglandi</td>
<td>USA</td>
<td>SRR19593550</td>
<td>25264</td>
</tr>
<tr>
<td>Avian Schistosomatidae lineage 2</td>
<td>W634</td>
<td>Chilina perrieri</td>
<td>AR</td>
<td>SRR19593549</td>
<td>7969</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. B</td>
<td>W399</td>
<td>Physa acuta</td>
<td>USA</td>
<td>SRR19593548</td>
<td>18677</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. C2</td>
<td>W402</td>
<td>Gyraulus parvus</td>
<td>USA</td>
<td>SRR19593547</td>
<td>18680</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. C4</td>
<td>W847</td>
<td>Biomphalaria glabrata</td>
<td>BR</td>
<td>SRR19593546</td>
<td>25514</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. C11</td>
<td>W607</td>
<td>Anas americana</td>
<td>USA</td>
<td>SRR19593545</td>
<td>25258</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. C11</td>
<td>W616A</td>
<td>Gyraulus sp.</td>
<td>USA</td>
<td>SRR19593544</td>
<td>19650</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. D</td>
<td>W342</td>
<td>Gyraulus parvus</td>
<td>CAN</td>
<td>SRR19593542</td>
<td>18619</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. G</td>
<td>W877</td>
<td>Ceratophyllus sp.</td>
<td>KE</td>
<td>SRR19593541</td>
<td>32612</td>
</tr>
<tr>
<td>Marinobilharzia patagonense</td>
<td>W637A</td>
<td>Siphonaria lessoni</td>
<td>AR</td>
<td>SRR19593540</td>
<td>18935</td>
</tr>
<tr>
<td>Avian Schistosomatidae lineage 3</td>
<td>C1</td>
<td>Chilina nequenensis</td>
<td>AR</td>
<td>SRR19593539</td>
<td>7970</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. M1</td>
<td>W216</td>
<td>Haminoea japonica</td>
<td>USA</td>
<td>SRR19593538</td>
<td>18660</td>
</tr>
<tr>
<td>Schistosomatidae sp. W688</td>
<td>W688</td>
<td>Indoplanorbis exustus</td>
<td>NP</td>
<td>SRR19593537</td>
<td>18710</td>
</tr>
</tbody>
</table>

**Outgroups**

| Vasotrema cf. robustum              | W411      | Physa sp.                 | USA      | SRR19593536   | 18690      |
| Vasotrema cf. robustum              | PS        | Physa sp.                 | USA      | SRR19593535   | 18715      |
| Spirochidiidae                      | W702      | Biomphalaria straminea    | BR       | SRR19593534   | 20804      |

**Published Genomes**

- Allobilharzia visceralis
- Schistosoma hematobium
- Schistosoma mansoni
- Schistosoma japonicum
- Schistosomatidum douthitti
- Trichobilharzia szidati
- Trichobilharzia regenti
- Echinostoma caproni

SAMEA2201407
SAMEA1920831
SAMEA242295
PRJEB461
PRJEB127
Table 3. Summary of sequence capture data. Sample IDs separated with an (/) indicate that samples were pooled to increase DNA quantities; in these samples DNA (µg) is additive. The Schistosomatid life stage is indicated as A (adult), M (miracidia) or C (cercariae). nq = not quantified.

<table>
<thead>
<tr>
<th>Taxa</th>
<th>Sample ID</th>
<th>Life Stage</th>
<th>DNA [µg]</th>
<th>No. of Contigs</th>
<th>No. of bp</th>
<th>UCEs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schistosomatidae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anserobilharzia brantae</td>
<td>W352/W340</td>
<td>C</td>
<td>0.562</td>
<td>1158</td>
<td>330,978</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>W335</td>
<td>C</td>
<td>nq</td>
<td>16,743</td>
<td>4,870,006</td>
<td>1334</td>
</tr>
<tr>
<td></td>
<td>W351</td>
<td>C</td>
<td>&gt;0.005</td>
<td>5907</td>
<td>1,836,149</td>
<td>551</td>
</tr>
<tr>
<td>Austrobilharzia varieglandis</td>
<td>W396/359</td>
<td>A</td>
<td>0.0352</td>
<td>145</td>
<td>35,487</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>W697</td>
<td>A</td>
<td>0.0193</td>
<td>8344</td>
<td>2,495,016</td>
<td>1440</td>
</tr>
<tr>
<td>Bilharziella polonica</td>
<td>W930</td>
<td>A</td>
<td>0.0155</td>
<td>17,081</td>
<td>5,897,771</td>
<td>1677</td>
</tr>
<tr>
<td>Bivitellobilharzia nairi</td>
<td>W465</td>
<td>M</td>
<td>0.078</td>
<td>18,455</td>
<td>5,052,402</td>
<td>1336</td>
</tr>
<tr>
<td></td>
<td>BN2011</td>
<td>M</td>
<td>nq</td>
<td>3696</td>
<td>1,162,332</td>
<td>47</td>
</tr>
<tr>
<td>Dendritobilharzia pulverulentia</td>
<td>W836</td>
<td>A</td>
<td>nq</td>
<td>26,978</td>
<td>8,405,837</td>
<td>1737</td>
</tr>
<tr>
<td></td>
<td>W926</td>
<td>A</td>
<td>0.0944</td>
<td>47,498</td>
<td>14,794,680</td>
<td>1631</td>
</tr>
<tr>
<td>Gigantobilharzia huronensis</td>
<td>W414/S13</td>
<td>C</td>
<td>0.0375</td>
<td>242,326</td>
<td>72,500,699</td>
<td>1650</td>
</tr>
<tr>
<td></td>
<td>W678</td>
<td>C</td>
<td>0.0325</td>
<td>48,623</td>
<td>16,046,122</td>
<td>1582</td>
</tr>
<tr>
<td>Heterobilharzia americana</td>
<td>W805</td>
<td>A</td>
<td>nq</td>
<td>1442,043</td>
<td>393,96,852</td>
<td>1784</td>
</tr>
<tr>
<td>Macrobilharzia macrobilharzia</td>
<td>W931</td>
<td>A</td>
<td>0.0472</td>
<td>3908</td>
<td>1,580,643</td>
<td>1590</td>
</tr>
<tr>
<td>Marinobilharzia patagonensis</td>
<td>W637</td>
<td>A</td>
<td>0.163</td>
<td>34,871</td>
<td>13,736,793</td>
<td>1799</td>
</tr>
<tr>
<td>Ornithobilharzia canaliculata</td>
<td>W393</td>
<td>A</td>
<td>0.015</td>
<td>31,569</td>
<td>8,873,303</td>
<td>1441</td>
</tr>
<tr>
<td>Schistosomatium douthitti</td>
<td>SAL95.60</td>
<td>C</td>
<td>nq</td>
<td>15,755</td>
<td>4,482,586</td>
<td>597</td>
</tr>
<tr>
<td>Schistosoma bovis</td>
<td>PM1</td>
<td>C</td>
<td>nq</td>
<td>6316</td>
<td>4,181,809</td>
<td>1873</td>
</tr>
<tr>
<td>Schistosoma nasale</td>
<td>W546</td>
<td>C</td>
<td>0.0107</td>
<td>41,764</td>
<td>14,799,490</td>
<td>1830</td>
</tr>
<tr>
<td>Schistosoma spindale</td>
<td>W545</td>
<td>C</td>
<td>0.0662</td>
<td>15,206</td>
<td>5,394,821</td>
<td>1885</td>
</tr>
<tr>
<td>Schistosoma edwardienne</td>
<td>W957</td>
<td>C</td>
<td>nq</td>
<td>21210</td>
<td>8,383,839</td>
<td>1800</td>
</tr>
<tr>
<td>Schistosoma sp.</td>
<td>W555</td>
<td>C</td>
<td>0.1455</td>
<td>15,865</td>
<td>7,245,729</td>
<td>1862</td>
</tr>
<tr>
<td>Trichobilharzia quercucladeae</td>
<td>W929</td>
<td>A</td>
<td>0.0224</td>
<td>70,532</td>
<td>21,183,008</td>
<td>1576</td>
</tr>
<tr>
<td>Trichobilharzia stagnicola</td>
<td>W233</td>
<td>C</td>
<td>0.0531</td>
<td>18,675</td>
<td>13,662</td>
<td>658</td>
</tr>
<tr>
<td>Trichobilharzia cf. szidati</td>
<td>W620A</td>
<td>C</td>
<td>nq</td>
<td>48,011</td>
<td>15,967,522</td>
<td>1123</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. A</td>
<td>W613</td>
<td>A</td>
<td>0.091</td>
<td>10,458</td>
<td>3,571,057</td>
<td>1511</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. B</td>
<td>W399</td>
<td>C</td>
<td>0.0905</td>
<td>9349</td>
<td>2,995,586</td>
<td>1115</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. C</td>
<td>W402</td>
<td>C</td>
<td>0.0845</td>
<td>15,168</td>
<td>4,939,724</td>
<td>849</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. C4</td>
<td>W847</td>
<td>C</td>
<td>nq</td>
<td>35,118</td>
<td>10,814,130</td>
<td>1732</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. C11</td>
<td>W607</td>
<td>C</td>
<td>0.265</td>
<td>34,656</td>
<td>53,303</td>
<td>1599</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. C11</td>
<td>W616A</td>
<td>C</td>
<td>0.084</td>
<td>32,003</td>
<td>10,790,620</td>
<td>1409</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. D</td>
<td>W342</td>
<td>C</td>
<td>0.43</td>
<td>13,126</td>
<td>3,804,186</td>
<td>171</td>
</tr>
<tr>
<td>Avian Schistosomatidae sp. G</td>
<td>W877</td>
<td>C</td>
<td>nq</td>
<td>1227</td>
<td>356,239</td>
<td>330</td>
</tr>
<tr>
<td>Avian Schistosomatidae lineage 2</td>
<td>W634</td>
<td>C</td>
<td>nq</td>
<td>6368</td>
<td>1,553,874</td>
<td>419</td>
</tr>
<tr>
<td>Avian Schistosomatidae lineage 3</td>
<td>C1</td>
<td>C</td>
<td>nq</td>
<td>2576</td>
<td>764,111</td>
<td>458</td>
</tr>
<tr>
<td>Schistosomatidae sp. W688</td>
<td>W688</td>
<td>C</td>
<td>0.0337</td>
<td>8006</td>
<td>3,890,904</td>
<td>1428</td>
</tr>
<tr>
<td>Non-Schistosomatidae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vasotrema cf. robustum</td>
<td>W411</td>
<td>C</td>
<td>nq</td>
<td>17,681</td>
<td>4,449,659</td>
<td>372</td>
</tr>
<tr>
<td>Spirorchidae</td>
<td>W702</td>
<td>C</td>
<td>nq</td>
<td>47,339</td>
<td>15,417,172</td>
<td>885</td>
</tr>
<tr>
<td>Vasotrema cf. robustum</td>
<td>PS</td>
<td>C</td>
<td>nq</td>
<td>41,958</td>
<td>12,949,767</td>
<td>473</td>
</tr>
</tbody>
</table>

All recovered loci were mapped to their presumptive chromosomal location within the *Schistosoma mansoni* genome. Of 139 UCE loci that were mapped to the Z-chromosome (Supplementary File S4), 85 were recovered for ≥ 70% of the taxa (n = 35) and used for phylogenetic reconstruction (Figure 2). Loci were represented in 56–90% of the taxa. Moderate Z-chromosome UCE loci representation was obtained in pivotal taxa such as *Macrobilharzia macrobilharzia* (63% of 139 loci), Schistosomatidae sp. W688 (62%), *Heterobilharzia americana* (68%), and *Bivitellobilharzia nairi* (60%).

Both the maximum likelihood and Bayesian inference analyses of Z-chromosome UCE loci yielded topologies congruent with the analysis of genome-wide loci (Figures 1 and 2). In analyses of the Z-chromosome UCE loci, however, nodes across a range of divergence times were better resolved than in the analyses of genome-wide loci. This increase in resolution was particularly pronounced in nodes with a bootstrap support of <100 or a
posterior probability of <1 in the analysis of genome-wide loci. Bayesian inference of Z-chromosome UCEs resolved all pivotal nodes (Table 1), including strong support for the position of *Macrobilharzia + Schistosomatidae* sp. W688 at the base of the SB clade. The corresponding maximum likelihood analysis of Z-chromosome UCEs also placed *Macrobilharzia + Schistosomatidae* sp. W688 at the base of the SB clade, but with weak support (66% of bootstrap replicates).

![Figure 2](image)

**Figure 2.** Bayesian inference analysis performed in BEAST 2 of 85 concatenated Z-chromosome (UCE) loci. Blue node circles indicate posterior probability values of 1, black circles indicate 0.99. Maximum likelihood (RAxML) analysis provided concordant support values with the following exceptions: bootstrap support for (SB + M) = 66%; ((SB + M) (HS + DAS)) = 93%; (Schistosoma c.f. szidati + T. stagnicolae) = 95%. SB = *Schistosoma + Bivitellobilharzia*; M = *Macrobilharzia + Schistosomatidae* sp. W688; HS = *Heterobilharzia + Schistosomatium*; DAS = *Anserobilharzia, Allobilharzia, Trichobilharzia, Dendritobilharzia, Gigantobilharzia, Bilharziella* and all sampled Avian Schistosomatidae sp. Tree was edited in Fig Tree v 1.4. The gray open circle denotes the major divergence of the derived schistosomatids. The two gray arrows denote the two major derived nodes leading to two simultaneous and independent radiations, one predominantly in birds, the other predominantly in mammals.

### 2.3. UCE Loci Sequence Similarity among Blood Flukes

This study included four non-schistosomatid blood fluke taxa (Table 2): three members of *Spirorchidae* (W411, W702 and PS) and one *Aporocotylidae* (W962, not included in phylogenetic analyses). Within the total supermatrix dataset of 554 loci, 33% (W411), 40 (PS) and 61% (W702) loci were recovered from these samples and were found to be phylogenetically informative. Although this varying ‘completeness’ is not necessarily a function of phylogenetic relatedness, it shows the recovery of a reasonable number of non-schistosomatid loci using a *Schistosoma*-derived bait set (Supplementary File S1). Additionally, a randomly selected (though uniform across samples) subset of 25 loci from...
Spirochiidae (W411, W702 and PS) and Aporocotylidae (W962) were on average 82.3% (79.50–85.13%) similar to selected Schistosomatidae (Table 4), with about 3.9% greater similarity in Spirochiidae than Aporocotylidae.

Table 4. Average nucleotide similarity (%) between Schistosomatidae, Spirochiidae and Aporocotylidae for 25 UCE loci.

<table>
<thead>
<tr>
<th></th>
<th>Spirochiidae</th>
<th>Aporocotylidae</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Trichobilharzia regenti</em></td>
<td>84.90</td>
<td>80.24</td>
</tr>
<tr>
<td><em>Heterobilharzia americana</em></td>
<td>84.30</td>
<td>79.50</td>
</tr>
<tr>
<td><em>Schistosoma mansoni</em></td>
<td>83.60</td>
<td>80.84</td>
</tr>
<tr>
<td><em>Macrobilharzia macrobilharzia</em></td>
<td>85.13</td>
<td>80.48</td>
</tr>
<tr>
<td>AO Clade</td>
<td>83.40</td>
<td>80.56</td>
</tr>
</tbody>
</table>

3. Discussion

This is the first study to apply reduced representation phylogenomic methods to address the interrelationships and patterns of diversification within the Schistosomatidae. We reconstructed and compared phylogenies based on 554 nuclear UCE loci (4,780,079 bases). We generated the first phylogenetic tree, for any helminth group, based on Z-chromosome-specific loci (sex chromosome). Based on material available from extant species, our results suggest that schistosomatids first appeared in marine birds and gastropods. They later colonized freshwater snails and both birds and mammals associated with freshwater. Two separate acquisitions of mammalian hosts are supported. From within a diverse freshwater-based lineage of avian schistosomatids, some representatives have secondarily colonized marine gastropod and avian hosts.

Past studies using few or single loci to reconstruct the generic relationships within Schistosomatidae often failed to resolve deeper nodes, particularly those significant to diversification by host switching. Herein, we report the resolution of several pivotal nodes and thus an improved understanding of schistosomatid diversification, mainly in the context of intermediate host use. Relative to other taxa [48,58,65], we noted the reduced recovery of UCE loci, similar to what was found in with the only other UCE digenean phylogeny [47], but a resolution of pivotal nodes was achieved in our study. Below we discuss how these results offer insight into the timing of schistosomatid diversification events and host use.

3.1. Resolution of Pivotal Nodes and Host Divergences

Resolved phylogenetic trees are necessary to estimate the timing of diversification events. However, without available fossil data, no primary calibration points exist to estimate divergence times for schistosomatids or other helminths. Parasitologists often use host divergence data to derive secondary calibration points [2,66], but this assumes that parasites have evolved within the bounds of particular hosts. This assumption might be sound where parasites and hosts show evidence of co-phylogeny as the result of a long co-evolutionary history [67–69]. However, co-cladogenesis among multi-host parasites appears to be uncommon [1,2,70], particularly in groups with evident switching among distantly related hosts, which erodes co-evolutionary signatures [71]. Using the timing of host divergence as a proxy for parasite divergence calibration therefore requires caution. With this in mind, discussed below are hypothesized host-divergence dates at pivotal nodes which may provide some hard bounds for particular nodes within the Schistosomatidae.

3.1.1. AO Clade

All UCE analyses support the AO clade as a sister to the rest of the Schistosomatidae, suggesting an early diverging position. Members of *Austrobilharzia* and *Ornithobilharzia* infect marine birds and snails. Schistosomatids have been hypothesized to have first diverged
from spirorchiids (Spirorchiidae), modern day parasites of marine or freshwater turtles, which are the well supported sister clade to Schistosomatidae [18,27,72]. Extant turtle lineages are estimated to have originated during the mid-late Triassic (240–200 mya, [73]), which may provide an upper limit for the origins of the Schistosomatidae. One of the hallmarks of schistosomatid biology is their adoption of fully intravascular habitats in endothe- rms, including adult worms living in the liver, portal vein and mesenteric veins where nutrient levels are high. Dioecy is viewed as a consequence of the need to have a muscular male to move against the prevailing flow of blood in the portal system and a slender-bodied female to deposit eggs in the intestinal wall. These essential features may have their origins in partially endothermic turtles such as leatherbacks, but may have flourished in fully endothermic hosts, including theropod dinosaurs and/or their descendants, birds. The closest relatives we know of for these proto-schistosomatids are members of the AO clade found in marine charadriiform birds (Figure 1) and marine caenogastropods (e.g., Potamididae, Batillariidae, Nassariidae, Littorinidae) [21]. The snail family Nassariidae is estimated to have diverged roughly 120 mya [74], while Charadriiformes are thought to have diverged between 23–16 mya [75]. Schistosomatids later colonized freshwater aquatic or amphibious habitats and diverged in the two major clades we see today, SB + M and DAS + HS.

3.1.2. Hypothesized Early Gastropod Host Use Prior to the (M + SB) and (HS + DAS) Split

Some evidence from contemporary gastropod host use among schistosomatids suggests that prior to the (M + SB) and (HS + DAS) split, the prevailing gastropod hosts may have been snails in the family Planorbidae (Figure 3). In support of this, (1) although natural snail hosts are unknown for Bivitellobilharzia, experimental hosts were found to be planorbids [76]; (2) planorbids are prominently (but not exclusively) represented as intermediate hosts among Schistosoma species; (3) at least one representative of the M clade is hosted by a planorbid (Schistosomatidae sp. W688) [41]; (4) the earliest known diverging member of the DAS clade, Bilharziella polonica, has a planorbid snail host, as do many other members of the this clade [21]. A role for lymnaeid snails in hosting ancestral schistosomatids should not be discounted either: (1) a few species of Schistosoma use lymnaeid snail hosts; (2) the HS clade use exclusively lymnaeids; (3) they are prominently represented as snail hosts in the DAS clade. Caenogastropods too deserve consideration, as snail hosts prior to the split between the SB + M and DAS + HS clades, insofar as all members of the early diverging Schistosoma japonicum species group utilized Pomatiopsidae snails. Davis [77] suggested that proto-Schistosoma had co-diversified with amphibious caenogastropods.

3.1.3. SB Clade

Branch lengths in the phylogeny (Figure 1) suggest that the SB + M and DAS + HS clades may have radiated simultaneously. The SB clade represents a substantial radiation within mammals (28 species). Contemporary members of the SB clade infect a broad range of mammalian groups: Bovidae, Canidae, Elephantidae, Felidae, Hippopotamidae, Hominidae, Rhinocerotidae, Suidae, and Rodentia. Relationships among members of the SB clade have been largely resolved [41,64,78]. Schistosoma and Bivitellobilharzia have been considered to be sister genera despite conflicting phylogenies (Table 1), and all UCE datasets support the monophyly of Afro-Eurasian mammalian schistosomatids. Snyder and Loker [25] hypothesized that Schistosoma originated in the mid-Miocene (16–11.6 mya), based largely on estimates by Davis [77] using fossil data to estimate the divergence of Pomatiopsidae snails, hosts to the extant S. japonicum clade. The suspected role of this snail family in the diversification of Schistosoma led to the use of mt DNA to estimate pomatiopsid divergence between 12–5 mya [79]. Going forward, the identification of the natural intermediate hosts of Bivitellobilharzia and Macrobilharzia represents important collection goals, as this will help to evaluate the likelihood of the various scenarios highlighted above—planorbid, lymnaeid or pomatiopsid species as ancestral hosts for proto-Schistosoma, with possible attendant changes in the timing of molecular divergence among schistosomatids [45].
Figure 3. Hypothesized ancestral intermediate host use and host-switching events within the Schistosomatidae. Circles, colored corresponding to the legend, representing hypothesized intermediate host use, were mapped onto the Figure 1 phylogeny. Open circles represent less certainty relative to solid circles.

3.1.4. Macrobilharzia (M Clade)

Although branch support values were higher for the Z-chromosome phylogeny (Figure 2) than the concatenated UCE phylogeny (Figure 1), both support an affinity between the M and SB clades (Figures 1 and 2). This suggests a derived rather than an early branching position for the M clade within the Schistosomatidae. Large ribosomal subunit (28S) phylogenies [21,39,41] place Macrobilharzia at the base of the SB + DAS clades, thus diverging earlier than the majority of avian schistosomatid species. Mitochondrial cox1 studies, however, provide less consistent results [39,40], possibly due to the saturation of third codon positions [80,81]. Overall, there is a lack of sufficient detail about intermediate and definitive host use patterns and the cercarial anatomy for Macrobilharzia. For instance, whereas the cercariae for the M clade member W688 possess eyespots, cercariae within the SB clade do not. It is not known if cercariae of Macrobilharzia spp. have eyespots or not.

3.1.5. DAS Clade

Efforts to describe schistosomatid diversity have yielded numerous novel DAS lineages [21,28,30,31,33]. Intermediate host switching is evidently recurring within DAS and can be hypothesized as a probable mechanism of diversification; at least eight intermediate host-switching events have occurred within the DAS clade, at least two of which have been from freshwater to marine environments [21]. Planorbidae species are predominant hosts within DAS, and are associated with Bilharziella, Dendritobilharzia, Avian Schistosoma...
matid C [31], Avian Schistosomatid D (this study), and Anserobilharzia (Figure 3). Caution as to snail host use among early, and likely extinct, representatives of the DAS clade is warranted because representatives of at least six other gastropod families are known to be exploited collectively by extant members of the DAS clade, which are not all represented in Figure 3 [21,29,30].

The monophyly of DAS is consistently supported in published phylogenies (Table 1), including this study. However, this study provided increased resolution within the DAS clade. Bilharziella polonica and Avian Schistosomatidae 2 were strongly supported as being basal to DAS. Bilharziella polonica infects a broad taxonomic range of water birds (Anseriformes, Ciconiiformes, Gruiformes, Podicipediformes), and this may suggest that ancestors to the DAS + HS clade similarly infected a broad host range, as might be expected during a transition from a marine to freshwater environment. UCE analysis (Figure 1) supports the breakdown of the remaining DAS taxa into three primary clades.

- Clade 1: Trichobilharzia, Allobilharzia and Anserobilharzia form a well-supported sub-clade within DAS, reported exclusively from Anseriformes, predominantly ducks. Published single loci studies do not consistently resolve the node connecting Trichobilharzia and Allobilharzia + Anserobilharzia [28,82–84]. This is a logical grouping as this diverse radiation appears to have occurred in Anseriformes [28,83,84], and Trichobilharzia demonstrates a strong association with anatids. The diversification of modern Anatidae has been estimated to have occurred between 25–5 mya [85], which might be used as a hard bound for the radiation of Trichobilharzia.

- Clade 2: A second sub-clade within DAS is composed of several lineages retaining the names used from previous studies, Avian Schistosomatidae C (C2, C4, and C11 were sampled, [31]). Avian Schistosomatidae C is a complex of several species distributed globally (Americas, Europe, and Africa), none of which have yet been formally named [31].

- Clade 3: Gigantobilharzia, Dendritobilharzia, and six undescribed genera form Clade 3. Branch lengths within this clade, specifically the distance between tips, are large, which may suggest missing taxa. Remarkably, the taxa in Clade 3 (Figure 1) are hosted by at least four different snail families (including one marine family) and three orders of birds (Passeriformes, Charadriiformes, and Anseriformes) recovered from Argentina, Kenya, and across North America.

Anseriformes are a common host group throughout DAS and are utilized by species of Dendritobilharzia, Avian Schistosomatid A, Avian Schistosomatid C, and Avian Schistosomatid D and Trichobilharzia, possibly the most speciose schistosomatid genus [21,28].

3.1.6. HS Clade

UCE analyses provided strong support for a sister relationship between the HS and DAS clades, and accordingly a second independent switch into mammals (as posited by Snyder and Loker [25]). Based on topology (Figures 1 and 2), one can hypothesize that DAS and HS shared a recent common ancestor, likely in what is now North America. Heterobilharzia and Schistosomatium are both monotypic mammal-infecting genera, exclusive to North America, and are strikingly depauperate relative to extant DAS or SB taxa.

HS may have been historically more speciose, as the loss of the North American mammalian megafauna led to the extinction of many lineages relative to the mammalian megafauna in Africa and Asia [86], which did not experience similar losses. Schistosomatium douthitti is a rodent parasite whereas Heterobilharzia americana is primarily a raccoon (Procyonidae) and dog (Canidae) parasite, though it has been reported from a broad spectrum of mammals [87–89]. Both depend solely on lymnaeid snail hosts. Species-level investigations are required to verify if H. americana is truly a single species and not a complex of cryptic species [89]. Correa et al. [90] reconstructed phylogenetic relationships among Lymnaeidae and demonstrated a distinct split among North American and Old World lymnaeids, and an accurate date for this split could provide a hard bound on when the HS clade evolved in North America.
3.2. Diversification via Extensive Intermediate Host-Switching

Across the Schistosomatidae, definitive host associations appear less constrained, at the level of host order or below, relative to intermediate host associations [8,91,92]. Our study supports the idea that population isolation via intermediate host-switching mediates schistosomatid speciation [21,93,94]. Relative to schistosomatids, which infect 16 families within Caenogastropoda and Heterobranchia [21], other prominent digenean families engage relatively few gastropod families: for instance, both the Fasciolidae and Paramphistomidae each exploit two gastropod families [95,96]). Schistosomatids are more similar in this regard to their two closely related blood fluke families, the Spirorchiidae and Aporocotylidae. Spirorchiids have been reported from Vetigastropoda, Caenogastropoda, Heterobranchia, and Annelida [97]. Aporocotylids infect Annelida, Bivalva, Caenogastropoda, Heterobranchia, and Neritomorpha [98].

Most obviously within the DAS clade, intermediate host-switching events are characterized by short inter-nodal branches and do not appear to be associated with deep divergence events. This suggests that such host-switches can occur relatively rapidly and frequently over evolutionary time, and genetic differences among lineages (especially for nuclear loci) have not had time to accumulate.

Why intermediate host-switches are seemingly so numerous within Schistosomatidae is an open question, one understandably hard to capture with experimental studies. Some hypotheses are as follows: (1) The facilitated host-switching hypothesis—intermediate host-switches may be facilitated by coinfections with other snail parasites [21,99]. For instance, at least some Austrobilharzia species seem to specialize in actively exploiting the presence of other trematode larvae to colonize their snail hosts [100], and in some cases prior trematode infections enable schistosomatids to exploit new host species [101]. (2) The ecological fitting hypothesis would suggest that schistosomatid larvae may retain pre-adaptations, evolved from ancestral host-use, that enable them to infect different gastropod lineages [5]). The evolutionary potential hypothesis (this study) proposes that schistosomatid species maintain high genetic diversity and large effective population sizes, favoring the presence of rare alleles that might confer infectivity in a new gastropod lineage, as observed in population level studies of Schistosoma japonicum [92,102,103], Trichobilharzia spp. [10], and to a lesser extent S. mansoni [103–107]. One expectation may be that schistosomatids of migratory avian hosts have large effective population sizes and high genetic diversity, and therefore adaptive potential [108,109], which may favor switching into new snail hosts [94]. That schistosomatid clades associated with migratory birds have high rates of host switching is likely also related to bird vagility [10,110], which frequently moves parasites to new habitats where they have the opportunity to encounter different snail lineages. For these reasons, one might hypothesize that host-switching is a predominant mode of diversification among parasites obligate to migratory hosts.

3.3. Phylogenomic Considerations and Future Directions

3.3.1. Phylogenomics of Schistosomatids

The diminutive size and location of schistosomatids within their hosts (venous or rarely arterial system) present significant challenges to next-generation sequencing applications and downstream phylogenomic analysis. Species of several novel, unnamed genera were not included in this dataset despite considerable effort, due to inadequate specimen quality; their future inclusion may increase resolution still more. We found that, in general, adult worms yielded the most DNA and average number of reads, and the best capture success for multiple loci. Notably, among cercarial samples, none of these measures improved, on average, by increasing the number of cercariae extracted, suggesting this pattern is not solely related to amount of starting tissue.

Modifications of our sequence capture and library preparation protocols (see the Methods section) increased the recovery rate of UCE loci (a three-fold increase). Future studies performing sequence capture on organisms with a low quantity of or degraded museum specimens should consider incorporating these modifications. Large amounts of sequence
data are fairly robust to missing data [111], and it has been demonstrated that incomplete taxon sampling is more unfavorable to phylogenetic inference than missing data [112]. However, the extent of missing data within our final alignments prevented species tree estimation [113–115] and statistical assessments of incomplete lineage sorting [43], both of which are the logical next steps to furthering our understanding of schistosomatid evolution. Moving forward, it would be advisable to design a new bait kit to specifically target the most common 1500 UCE loci recovered and increase bait tiling density to improve recovery of UCE loci, thereby reducing missing data.

This study supports the utility of Z-chromosome loci for the phylogenetic reconstruction of schistosomatids. Sex-chromosome markers have proven to be valuable phylogenetic tools among vertebrates [115,116], often performing better than autosomal markers in resolving difficult nodes [43]. The schistosomatid Z-chromosome is the homogametic male chromosome (females ZW), and thereby has a lower effective population size, reduced recombination, and faster sorting rate relative to autosomal markers [115]. Sex-chromosome markers have been shown to improve resolution when incomplete lineage sorting is prevalent [43]. As taxon sampling increases, Z-specific loci may prove to be useful markers for future diversification studies. An ideal strategy might be to design a bait set targeting a greater number of Z-specific loci.

Our analyses (Table 4, Supplementary File S1) demonstrate that our bait set contained phylogenetically informative loci across blood flukes, and even across digeneans. We conclude that the application of UCE loci shows promise in resolving higher level taxonomic questions, such as relationships within Diplostomoidea [47,61] or across Digenea [117,118].

3.3.2. Guiding Future Collection Efforts

Additional schistosomatid biodiversity remains to be discovered. Brant and Loker [21] posit the likelihood of undiscovered schistosomatid lineages within marine environments; several studies support this within the AO clade [119–122]. Surveys of marine gastropod genera such as Haminoea [121] and Siphonaria [30,35] (Figure 1) have yielded schistosomatids, but these represent more recent divergences rather than early branching members of the AO clade (Figure 3). The recent, surprising discovery of an aporocotylid fish blood fluke from a dugong [123] provides further impetus to search for additional schistosomatids in marine environments.

The long branch lengths between Macrobilharzia and Schistosomatidae sp. W688 may be suggestive of missing taxa; unconfirmed species of Macrobilharzia have been reported from Africa and Asia [41,124]. Macrobilharzia macrobilharzia infects suliform birds, which have not been sampled as extensively as waterfowl, and would be the logical place to continue investigations into Macrobilharzia. Another possibility is that proto-M clade members may have once been represented by many more taxa inhabiting what is recognized as an ancient group of birds, which has been in existence since at least the Paleogene, 66–56 mya [76].

4. Materials and Methods

4.1. Taxon Sampling

Specimens used for this study are vouchered at the Parasites Division, Museum of Southwestern Biology, University of New Mexico, in Albuquerque New Mexico USA. Locality information, host data and museum catalogue numbers are summarized in Table 2 and can be accessed via the Arctos database.

Schistosomatid samples were collected as described in Brant and Loker [28] and Ebbs et al. [10]. Adult and larval schistosomatids ( cercariae, sporocysts and miracidia) were collected between 1995–2017, and were preserved in 95% ethanol or RNAlater. Thirteen of the 17 named genera and 11 currently undescribed lineages were included in this study. The undescribed lineages are related to novel genera or species that lack adult specimens or adult fragments with morphological features for discrimination. Most of the specimens used in this UCE study have been included in previously published schistoso-
matid molecular phylogenies (28S, ITS1, cox1 markers) and have originated from various collectors over the past two decades. This manuscript retained the taxa designations and/or collector IDs from the original publications (see Table 2) to facilitate the cross comparison of tree topologies and taxon identification.

UCE probes were designed using the Schistosoma mansoni genome (described in detail below); the resultant set was used to mine published genomes. Fourteen schistosomatid taxa and five outgroup taxa (Supplementary File S1) were mined for UCE loci and included within this study. The number of recovered loci relates to the phylogenetic similarity of the taxa with S. mansoni and the quality and completeness of the published SRA (Sequence Read Archive) data.

4.2. Sample Preparation

Several extraction protocols were used to obtain template DNA of sufficient quantity and quality, which is often a limiting factor in the number and quality of the UCE loci recovered from museum specimens [125,126]. Most samples were extracted using silica-based minipreps from either the QIAamp DNA Micro Kit (Qiagen), or the DNeasy Blood and Tissue kit (Qiagen). The QIAamp DNA Micro Kit, which is optimized for small amounts of tissue, performed best across schistosomatid samples. Sample DNA was quantified using Qubit Fluorometric quantification (Thermo Fisher, Waltham, MA, USA) using manufacturer buffers and protocols. Samples selected for sequence capture yielded >0.005–0.5 micrograms of DNA. Sample quality was assessed using a Bioanalyzer (Aligent 2100), however the DNA quantity was generally too low to accurately gauge the fragment size, which varied across samples, with some samples showing significant amounts of degradation and others relatively little. To prevent further degradation, vortexing, sonication, and the thawing/re-freezing of samples was minimized during handling and DNA extraction.

4.3. Sequence Capture of Ultra-Conserved Elements

In total, 40 samples were selected for targeted sequence capture of UCEs (Tables 2 and 3), specifically to address relationships within the Schistosomatidae and to resolve deeper nodes that have evaded phylogenetic placement in previous studies (Table 1). A custom bait set (18,550 baits, 120 nucleotides in length, 2× tiling density) was designed using the Schistosoma mansoni genome as a reference. Approximately 4000 UCE loci were targeted. The bait set was manufactured by Arbor Biosciences (www.arborbiosci.com, accessed on 16 February 2017).

Library enrichment procedures for the MYcroarray MYBaits kit (MYbaits® Manual v. 3.02) were followed, but with several modifications to the standard sequence capture and library preparation protocols to accommodate low amounts of DNA and a variable fragment size. To optimize sample preparation protocols, our samples were divided into four distinct sequence capture experiments and sequencing runs (R1, R2, R3 and R4). R1 used standard sticky-end library preparation coupled with standard amplification polymerase. R2–R4 employed blunt end library preparation chemistry and an uracil non-stalling amplification polymerase. This step aimed to reduce adapter dimers, which were abundant in R1 samples. For all runs, a size-selection step following library preparation was not performed due to low DNA quantity. Between R1 and R2–R4, hybridization temperatures were modified (62 or 65°C, respectively). Post-capture libraries were amplified for 12 cycles. Sequencing of paired-end, 100 bp reads was conducted on an Illumina HiSeq 2000. All sequence data for taxa listed in Table 2 have been made available on NCBI GenBank.

4.4. Processing and Alignment of UCE Data

Quality control of the raw reads included trimming adapter contamination and low-quality bases from reads, using the program Trimmomatic [127] and a 4-bp sliding window, quality score 20, and 36-bp minimum read length. Clean reads were processed following the PHYLUCE software package [128], including contig assembly using Trinity [129]. Assembled contigs were then matched to the probe set using lastz [130]. Previously
published genomes were downloaded from the NCBI ftp site. For species lacking assembled draft genomes, but for which data were available, we downloaded the Sequence Read Archive (SRA) (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3013647/, accessed on 30 January 2017) using the NCBI SRA-toolkit (https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=toolkit_doc, accessed on 30 January 2017) in fastq format (fastq-dump in SRA toolkit) and trimmed the data following the steps described above. Draft genome sequences were assembled using the MEGAHIT program [131]. UCE loci were mined from genome sequences following the PHYLUCE protocol. Samples with low capture success, where fewer than 100 UCE loci were recovered, were excluded from subsequent analyses. Alignments were created using MAFFT v. 2.2.7 [132]. Among the outgroup taxa mined for UCE loci (Table 3), the *Echinostoma caproni* genome (PRJEB1207) had the highest number of shared loci and was selected as the outgroup.

All the shared UCE loci for Schistosomatoidea + outgroup (*n* = 47,554 loci, 4,780,079 bases) were aligned for the purposes of phylogenetic reconstruction (Figure 1). A second dataset was limited to loci located on the schistosomatid sex chromosome, the Z-chromosome (*n* = 35, 85 loci, 937,745 bases). Sex chromosomes, due to their reduced recombination rates and effective sizes, have been shown to make excellent phylogenetic markers, often resolving nodes that autosomal loci fail to resolve [43]. Based on our UCE probe design, 139 UCE loci were mapped to the *Schistosoma mansoni* Z-chromosome (PHYLUCE [50]). Of the 139 recovered, 85 loci were shared among ≥70% of the sampled taxa and were aligned for phylogenetic reconstruction. For all datasets, loci were concatenated into a single supermatrix [50,128,132,133].

### 4.5. Alignment Building and Phylogenetic Reconstruction

All alignments were unpartitioned and analyzed in RAxML v.8.0.19 [60] using the GTRGAMMA model of evolution, with 500 thorough bootstrap replicates [50,56] for maximum likelihood analysis. Bayesian inference analysis was performed in BEAST 2 [134] consisting of two replicated runs with four Markov chain Monte Carlo (MCMC) chains, as well as one cold and three heated chains. Each analysis ran for 10,000,000 generations and was sampled every 1000 generations. Likelihood parameters and convergence between runs were assessed using the program Tracer v.1.6 [135] based on EES values greater than 200. The first 2500 trees from each analysis were discarded as burnin. The resulting phylogenetic trees were visualized and manipulated using Fig Tree v. 1.4.4 (https://github.com/rambaut/figtree/ accessed on 30 January 2017).

### 5. Conclusions

Through the analysis of 554 nuclear UCE loci, and a subset of 85 Z-chromosome specific UCE loci, we were able to resolve many pivotal interrelationships within Schistosomatidae, representing the most comprehensive family-level phylogeny to date. Some nodes failed to be resolved or were weakly supported. Further resolution of the two primary radiations (SB + M and DAS + HS) resulting in derived schistosomatid diversity, may be challenging for two possible reasons. (1) Contemporary lineages might have radiated simultaneously and rapidly, resulting in incomplete lineage sorting [42], which could have led to a hard polytomy [136,137]. As mentioned previously, population studies on members of both primary radiations suggest high genetic diversity and effective population sizes, both of which would contribute to incomplete lineage sorting [42,138,139]. (2) Colonization–extinction dynamics are common within parasites [1], and contemporary lineages might be more closely related to one or many extinct lineages than to their most closely related extant taxa. A combination of these factors might work in concert to make further resolution of Schistosomatidae, and potentially other helminth taxa, difficult. Nevertheless, a remarkable history of intermediate host-switching is evident, and indicative of a role as a primary driver of schistosomatid diversification.
**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/pathogens11070769/s1, Supplementary File S1: Table of all published digenean genomes mined for UCE loci; Supplementary File S2: Likelihood scores of RAxML phylogenies with varying levels of missing data; Supplementary File S3: Excel file of the presence and absences of the 554 UCE loci by taxa; Supplementary File S4: Excel file of the presence and absences of the 85 Z-chromosome specific UCE loci by taxa.


**Funding:** This work was primarily funded through a National Science Foundation, (NSF) USA grant to S.V.B. (DEB 1021427) and a National Institutes of Health grant R37AI101438 to E.S.L. S.A.L. was supported by NSF grant DEB 1845021. Funding to HAP National Council for Scientific and Technological Development (CNPq), Brazil (research scholarship to HAP).

**Data Availability Statement:** The data presented in this study are openly available on the NCBI Sequence Read Archive.

**Acknowledgments:** The authors would like to thank three anonymous reviewers for insightful comments regarding this manuscript. Thank you to the following for invaluable help collecting schistosomatid specimens; Museum of Southwestern Biology Bird Division, Andy Johnson, Robert Dickerman, and Chris Witt; National Museum Bloemfontein: Nick Nuttall and Davie H. de Swardt; James D. Liberatos; Robert E. Wilson; Martina Laidemitt; Keith Keller; Emily Sarvis. We acknowledge technical support from the University of New Mexico Molecular Biology Facility, which is supported by National Institutes of Health, USA grant P30GM110907 from the Institute Development Award program of the National Center for Research Resources, USA.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

3. Glodosky, C.M.; Sandland, G.J. Assessing host competency between native and invasive snail species exposed to the native parasite *Echinostoma revolutum*. *Aquat. Invasions* 2014, 9, 1. [CrossRef]


25. Snyder, S.D.; Loker, E.S. Evolutionary relationships among the Schistosomatidae (Platyhelminthes: Digenea) and an Asian origin for *Schistosoma*. *J. Parasitol.* 2000, 86, 283–288. [CrossRef]


30. Brant, S.V.; Loker, E.S.; Casalins, L.; Flores, V.R. Phylogenetic placement of a schistosome from an unusual marine snail host, the false limpet (*Siphonaria lessoni*) and gulls (*Larus dominicanus*) from Argentina with a brief review of marine schistosomes from snails. *J. Parasitol.* 2017, 103, 75–82. [CrossRef] [PubMed]


32. Gordy, M.A.; Hanington, P.C. A fine-scale phylogenetic assessment of digenean trematodes in central Alberta reveals we have yet to uncover their total diversity. *Ecol. Evol.* 2019, 9, 3153–3238. [CrossRef]

33. Flores, V.R.; Viozzi, G.; Casalins, L.; Loker, E.S.; Brant, S.V. A new schistosome (Digenea: Schistosomatidae) from the nasal tissue of South America black-necked swans, *Cygnus melancoryphus* (Anatidae) and the endemic pulmonate snail *Chilina gibbosa*. *Zootaxa* 2021, 4948, 404–418. [CrossRef] [PubMed]


35. Lorenti, E.; Brant, S.V.; Gilardoni, C.; Diaz, J.I.; Cremonete, F. Two new genera and species of avian schistosomes from Argentina with proposed recommendations and discussion of the polyphylectic genus *Gigantobilharzia* (Trematoda, Schistosomatidae). *Parasitology* 2022, 149, 1–20. [CrossRef]


40. Oyarzún-Ruiz, P.; Thomas, R.; Santodomingo, A.; Collado, G.; Muñoz, P.; Moreno, L. Morphological, behavioral, and molecular characterization of avian schistosomes (Digenea: Schistosomatidae) in the native snail *Chilina dombequana* (Chilinidae) from Southern Chile. *Pathogens* 2022, 11, 332. [CrossRef]

41. Devkota, R.; Brant, S.V.; Thapa, S.; Loker, E.S. Two avian schistosome cercariae from Nepal, including a *Macrobiharzia*-like species from *Indoplanorbis exustus*. *Parasitol. Int*. 2014, 63, 374–380. [CrossRef]


48. McCormack, J.E.; Harvey, M.G.; Faircloth, B.C.; Crawford, N.G.; Glenn, T.C.; Brumfield, R.T. A phylogeny of birds based on over 1500 loci collected by target enrichment and high-throughput sequencing. *PLoS ONE* 2013, 8, e54848. [CrossRef] [PubMed]


50. Faircloth, B.C.; McCormack, J.E.; Crawford, N.G.; Harvey, M.G.; Brumfield, R.T.; Glenn, T.C. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Syst. Biol.* 2012, 61, 717–726. [CrossRef]


54. Parada, A.; Hanson, J.; D’Eia, G. Ultraconserved Elements improve the resolution of difficult nodes within the rapid radiation of neotropical sigmodontine rodents (*Cricetidae: Sigmodontinae*). *Syst. Biol.* 2021, 70, 1090–1110. [CrossRef] [PubMed]


59. Blaier, B.B.; Lloyd, M.W.; Guillory, W.X.; Brady, S.G. Sequence capture and phylogenetic utility of genomic ultraconserved elements obtained from pinned insect specimens. *PLoS ONE* 2016, 11, e0161531. [CrossRef]


62. Barker, S.C.; Blair, D. Molecular phylogeny of *Schistosoma* species supports traditional groupings within the genus. *J. Parasitol.* 1996, 82, 292. [CrossRef]
63. Devkota, R.; Brant, S.V.; Loker, E.S. A genetically distinct *Schistosoma* from *Radix luteola* from Nepal related to *Schistosoma turkestanicum*: A phylogenetic study of schistosome and snail host. *Acta Trop.* 2016, 164, 45–53. [CrossRef]


89. Loker, E.S.; Dolginow, S.Z.; Pape, S.; Topper, C.D.; Alda, P.; Pointier, J.P.; Ebbs, E.T.; Sanchez, M.C.; Verocai, G.G.; De Jong, R.J.; et al. An outbreak of canine schistosomiasis in Utah: Acquisition of a new snail host (Galba truncatula) by Heterobilharzia americana, a pathogenic parasite on the move. One Health 2021, 13, 100280. [CrossRef]
100. Lu, D.B.; Wang, T.P.; Rudge, J.W.; Donnelly, C.A.; Fang, G.R.; Webster, J.P. Evolution in a multi-host parasite: Chronobiological circadian rhythm and population genetics of Schistosoma japonicum cercariae indicates contrasting definitive host reservoirs by habitat. Int. J. Parasitol. 2009, 39, 1581–1588. [CrossRef]
101. Walker, J.C. Austrobielharzia terryi gen. nov: A schistosome dominant in interspecific interactions in the molluscan host. Int. J. Parasitol. 1979, 9, 137–140. [CrossRef]
107. Walker, J.C. Austrobielharzia terryi gen. nov: A schistosome dominant in interspecific interactions in the molluscan host. Int. J. Parasitol. 1979, 9, 137–140. [CrossRef]
111. Walker, J.C. Austrobielharzia terryi gen. nov: A schistosome dominant in interspecific interactions in the molluscan host. Int. J. Parasitol. 1979, 9, 137–140. [CrossRef] [PubMed]


130. Harris, R.S. Improved Pairwise Alignment of Genomic DNA; The Pennsylvania State University: State College, PA, USA, 2007.


