Molecular Phylogenetics and Evolution 126 (2018) 371–381

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Resolving evolutionary changes in parasite life cycle complexity: Molecular phylogeny of the trematode genus *Alloglossidium* indicates more than one origin of precociousness

Emily L. Kasl[a,b,*], William F. Font[c], Charles D. Criscione[b]

[a] Department of Biology, University of North Alabama, Florence, AL, USA
[b] Department of Biology, Texas A&M University, College Station, TX, USA
[c] Department of Biological Sciences, Southeastern Louisiana University, Hammond, LA, USA

ARTICLE INFO

Keywords:
Complex life cycle
Precocious development
Progenetic
Ancestral trait reconstruction
Character evolution

ABSTRACT

The evolutionary causes and consequences of changes in complex life cycles are of central importance in parasitology. However, data remain limited because in part, knowledge on phylogenetic relationships among species that differ in life cycle patterns remains scarce. We present a molecular phylogeny of the trematode genus *Alloglossidium*, which contains several species that display precocious (a.k.a., progenetic) life cycles (i.e., maturation in what is typically regarded as an intermediate host). The molecular phylogeny contrasts with previous morphological and life-history based phylogenetic hypotheses. In particular, a precocious life cycle wherein leeches are used as final hosts evolved early in the history of the genus. Among the remaining species, which are a separate clade, a three-host life cycle using ictalurid catfishes is ancestral. Furthermore, there are at least two additional independent evolutionary events that lead to a precocious life cycle where a catfish host is lost and a crustacean is used as a final host. We conclude with a discussion on how existing hypotheses on the evolution of precociousness, and parasite life cycle complexity in general, may or may not relate to the patterns observed in genus *Alloglossidium*.

1. Introduction

The addition or removal of a host from a parasite's life cycle is not a trivial evolutionary event; life cycle transitions can have major consequences for parasite transmission, behavior, physiology, development, and mating systems (e.g., Kasl et al., 2015). Moreover, complex life cycles involving trophic transmission are regarded as a major adaptive peak for parasitism (Lafferty and Kuris, 2002; Poulin and Randhawa, 2015). Consequently, the evolution of changes in life cycle complexity is a longstanding topic in parasitology (Poulin, 2007) and there is now renewed emphasis on elucidating the factors driving the selection for, and maintenance of, such complexity (Choisy et al., 2003; Parker et al., 2003; Hammerschmidt et al., 2009; Benesh et al., 2013, Parker et al., 2015a,b; Auld and Tinsley, 2015). However, studying the causes and consequences of complex life changes hinges on knowledge of the evolutionary order of life cycle transitions, which in itself, requires knowledge of phylogenetic relationships among species that differ in life cycle patterns. Unfortunately, such relationships are still not known among many parasitic taxa, especially within groups that display life cycle variation.

Digenean trematodes are notable for having some of the most complex life cycles, typically incorporating both free-living and parasitic developmental stages and almost always including both asexual and sexual reproduction (Cribb et al., 2003; Olson et al., 2003). Across the Digenea, a 3-host life cycle is the most commonly found pattern (Cribb et al., 2003, Olson et al., 2003, see Fig. 1A as an example). In contrast to this common 3-host pattern, however, some species exhibit what has been termed a "progenetic" life cycle (Cribb et al., 2003; Lefebvre and Poulin, 2005a). Hereafter, we refer to species exhibiting such patterns as "precocious" rather than progenetic to reflect the parasite's early onset of adult development (attaining sexual maturity) within what is typically deemed an intermediate host. Some of these precocious species have obligate 2-host patterns where it is presumed that the 3rd host, i.e., the previous final host, was lost, whereas other species are facultative in having a 2- or 3-host pattern (Lefebvre and Poulin, 2005b). It is these precocious life cycle variations that are the focus of our study. In particular, we are interested in the evolution of precocious life cycles within the digenean genus *Alloglossidium*.

The genus *Alloglossidium* Simer, 1929 is comprised of 18 nominal species (Table 1). It is a uniquely appropriate system with which to

* Corresponding author at: Department of Biology, University of North Alabama, Florence, AL, USA.
E-mail address: ekasl@una.edu (E.L. Kasl).

https://doi.org/10.1016/j.ympev.2018.04.027
Received 10 January 2018; Received in revised form 17 March 2018; Accepted 17 April 2018
Available online 27 April 2018
1055-7903/

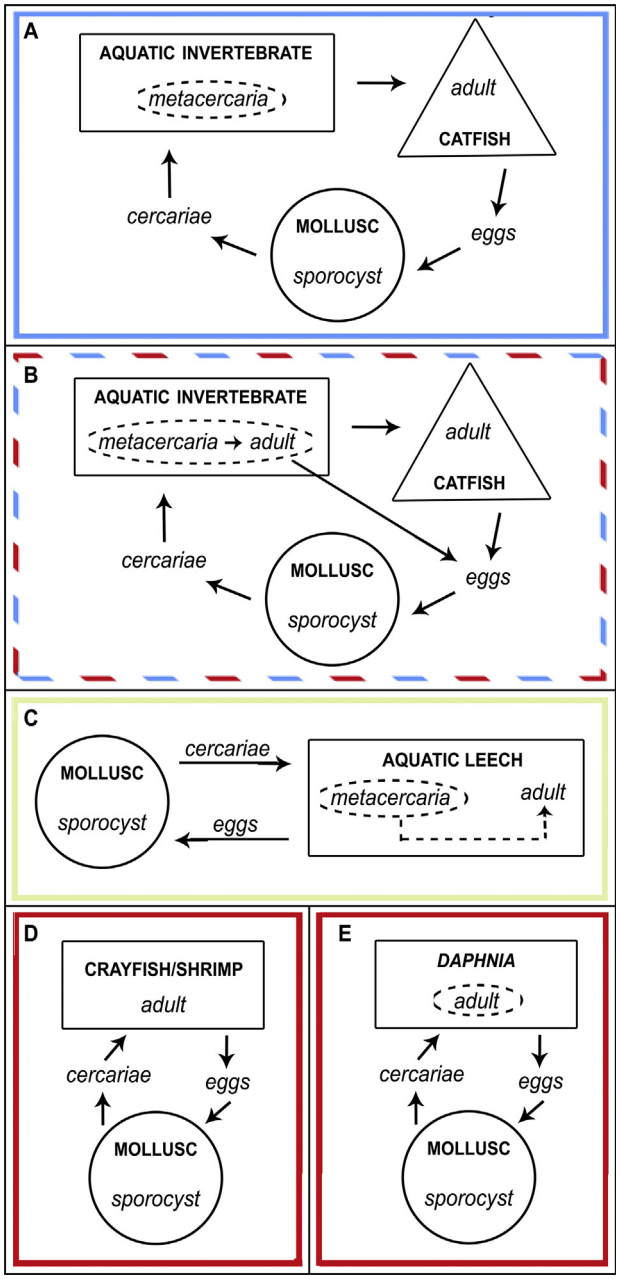

Fig. 1. Diagram of *Alloglossidium* life cycle patterns (see Table 1 for species associations). Hosts are denoted using bold uppercase type. Parasite life history stages are denoted using italicized lowercase type. A. In the obligate 3-host life cycle, the parasite is ingested by a molluscan 1st intermediate host, undergoes asexual reproduction leading to the release of larval cercariae. Cercariae subsequently penetrate an invertebrate 2nd intermediate host, developing into an encysted (indicated by dotted lines) metacercarial stage. Upon ingestion of the 2nd intermediate host by the final host (an ictalurid catfish), the parasite excysts, migrates to the intestines, and becomes sexually mature, thereby completing the life cycle. B. In the facultative precocious life cycle (found in *A. progeneticum*), the parasite becomes sexually mature while still encysted in a crayfish 2nd intermediate host (i.e. a fish host is not needed). However, encysted adults can still be trophically transmitted to an ictalurid host and is therefore considered a facultative 2- or 3-host life cycle. C–E show life cycles for species with obligate precocious life cycles. C. Species utilizing leeches as the final host maintain all life history stages (i.e., the encysted metacercarial stage), first encysting, then excysting and migrating to the gut within the same host individual to complete development. D. Species reported from crayfish and shrimp have lost the encysted metacercarial stage, instead persisting freely in the antennal glands of their respective hosts. E. *Alloglossidium anomophagis* infects the body cavity of *Daphnia* and becomes an adult while remaining encysted.

address questions regarding evolutionary changes in life cycle complexity because it contains species with 3-host life cycles and with precocious life cycles. Moreover, roughly 15% of the known precociously developing digenean species occur within this genus (Lefebvre and Poulin, 2005a). Five species have life cycles reflecting the typical 3-host digenean pattern where an ictalurid catfish is the final host (Table 1, Fig. 1A, but see discussion for caveats on undiscovered life history). One species, *Alloglossidium progeneticum*, is capable of facultatively incorporating either 2- or 3-hosts in its life cycle because it can precociously develop while still encysted in an "intermediate" crayfish host (Table 1, Fig. 1B). The remaining 12 species exhibit one of three obligate precocious life patterns classified by the type of final host (i.e., leech or freshwater crustacean) and the presence or absence of an encysted metacercarial stage (Table 1, see Fig. 1C–E for elaboration). Because of this life cycle variation found among species, the genus *Alloglossidium* was recognized early on as a potential model system with which to study developmental constraints and integrative biological components (e.g., the origin, stasis, and diversification) of complex life cycles (Brooks, 2003). While past studies have addressed the evolution of life cycle changes in this genus (discussed below), interpretations are limited in part due to the finding of cryptic species in the genus (e.g., Tkach and Mills, 2011) and exclusion of *A. anomophagis*.

The first two studies to address the phylogeny of species in the genus *Alloglossidium* defined interspecific relationships through logical deductions. Font (1980) suggested a 3-host life cycle was the ancestral condition for the genus due to the ubiquity of this pattern among digeneans. He deduced there was a subsequent split leading to the speciation of *Alloglossidium* maturing in crustacean hosts (through a progressive transition via the facultative precocious pattern) and to those species maturing in leeches. In contrast, Riggs and Ulmer (1983) postulated that the obligate 2-host pattern in leeches was the ancestral condition. They argued that leeches predated catfishes in the fossil record and that the similarity between the gut lumen of leeches and the intestines of catfishes enabled a host-switching event from leeches to catfishes. Species associated with crustacean final hosts were not addressed by Riggs and Ulmer (1983).

Previous method-based phylogenetic hypotheses were based on a combination of morphological and life history characters where the most recent analysis proposed a single transition from a 3- to 2-host crustacean life cycle pattern, followed by a host switching event from crustaceans to leeches (Fig. 2; Carney and Brooks, 1991; Smythe and Font, 2001; but see Brooks, 2003 for methodological corrections). However, incorporating the same life history traits to both reconstruct the phylogeny and assess their ancestral changes in life cycle pattern has the potential to confound the analysis. Ideally, when testing hypotheses related to trait evolution, a phylogeny constructed from data independent of the traits themselves would be preferred.

The aim of our study was to construct the first molecular-based phylogeny with which to study the evolutionary relationships within the genus *Alloglossidium*. This phylogeny subsequently provides an independent framework to assess the character evolution of life cycle pattern transitions. In particular, we compare the results of the molecular-based phylogeny to the existing deductive or morphology-based hypotheses on complex life cycle evolution in the genus *Alloglossidium*. In doing so, we elucidate those parameters essential to study the evolution of life cycle complexity, i.e., when a particular life cycle pattern developed and the number of transitions among possible patterns. Lastly, we discuss how existing hypotheses on the evolution of precociousness and parasite complex life cycles in general may or may not relate to the patterns observed in genus *Alloglossidium*.

2. Materials and methods

2.1. Taxon sampling and outgroup selection

Sampling of *Alloglossidium* species was conducted as part of a large-

Table 1
Sample ID, species name (based on morphological identification), definitive host name, collection locality, latitude (N), longitude (W), and GenBank accession numbers for *Alloglossidium* species used in this study. The superscripts on the species names correspond to the life cycle pattern insets on Fig. 1.

ID	Species	Definitive Host	Collection locality	Latitude (N)	Longitude (W)	GenBank Accession No.	
						rDNA	ND1
AN01	*A. anomophagis*[E]	*Daphnia obtusa*[*]	Houston Coastal Center, TX[^]	29°43.34394′	95°31.1331′	MH041376	MH041326
AN02	*A. anomophagis*[E]	*Daphnia obtusa*[*]	Houston Coastal Center, TX[^]	29°43.34394′	95°31.1331′	MH041377	MH041327
CA01	*A. cardicolum*[D]	*Procambarus acutus*[*]	Walnut Rd, MS	32°40.354′	89°43.907′	MH041378	MH041328
CA02	*A. cardicolum*[D]	*Procambarus acutus*[*]	Rosedale, LA[^]	30°27.727′	91°25.023′	MH041379	MH041329
DO01	*A. dolandi*[D]	*Procambarus acutissimus*	Jackson Br. @ Williams Store Rd, GA[^]	32°43.095′	81°30.133′	MH041380	MH041330
DO02	*A. dolandi*[D]	*Procambarus epicyrtus*[*]	Jackson Br. @ Newington Hwy, GA[^]	32°42.930′	81°29.362′	MH041381	MH041331
GR01	*A. greeri*[D]	*Cambarellus schudfeldti*[*]	Montrose, LA[^]	31°34.55′	92°59.85′	MH041387	MH041337
GR02	*A. greeri*[D]	*Cambarellus schudfeldti*[*]	Vatcherie, LA	29°54.729′	90°43.713′	MH041388	MH041338
GR03	*A. greeri*[D]	*Cambarellus schudfeldti*[*]	Rockwood, IL	37°43.546692′	89°27.809653′	MH041389	MH041339
RE01	*A. renale*[D]	*Palaemonetes kadiakensis*[*]	Olustee Creek, AL	31°56.65′	86°7.133333′	MH041385	MH041335
RE02	*A. renale*[D]	*Palaemonetes kadiakensis*[*]	Choctaw Rd, LA	29°51.478′	90°45.281′	MH041386	MH041336
PR01	*A. progeneticum*[A]	*Ameiuras natalis*	Crooked Creek, AR	36°14.116′	92°42.763′	MH041383	MH041333
PR02	*A. progeneticum*[B]	*Procambarus spiculifer*[*]	Calls Creek, GA[^]	33°53.310′	83°22.918′	MH041382	MH041332
PR03	*A. progeneticum*[A]	*Ameiurus natalis*	Gus Engeling WMA, TX	31°55.687′	95°53.279′	MH041384	MH041334
CO01	*A. corti*[A]	*Ameiurus melas*	Whisky Bay, LA	30°23.479′	91°20.826′	MH041392	MH041342
CO02	*A. corti*[A]	*Ameiuras natalis*	Greenwood, LA	32°28.025′	93°58.905′	MH041393	MH041343
CO03	*A. corti*[A]	*Ictalurus punctatus*	Brazos River, Bryan, Texas	30°37.705167′	96°32.659333′	MH041394	MH041344
FO01	*A. fonti*[A]	*Ameiurus melas*[*]	Trout Brook, VT	43°53.36′	72°41.426667′	MH041395	MH041345
FO02	*A. fonti*[A]	*Ameiurus melas*[*]	Sixmile Creek, WI	43°8.430613′	89°25.553496′	MH041396	MH041346
FO03	*A. fonti*[A]	*Ameiurus melas*	Trinity River, TX	30°16.578′	94°47.763′	MH041398	MH041348
FO04	*A. fonti*[A]	*Ameiurus melas*	Greenwood, LA	32°28.025′	93°58.905′	MH041397	MH041347
MS01	*A. kenti* MS[A]	*Ictalurus punctatus*[*]	Money Bayou, Money, MS[^]	33°39.493′	90°12.528′	MH041405	MH041355
MS02	*A. kenti* MS[A]	*Ictalurus furcatus*	Tallahatchie River @ Money, MS (N)	33°43.120′	90°13.174′	MH041404	MH041354
MS03	*A. kenti* MS[A]	*Ictalurus furcatus*	Tallahatchie River @ Money, MS (S)	33°35.6874′	90°11.6990′	MH041406	MH041356
TK01	*A. kenti* TK[A]	*Ictalurus punctatus*[*]	Little Brazos, TX	30°38.485′	96°31.222333′	MH041399	MH041349
TK02	*A. kenti* TK[A]	*Ictalurus punctatus*[*]	Little Brazos, TX	30°38.485′	96°31.222333′	MH041400	MH041350
TK03	*A. kenti* TK[A]	*Ictalurus punctatus*[*]	Crooked Creek, AR	36°14.116′	92°42.763′	MH041401	MH041351
NY01	*Alloglossidium n.* sp. *1*[A]	*Ictalurus punctatus*[*]	Oneida Lake, NY	43°11.005′	75°59.636′	MH041402	MH041352
NY02	*Alloglossidium n.* sp. *1*[A]	*Ictalurus punctatus*[*]	Oneida Lake, NY	43°11.005′	75°59.636′	MH041403	MH041353
FL01	*A. floridense*[A]	*Noturus gyrinus*	Spring Run, FL[^]	29°°51.533833′	82°44.094′	MH041390	MH041340
FL02	*A. floridense*[A]	*Noturus leptacanthus*[*]	Spring Run, FL[^]	29°51.533833′	82°44.094′	MH041391	MH041341
GE01	*A. geminum*[A]	*Ameiurus nebulosus*[*]	Oneida Lake, NY[^]	43°11.005′	75°59.636′	MH041409	MH041359
GE02	*A. geminum*[A]	*Ameiurus nebulosus*[*]	Oneida Lake, NY[^]	43°11.005′	75°59.636′	MH041410	MH041360
CD01	*Alloglossidium n.* sp. *2*[A]	*Ictalurus punctatus*[*]	St. Lawrence River, Quebec, Canada[^]	45°18.96′	73°52.74′	MH041407	MH041357
CD02	*Alloglossidium n.* sp. *2*[A]	*Ictalurus punctatus*[*]	St. Lawrence River, Quebec, Canada[^]	45°18.96′	73°52.74′	MH041408	MH041358
RI01	*A. richardsoni*[C]	*Haemopis plumbea*[*]	Zippel Bay, Lake of the Woods, MN	48°52.369′	94°51.549′	MH041411	MH041361
RI02	*A. richardsoni*[C]	*Haemopis marmorata*	Falcon Lake, Canada	49°40.701′	95°18.867′	MH041412	MH041362
MA01	*A. macrobdellensis*[C]	*Macrobdella decora*	Boot Bog, WI	45°5.582′	91°20.178′	MH041413	MH041363
MA02	*A. macrobdellensis*[C]	*Macrobdella ditetra*[*]	Whisky Bay, LA[^]	30°23.479′	91°20.826′	MH041414	MH041364
HA01	*A. hamrumi*[C]	*Macrobdella decora*[*]	Storrs, U. Connecticut, CT	41°49.0495′	72°15.54283′	MH041415	MH041365
HA02	*A. hamrumi*[C]	*Macrobdella decora*[*]	Storrs, U. Connecticut, CT	41°49.0495′	72°15.54283′	MH041416	MH041366
HI01	*A. hirudicola*[C]	*Haemopis plumbea*	Zippel Bay, Lake of the Woods, MN	48°52.369′	94°51.549′	MH041417	MH041367
HI02	*A. hirudicola*[C]	*Macrobdella decora*	Zippel Bay, Lake of the Woods, MN	48°52.369′	94°51.549′	MH041418	MH041368
TU01	*A. turnbulli*[C]	*Haemopis grandis*[*]	Zippel Bay, Lake of the Woods, MN[^]	48°52.369′	94°51.549′	MH041423	MH041373
TU02	*A. turnbulli*[C]	*Haemopis grandis*[*]	Falcon Lake, Canada	49°40.701′	95°18.867′	MH041424	MH041374
SC01	*A. schmidti*[C]	*Haemopis grandis*[*]	Falcon Lake, Canada[^]	49°40.701′	95°18.867′	MH041419	MH041369
SC02	*A. schmidti*[C]	*Haemopis grandis*[*]	Lake Kabetogama, MN	48°26.839′	93°2.951′	MH041420	MH041370
GU01	*Alloglossidium n.* sp. *3*[C]	*Macrobdella ditetra*[*]	Gus Engeling WMA, TX[^]	31°55.687′	95°53.279′	MH041421	MH041371
GU02	*Alloglossidium n.* sp. *3*[C]	*Macrobdella ditetra*[*]	Gus Engeling WMA, TX[^]	31°55.687′	95°53.279′	MH041422	MH041372

[*] Indicates type host.
[^] Indicates type locality.

scale biodiversity survey of the eastern two-thirds of the United States and southern Canada. All nominal species, with the exception of *Alloglossidium demshini* (described post-survey; Tkach et al., 2013) were collected (Table 1), killed with hot water or heat-fixed underneath a coverslip, and stored in 70% or 95% ethanol for further molecular work. All specimens were identified using the most complete morphological descriptions available in the literature and, when possible, compared to molecular sequences deposited on GenBank (Tkach and Mills, 2011: JF440783.1, *A. corti*; JF440765.1and JF44767.1, *A. fonti*; JF440771.1, *A. geminum*; JF440808.1, *A. kenti*; Kasl et al., 2014: KC812276.1, *A. floridense*) for further delimitation. In order to definitively connect molecular sequence data to morphological descriptions, efforts were made to obtain specimens from the type hosts and/or type localities. Though, it should be noted that this was not always possible due to extirpation of hosts at the type locality (i.e., the loss of the madtom hosts used by *A. corti* in Lake Mendota, WI; Lyons, 1989), habitat destruction of the type locality (i.e., *A. renale*; personal observation), or even a lack of type locality (i.e., *A. hirudicola* was originally described from leeches purchased from a biological supply company). All individual specimens were collected from different host individuals. When possible, specimens were also chosen to represent known geographical ranges (Table 1). Additionally, new lineages (likely representing new species) found during the surveys were included. Three outgroup taxa were used. *Brachycladium goliath* (Briscoe et al., 2016: KR703279) was chosen due to the completeness of sequence information on Genbank, with coverage for both the mitochondrial and ribosomal regions of interest. *Paramacroderoides echinus*, from longnose gar, *Lepisosteus osseus*, in the Little Brazos River, TX (TX; 30°38.485N 96°31.222333′W; rDNA: MH041375; mtDNA: MH041325), was chosen due to the accessibility of obtaining specimens for molecular use and

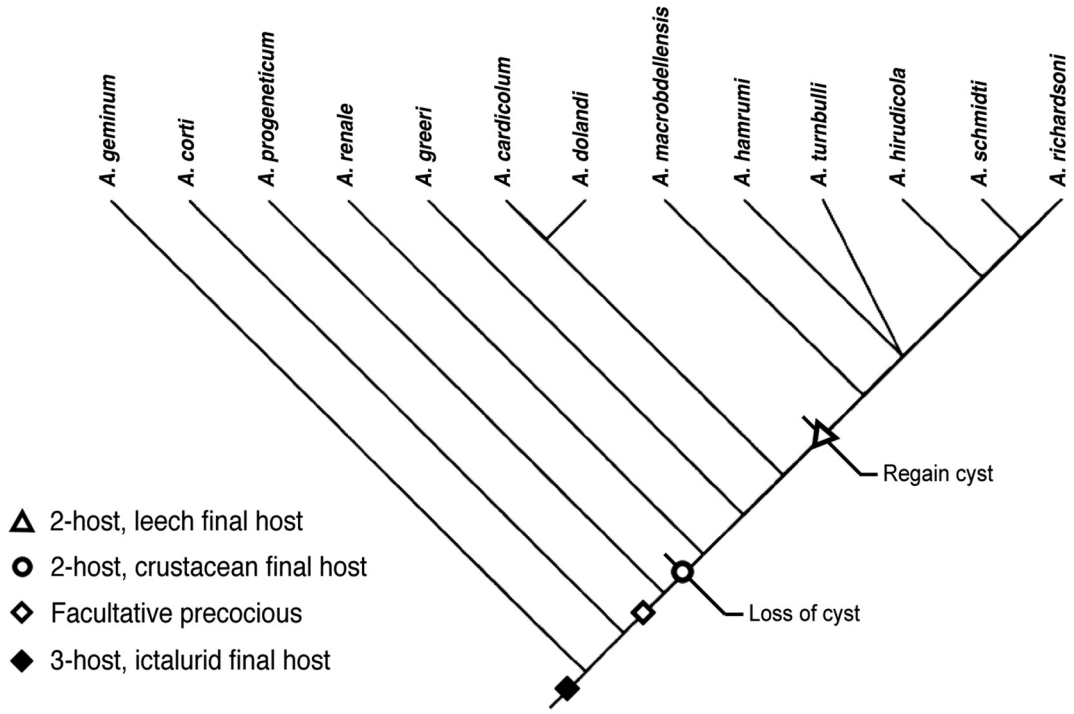

Fig. 2. The most recent morphological and life-history based phylogenetic tree of the genus *Alloglossidium.* Analysis was conducted by Smythe and Font (2001) (but, see also Brooks (2003)). The figure is slightly altered from Brooks (2003) and adapted with permission of the Journal of Parasitology.

the historical placements of both *Alloglossidium* and *Paramacroderoides* within family Macroderoididae McMullen, 1937 (Bray et al., 2008). The final outgroup, *Magnivitellinum simplex* (Hernández-Mena et al., 2016: KU535683), was only obtainable as a partial sequence (28s rDNA). Nevertheless, *M. simplex* was chosen due to recent molecular support indicating that *Magnivitellinum* is the sister taxon to *Alloglossidium* (Hernández-Mena et al., 2016).

2.2. Extraction, DNA amplification, and sequencing

DNA was extracted by placing individual worms in 200 µl of 5% chelex containing 0.2 mg/ml of proteinase K, incubated for 2 hr at 56 °C and boiled at 100 °C for 8 min. Two unlinked markers were used for identification and analysis of phylogenetic relationships: (1) a region of the ribosomal DNA (rDNA) complex that ranged from the 3′ end of the 18s (309 bp) through the first internal transcribed spacer (ITS1), 5.8s, ITS2 and about 1360 bp of the 5′ end of the 28s gene (domains D1-D3), and (2) a 676 base pair region of the mitochondrial NADH-dehydrogenase subunit 1 gene (ND1). Polymerase chain reaction (PCR) amplifications were performed using 25 µl reactions. The 18s-ITS1 portion of the nuclear ribosomal sequence (s18 forward primer, 5.8s1 reverse primer; Table S1) was obtained using a reaction consisting of 3 µl of template extract, 16.25 µl water, 2.5 µl 10x buffer, 1.5 µl $MgCl_2$ [25 mM], 0.5 µl dNTP [10 mM/each], 0.5 µl of each primer [20 µM], and 0.25 µl of Taq polymerase (Omega Bio-Tek Inc., Norcross, GA) [5 units/µl], and a thermocycling profile of 95 °C for 3 min, once; 94 °C for 45 sec, 60 °C for 30 sec, 72 °C for 60 sec, 35 times; 72 °C for 7 min, once. The ITS1-5.8s-ITS2-28s portion of the nuclear ribosomal sequence (CC41 forward primer, 1500R reverse primer; Table S1; thermocycler profile described in Olson et al., 2003) and the ND1 region (MB352 forward primer, CC28 reverse primer; MB411 forward primer, CC29 reverse primer; Table S1; thermocycler profile described in Criscione and Blouin, 2004) were obtained using the same reaction recipe with the exception that 15.25 µl water and 2.5 µl $MgCl_2$ [25 mM] were used. If the above protocol was not successful in obtaining ND1 sequences (particularly in the case of the leech-associated species and *A. floridense* and *A. dolandi*) a combination of gel extractions, using the Ultra Clean Gel Spin DNA extraction kit (MO BIO Laboratories, Inc., Solana Beach, California), and alternate sequencing primers were used (Table S1). PCR products were purified using the Ultra Clean PCR clean-up Kit (MO BIO Laboratories, Inc., Solana Beach, California) and sent to the DNA Analysis Facility on Science Hill at Yale University (New Haven, Connecticut) for sequencing.

2.3. Phylogenetic analyses

At least 2 individuals per species (as defined by morphological identifications) were used to assess the phylogenetic relationships within the genus *Alloglossidium*. These individual specimens were obtained from unique hosts and, when possible, were chosen to reflect some of the known host diversity (different host species) and/or geographic variation (different locations) associated with each species. Contiguous sequences from individuals were assembled using Sequencher™ (GeneCodes Corp., ver. 4.1.4), and submitted to GenBank (see Table 1 for accession numbers). Both nuclear and mitochondrial sequences were aligned using Clustal W, and manually edited, within the BioEdit program, version 7.1.8 (Hall, 1999). The alignment data are also included as a supplemental file (Supplementary Data 1).

To infer phylogenetic relationships, identical sequences were condensed into a single taxonomic unit and used in the analysis of 3 data sets: rDNA-only (18s-ITS1-5.8s-ITS2-28s), ND1-only, and a concatenated rDNA-ND1 set. The ITS1 and ITS2 regions were not alignable between outgroup and ingroup taxa and were thus treated as missing data for the outgroup taxa. Within the ingroup taxa, ITS regions could be aligned unambiguously, baring a few minor exceptions. In particular, the ITS regions contained a number of short indels (1–14 bp in length) that, though unalignable in themselves, clearly showed evidence of phylogenetic signal (Table S2). Consequently, unambiquous alignable nucleotides flanking the short indel regions were used for demarcation. Subsequently, indel “blocks” were characterized as multicharacter states if there were different haplotypes within the block itself (Table S2). Because maximum likelihood methods cannot simultaneously use both multicharacter traits (i.e. indel regions) along with nucleotide data, we applied a Bayesian framework to reconstruct

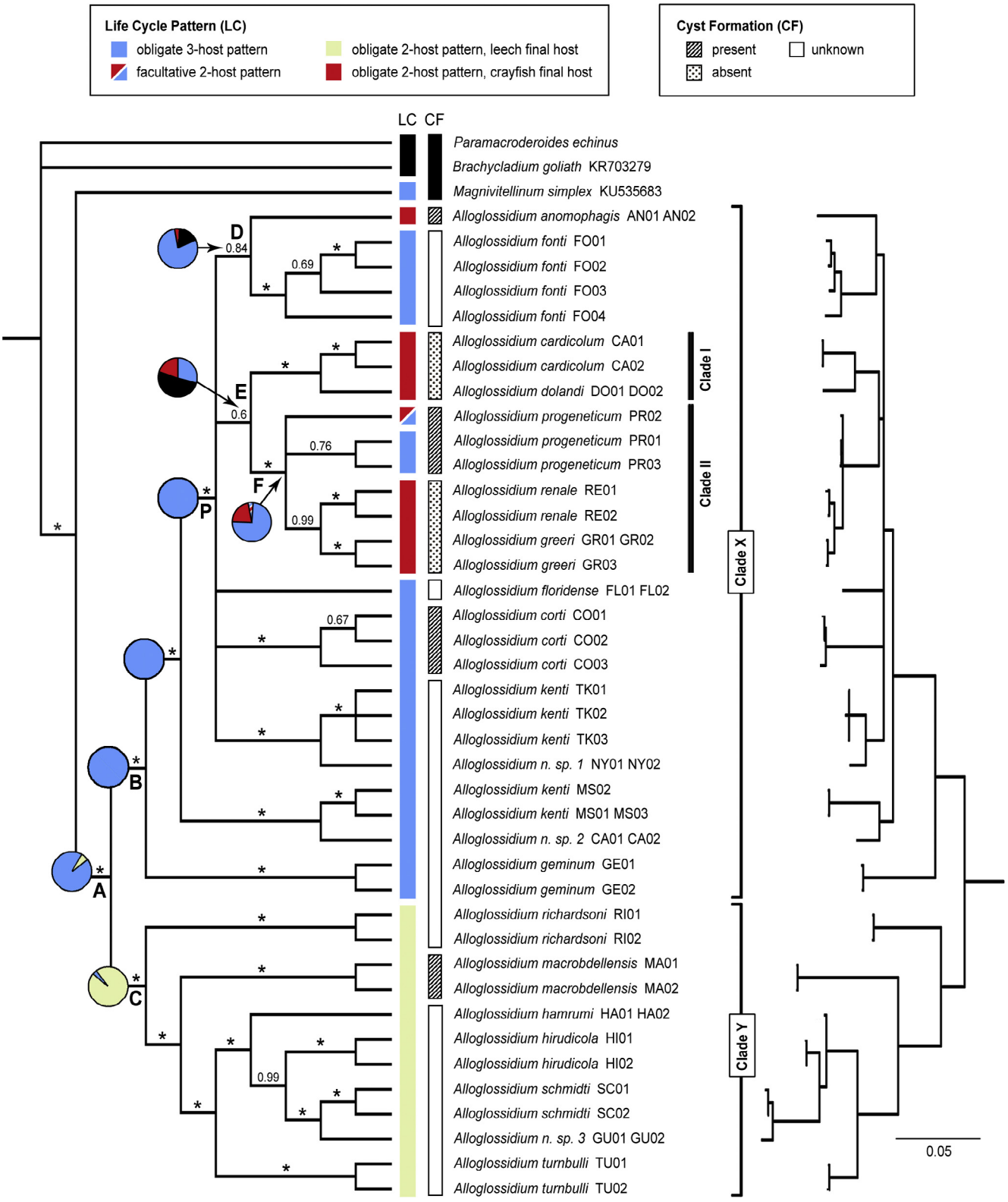

Fig. 3. Bayesian inference 50% majority rule consensus cladogram (left) and phylogram (right) obtained from the concatenated (ribosomal and mitochondrial) data set as analyzed in MrBayes. Bayesian posterior probabilities are shown at each node, with asterisks indicating probabilities of 1.0. Pie diagrams show the results of the ancestral trait reconstructions for nodes of interest (discussed further in text). A black proportion of a pie diagram shows the proportion of sampled trees where a node is absent. Colored proportions indicate the average frequency of the ancestral state inferred using the maximum likelihood method and the average frequency across trees calculation (see Methods).

the *Alloglossidium* phylogeny. Data blocks were set up according to rDNA gene (18s, ITS1, 5.8s, ITS2, 28s), rDNA indels (included in a single block), or ND1 codon position (1st, 2nd, 3rd). To demarcate the rDNA regions, we compared our sequences to rDNA annotation of two digenean species of the genus *Diplostomum* (Brabec et al., 2015). PartitionFinder v.1.1.0 (Lanfear et al., 2012) was used to determine the best partition scheme, according to the Bayesian Information Criteria (BIC), and model of evolution for each partition. The following parameters were used in ParitionFinder searches: branch lengths = unlinked; models = mrbayes; model selection = BIC; data blocks defined by gene (rDNA) and codon position (ND1 mtDNA); and search = all. The partitioning schemes of the individual rDNA and ND1 datasets were subsequently applied to the applicable regions of the concatenated dataset.

Bayesian phylogenetic reconstructions were conducted with the MrBayes v3.2 (Huelsenbeck and Ronquist, 2001) package. So that each partition was allocated its own set of parameters, statefreq (state frequency), revmat (transition matrices), pinvar (proportion of invariant sites), and shape (gamma shapes) were coded as unlinked and the ratepr parameter was set to variable to allow for the evolution of partitions under different rates. PartionFinder identified 4 partitions for the rDNA dataset: (1) 18s and 5.8s (K2P; Kimura, 1980), (2) ITS1 (SYM + Γ; the GTR model with stationary state frequencies fixed to be equal), (3) ITS2 (K2P), and (4) 28s (GTR + Γ). The multistate/binary traits were included as a separate partition and estimated using Γ rate only. For the mtDNA dataset, PartitionFinder identified 2 partitions: (1) 1st and 2nd codons combined (HKY + Γ; Hasegawa et al., 1985) and (2) 3rd codon alone (GTR + I + Γ). Two parallel runs each with 4 chains were run over 4 million generations, sampled every 2000 steps. The run length was deemed appropriate based on the convergence of the parallel runs (standard deviation of split frequencies < 0.01). Also, for all model parameters, the potential scale reduction factors approximated 1, thus indicating a good sample from the posterior probability distribution. After reaching stability, the first 25% of trees were discarded as burn-in and calculated a 50% majority-rule consensus tree.

2.4. Ancestral reconstruction of the evolution of life cycle complexity

Character states associated with life cycle complexity and coded for use in ancestral state reconstructions were (0) obligate 2-host crustacean, (1) facultative precocious, (2) obligate 3-host, and (3) obligate 2-host leech. The life history states of the outgroup taxa *B. goliath* and *P. echinus* were coded as missing data as there is still ambiguity for digenean relationships, due to sparse taxa sampling, within the superfamily Plagiorchioidea (to which *Alloglossidium* belongs). Hence, we did not want the character states of these two taxa to influence results. However, the findings of two recent studies based on 28s rDNA indicate that the genus *Magnivitellinum* is sister to the genus *Alloglossidium* (Hernández-Mena et al., 2016; Sokolov and Shchenkov, 2017). *Magnivitellinum* spp. infect fish, including catfishes (Order Siluriformes) as the definitive host. Hence, we coded *M. simplex* as having an obligate 3-host state. We return to this in the discussion. Ancestral state reconstruction was carried out on the 50% majority consensus tree using the Trace Character Over Trees (TCOT) command in Mesquite v3.04 (Maddison and Maddison, 2010). The TCOT command was chosen to take into account uncertainty (i.e. node presence or absence summarized across a series of trees) when determining the most likely ancestral condition, rather than limiting the reconstruction to the consensus tree alone (Trace command). Ancestral character evolution was assessed by using the maximum likelihood method and the average frequency across trees calculation (sampled from the trprobs output [2143 trees]), which provides the average probability of a state across all of the trees possessing that node. Transitions between each character state were coded as equally likely. A probability ≥ 0.60 for a character was considered to be the most probable state at that node (Feutry et al., 2013; Chauhan and Pandey, 2014; Carvalho-Sobrinho et al., 2016).

As will be seen in the results (see below), some of the more terminal nodes did not have strong support. Hence, there is less certainty in assessing the number of state transitions, especially the number of times an obligate 3-host state goes to an obligate 2-host crustacean state. As an additional means to assess the number of state changes, we used the Summarize State Changes Over Trees command in Mesquite. This method calculates the average number of times a transition (from any one state to another) occurs across a set of trees (sampled from the trprobs output [2143 trees]). We did this with both the parsimony and likelihood settings. The likelihood setting requires a threshold cutoff (we used the default of 2) based on the likelihood scores of state changes. We note as the threshold is decreased, the results approach that found under the parsimony setting. Hence, we view the results under the likelihood setting as more of a lower bound estimate in terms of the average number of times a particular transition occurs.

2.5. Phylogenetic hypothesis testing

As noted above, some of the more terminal nodes did not have strong support (in particular, Node E of Fig. 3; see Results). As Node E was critical for determining a minimum of 2 transitions from an obligate 3-host state to an obligate 2-host crustacean state, we performed a topology test of this node (following Brown and Thompson, 2016). A single-origin hypothesis was defined as the positive constraint (H1) and defined such that all precocious species formed a forced monophyly. In contrast, the negative constraint (H0) was defined based on relationships supported in the concatenated phylogeny, i.e., species associated with Clades I and II form a clade (Node E, Fig. 3) that excludes *A. anomophagis* from being sister to the clade defined by Node E or any of the taxa within Node E. Thus, support for the negative constraint would suggest at minimum 2 origins of precocious development. Due to the topological differences associated with the relationships of precociously developing species in the rDNA and mtDNA datasets, we tested the same positive and negative constraints in both datasets. To test whether the positive constraint was preferred over the negative constraint, we estimated the marginal likelihoods for each model using the stepping-stone sampling approach (Xie et al., 2011) implemented in a modified version of MrBayes (mb_topoFix; Brown and Thomson, 2016), which prevented topology moves from being incorrectly shut off during different runs. For the stepping stone approach, we set the number of generations at 2,000,000, the alpha-shape parameter to 0.4, the sampling frequency to 1000, nsteps to 50, and the default for the burn-in (resulting in 39,000 generations per step). Convergence was determined based on the standard deviation of split frequencies. To determine support for a particular phylogenetic hypothesis we used the difference between the marginal likelihoods to calculate the 2ln Bayes Factor (2ln B10). Strong significance was determined according to the recommendations of Kass and Rafferty (1995) as values of 2ln B10 > 10.

2.6. Analysis of the “progeneticum” clade

Initial findings of a strongly supported clade containing *A. progeneticum*, *A. renale*, and *A. greeri* (see Results), coupled with previous research showing high intraspecific variation in life cycle pattern of *A. progeneticum* (Kasl et al., 2015), led us to further analyze the evolution of precociousness within the “progeneticum” clade (Clade II, Fig. 3). The ND1 mtDNA sequences of all known *A. progeneticum* haplotypes, which were collected across the southeastern U.S.A. (GenBank accession nos. KT455707-KT45582, Kasl et al., 2015; KU728083-KU728090, McAllister et al., 2016), were analyzed alongside specimens of *A. progeneticum, A. renale,* and *A. greeri* from this study. *Alloglossidium kenti* (from Mississippi) and *A. geminum* were chosen as outgroup taxa due to their position as either the closest supported ancestor to Clade II (*A. kenti* from Mississippi) or as the most basal taxa to all *Alloglossidium* species in fish and crustacean hosts (*A. geminum*). A 676 bp dataset was

created and the Bayesian reconstructions were analyzed using the same parameters specified for the ND1-only dataset above.

3. Results

3.1. Molecular variation

Among *Alloglossidium* sequences only, the alignment length of the rDNA was 2602 nucleotides. As noted above there were 12 short indel regions (Table S2) that were treated as multicharacter or binary states. Excluding these sites plus a single uninformative insertion site, there were 2555 nucleotides in the rDNA alignment. We briefly note here that phylogenetic analyses of the rDNA alone with or without the inclusion of the 12 indel regions coded as multicharacter or binary states yielded very similar topologies. The only differences were in nodal support values for a few tip relationships that altered the consensus trees of the rDNA data alone (compare Figs. S1 and S2). Hence, from here on we will only refer to the results of the data where the12 indel regions were included in the analyses. Twenty-four haplotypes were found from 49 individuals, corresponding to 17 of the nominal species included in this study and 3 new species yet to be described: *Alloglossidium* n. sp. 1, *Alloglossidium* n. sp. 2, and *Alloglossidium* n. sp. 3 (Fig. S1). These "cryptic" species were identified via concordance between mtDNA and rDNA marker patterns, and have subtle differences in morphological traits. Formal species descriptions are beyond the scope of this paper and will be addressed in future manuscripts. With the exception of 3 lineages, all rDNA sequences were identical to other specimens sharing the same morphological ID. Specimens identified as *A. richardsoni* were observed to have a single A/G substitution in the ITS1 region. Among individuals identified as *A. fonti*, 3 rDNA haplotypes were present; these haplotypes appeared to be associated with a geographic split between the northern (FO01 = Vermont, FO02 = Wisconsin) and southern (Louisiana/Texas, FO03 = FO04) extent of the species. Most interesting is the case of *A. kenti*, morphologically identified, but molecularly associated with 2 highly divergent clades (Fig. S1). Specimens sampled from the type locality of *A. kenti* (Mississippi, MS01-MS03) did not group as sister to the specimens retrieved from Texas or Arkansas (TK01-TK03, Figs. 3, S1). In contrast, the TK01-TK03 specimens had 28*s* rDNA sequence identical to that reported by Tkach and Mills (2011). Henceforth, we denote the *A. kenti* lineages as *A. kenti* MS – for those obtained from the Mississippi type locality – and *A. kenti* TK – for those whose rDNA corresponded to the sequence used to resurrect the species (Tkach and Mills, 2011). A future manuscript will provide taxonomic revisions (including morphology) within the genus.

The ND1 mtDNA region consisted of 676 bp, contained no gaps in either ingroup or outgroup species, and yielded 40 unique haplotypes among the ingroup specimens. Within both the ingroup and outgroup, no premature stop codons were found after translation of the sequences (using amino acid translation code 9 on GenBank). Table S3 reports pairwise ND1 *p*-distances. We draw attention to a few special cases. Among *A. fonti* specimens, three individuals (FO01, Vermont; FO02, Wisconsin; and FO03, Texas) have *p*-distances between 2.2 and 3.3% when compared against each other, but pairwise comparisons with the fourth individual (FO04, Louisiana) yields *p*-distances of 5.5–5.9%. Additionally, among the individuals associated with the *A. kenti* TK designation, the 3 individuals from Texas (TK01, TK02) and Arkansas (TK03) differed from each other by < 1%. Interspecific comparisons of nominal species at the ND1 typically showed a *p*-distance between 9.6 and 21.4% (Table S3). However, comparisons of *A. hamrumi*, *A. hirudicola*, and *A. turnbulli* were found to have *p*-differences between 1.6 and 4.9%, ranging as low as 1.6–2.4% between *A. hamrumi* and the two *A. turnbulli* specimens (Table S3). After accounting for identical sequences, the 49 original *Alloglossidium* individuals were condensed down to 40 unique sequences in the concatenated rDNA-ND1alignment.

3.2. Phylogenetic relationships

Bayesian analyses for the three datasets (rDNA-only, mtDNA-only, and concatenated) produced largely congruent trees at the base. Across all analyses there was strong support (Pp = 1.0) for the monophyly of all the specimens deemed to belong to *Alloglossidium* (Figs. 1, S1, S3). The concatenated and rDNA data provided strong support (Pps = 1.0) for a divergence between two distinct clades. Clade X contains all *Alloglossidium* spp. utilizing fish or crustacean final hosts, and Clade Y contains all *Alloglossidium* spp. utilizing leech final hosts (Figs. 3, S1). Though support is not as strong (Pps ~ 0.75), the mtDNA data showed the same split (Fig. S3). Within Clade Y, the concatenated and rDNA data provided strong support (Pps = 1.0) for a basal split involving *A. richardsoni* followed by *A. macrobdellensis* (Figs. 3, S1). These patterns were also observed with the mtDNA, but with less support (Pps = 0.74 and 0.9, respectively, Fig. S3). Within Clade X, *A. geminum* shows a basal split with the remaining species with good support in the concatenated and rDNA data (Pp = 1.0, Figs. 3, S1); this is also a pattern seen with the mtDNA but with less support (Pp = 0.75, Fig. S3). A clade containing *A. kenti* MS samples as well as *Alloglossidium* n. sp. 2 (samples from Canada) shows the next most basal split in Clade X with the concatenated and rDNA data (Figs. 3, S1), but this is not seen with the mtDNA data (Fig. S3). Additionally, 2 internal clades within Clade X (Clade I: *A. cardicolum* + *A. dolandi*; Clade II: *A. progeneticum* + *A. renale* + *A. greeri*) were well supported across all analyses (Pps ~ 1.0; Figs. 3, S1, S3).

It is important to note that a few topological incongruencies were observed among the internal relationships of both major clades (Figs. 3, S1, S3). In Clade X, the sister taxa to *A. anomophagis* differed depending on analysis (concatenated and rDNA = *A. fonti*; mtDNA = *A. floridense*), though the node had high support in all three datasets (concatenated: Pp = 0.84; rDNA and mtDNA: Pp > 0.95). Furthermore, the sister taxa to Clade II varied: both the concatenated and mtDNA analyses suggesting Clade I was the most likely candidate to be sister to Clade II (Pp > 0.60) while *A. floridense* was found to be the most closely related species when only considering the rDNA data (Pp = 0.62). Within Clade Y, the position of *A. turnbulli* differed between the rDNA and mtDNA analyses. In particular, the rDNA showed extensive divergence (see branch length in Fig. S1), but the mtDNA of *A. turnbulli* was very similar to that of *A. hamrumi* (*p*-distance = 1.6–2.4% (Fig. S3). While recognizing these incongruencies as potential caveats (which we return to in the Discussion), we defer to the more collective dataset of the concatenated rDNA-mtDNA as the most likely phylogenetic reconstruction for the genus. Therefore, the concatenated rDNA-mtDNA serves as the basis for phylogenetic hypothesis testing and ancestral reconstructions (Fig. 3).

3.3. Inference of ancestral states

Node A (Fig. 3), which was present in all 2143 sampled trees, shows the obligate 3-host pattern as the ancestral state to the genus *Alloglossidium* (average frequency = 0.95). Nodes B and C were present in 2140 trees and were found to have strong support for a particular life history, corresponding to the deep divergence between *Alloglossidium* species in fish/crustaceans and those species in leeches (Fig. 3). Node B, which contained all species utilizing fish or crustacean final hosts (Clade X), favored an ancestral 3-host pattern (average frequency = 0.99) while node C (Clade Y) had strong support for the 2-host, leech final host life cycle (0.98). Among the internal relationships within Clade X, an obligate 3-host life cycle was always supported as the ancestral condition. However, nodes D and E only appeared in a subset of trees (1721 and 1098, respectively). Among the trees that had nodes D and E, there was majority support for an obligate 3-host life cycle (0.93 for Node D and 0.60 for Node E). Finally, the most likely ancestral life history for node F (associated with Clade II) was also found to be the 3-host condition (0.76).

Because of the weaker support for the nodes noted above, there is less certainty in assessing the number of transitions going from an obligate 3-host state to an obligate 2-host crustacean state. However, the uncertainty can be taken into account by using the Summarize State Changes Over Trees to estimate the average number of times a transition occurs from one state to another across a set of probable trees from the Bayesian analysis. The likelihood based analysis estimated an average of 1.7 whereas the parsimony estimated an average of 2.6 times an obligate 3-host state changed to an obligate 2-host crustacean state.

3.4. Phylogenetic hypothesis testing

In the concatenated dataset, a comparison of marginal likelihoods obtained via the stepping stone approach (positive constraint, $H_1 = -14448.07$; negative constraint, $H_0 = -14440.10$) found very strong support against a forced monophyletic clade containing all precociously developing species compared to excluding *A. anomophagis* ($2\ln B_{10} = 15.94$). Results from the individual rDNA (H_1: −7788.44; H_0: −7781.83; $2\ln B_{10} = 13.22$) and mtDNA (H_1: −6687.68; H_0: −6679.52; $2\ln B_{10} = 16.32$) datasets also found strong support against a single origin of precociousness compared to excluding *A. anomophagis*.

3.5. Results from the analysis of the "progeneticum" clade

The mtDNA-only analysis of the internal relationships of Clade II (*A. progeneticum*, *A. renale*, and *A. greeri*), resulted in a polytomy across the ingroup taxa (Fig. 4). The lack of resolution is due to the shallow divergence in the mtDNA, possibly an indication of more recent events. We note that the non-precocious *A. progeneticum* (those with obligate 3-host life cycles) were by far the most prevalent life history pattern among populations. Only two precociously-developing lineages, geographically separated into either the Oconee River or Flint River drainages of Georgia, were identified (Fig. 4). Finally, *Alloglossidium renale* and *A. greeri* formed a highly supported clade (Pp = 1.0) nested within the non-precocious *A. progeneticum* lineages (Fig. 4).

4. Discussion

4.1. Interrelationships of Alloglossidium spp.

A molecular phylogeny of the genus *Alloglossidium* is a necessary first step towards investigating questions related to the evolution of changes in life cycle complexity. Without wading into the debate on the superiority of molecular (Scotland et al., 2003) or morphological data (rebuttal, Jenner, 2004) in creating the most valid phylogenetic reconstructions, the molecular phylogeny provides one major advance

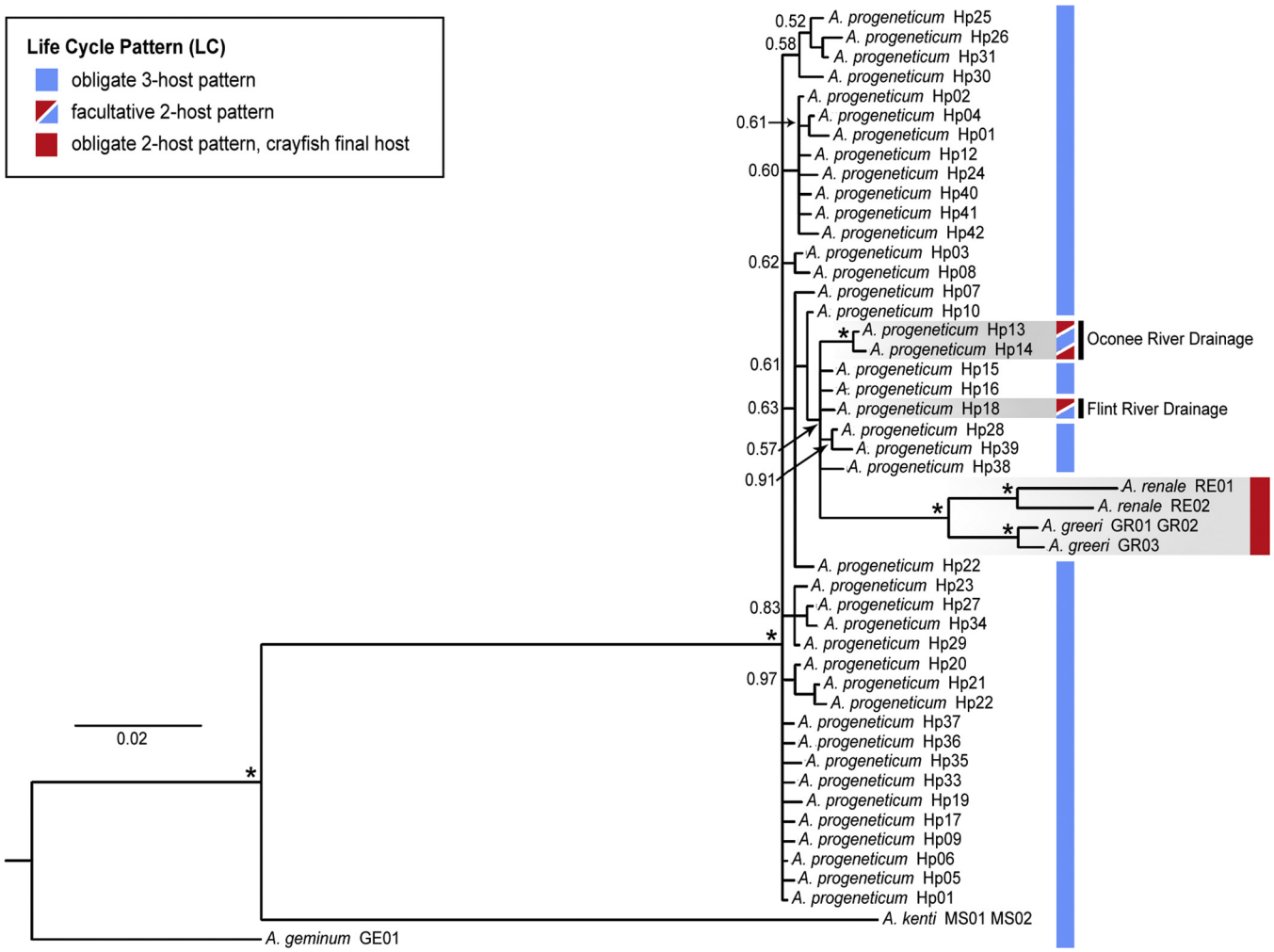

Fig. 4. The 50% majority rule consensus tree obtained from the Bayesian analysis of mitochondrial (ND1 mtDNA) haplotypes associated with *A. progeneticum*, *A. renale*, and *A. greeri*. Numbers at nodes correspond to the Bayesian posterior probabilities; * denote nodal support values of 1.0. *Alloglossidium geminum* and *A. kenti* were chosen as outgroups representing obligate 3-host life cycle patterns. Colors denote life cycle pattern. Facultative precocious lineages of *A. progeneticum* associated with the Oconee and Flint River drainages in Georgia are discussed further in text.

over the previously proposed morphological and life history-based phylogenies of *Alloglossidium*: a means of creating an independent assessment of relationships that can subsequently be used to address character trait evolution. The use of 2 independently evolving genes (rDNA and mtDNA) to construct the molecular phylogeny does support some previous findings based on morphology, but also shows some dramatic differences.

The molecular phylogeny does support Smythe and Font's (2001) placement of *A. dolandi*, *A. cardicolum*, and *A. richardsoni* within the genus *Alloglossidium*. It also supports the inclusion of *A. anomophagis* in the genus (Poinar et al., 1995). Both the molecular-based phylogenies (presented here) and the morphological-based phylogeny of Smythe and Font (2001) show a monophyletic clade (Clade Y) for the species with a 2-host pattern that use a leech final host. Also, Clade I, which included *A. cardicolum* and *A. dolandi*, was the only other relationship found to be congruent between the molecular and morphological phylogenies (compare Figs. 2 and 3).

The molecular phylogeny diverged from the findings of Smythe and Font (2001) in a number of notable ways. The concatenated dataset supported an early divergence between all species in leeches (Clade Y) and the remaining species in the genus (Clade X) where *A. geminum* had a basal split with the rest of the members of Clade X (Fig. 3). In contrast, Smythe and Font (2001) found *A. geminum* to be at the basal split to all other *Alloglossidium* species and where all species in leech final hosts formed a more recently diverged clade (Fig. 2 vs. Fig. 3). Furthermore, Smythe and Font (2001) found *A. richardsoni* was one of the most recently derived species using a leech host, whereas the molecular data indicated it had a basal split with the remaining species in the leech host clade (Clade Y, Fig. 3). Finally, sequence data supported the delineation of a "*progeneticum*" clade (Clade II; comprised of *A. progeneticum*, *A. renale*, and *A. greeri,* where the latter two are sister taxa). In contrast, the results of Smythe and Font (2001) do not show any of these three species being sister to one another, but rather as a series of sequential splitting events (Fig. 2).

4.2. Reconstructing the evolution of life cycle complexity

When using the morphology and life history-based tree of Smythe and Font (2001), the most parsimonious evolution of life cycle complexity was as follows: (1) the 3-host pattern (catfish final host) was ancestral, (2) the loss of the fish host occurred according to a progressive event from a facultative precocious to an obligate 2-host pattern with a crustacean final host (though it is equally parsimonious that the facultative precocious life cycle evolved on the *A. progeneticum* lineage itself, especially given the recent finding of non-precocious populations of this species; Kasl et al., 2015), and (3) a host-switching event from crustacean to leech final hosts. We also note that *A. renale*, *A. greeri*, *A. cardicolum*, and *A. dolandi* do not form a cyst (i.e., there is no encysted metacerarial stage). Hence, the phylogeny of Smythe and Font (2001) would lead to the interpretation of the loss of the cyst trait after the split with *A. progeneticum* and then the subsequent regaining of the cyst trait on the branch leading to the species using leech hosts.

The topological differences observed in the molecular-based phylogeny (Fig. 3) compared to the morphology/life history-based phylogeny (Fig. 2) lead to dramatic alterations in the interpretation of the life cycle changes observed in the genus *Alloglossidium*. Most notably, the molecular analyses showed a well-supported early divergence between the clades with species in leech hosts (Clade Y) and with all other *Alloglossidium* species (Clade X). Hence, the precocious life cycle that evolved in leeches happened early in the history of *Alloglossidium* (rather than late as inferred in Fig. 2) and independent of other precocious events found within Clade X (rather than a single precocious event followed by a host switch from crustacean to leech as inferred in Fig. 2). Within Clade X, there is strong support that the ancestral state of this clade is a 3-host pattern using ictalurid catfishes as a host (Fig. 3). At Node P, which shows a 3-host catfish state, however, there is a polytomy that includes the remaining species that show a precocious life cycle. Taking the concatenated-based tree along with the ancestral reconstruction analysis at face value, there are three independent evolutionary events that lead a precocious life cycle with the loss of the catfish host. The first is with a species that infects the body cavity of water fleas as a final host, *A. anomophagis* (node D, Fig. 3), which is sister to *A. fonti* (a 3-host catfish species). Second, the ancestral trait of Clade II, the "progeneticum" clade, (node F, Fig. 3) is most probably a 3-host catfish life cycle. Thus, the clade containing *A. renale* and *A. greeri*, which infect shrimp and crayfish, respectively, as final hosts, represents another independent loss of a catfish host. The third loss occurs on the branch leading to Clade I, which contains two species that infect crayfish as final hosts, *A. cardicolum* and *A. dolandi*. Here we take special note that although the last four mentioned species infect the antennal glands of crayfish, they inhabit the antennal gland in very different ways. Both *A. renale* and *A. greeri* reside inside the cavity of the gland (like a pea in a pod, personal observations), whereas *A. cardicolum* and *A. dolandi* reside in the nephridial tubules of the antennal gland (Turner and McKeever, 1993). These sites of infection patterns may explain the vastly different morphologies of these species (bulbous-oval for the former and filiform-long for the latter).

The "progeneticum" clade (Clade II, Fig. 3) itself is interesting in possibly having multiple independent life cycle transitions. First, it is intriguing that the *A. renale-A. greeri* clade is nested within *A. progeneticum* samples (based on the mtDNA alone, Fig. 4), most of which originated from locations that show no evidence of precocious development (Kasl et al., 2015). Hence, the *A. renale-A. greeri* clade may be descended from an *A. progeneticum* population. Second, we have only found the facultative precocious populations of *A. progeneticum* from two independent river drainages in Georgia and these samples have lower genetic divergence to samples from non-precocious populations than they do to one another (see networks given in Kasl et al., 2015, McAllister et al., 2016). At face value, these results suggest possible separate origins of facultative precociousness among *A. progeneticum* populations. Nonetheless, the shallow divergence of the mtDNA data alone precludes definitive assessment of two origins of a facultative life cycle among populations of *A. progeneticum* and more data will be needed to test this hypothesis.

With regard to the ancestral state of the genus *Alloglossidium*, the probability of the ancestral state for Node A favors the 3-host pattern as the ancestral condition (Fig. 3). This result is due to the fact that *M. simplex*, which has a 3-host character state, is sister to the genus *Alloglossidium* (Hernández-Mena et al., 2016; Sokolov and Shchenkov, 2017). As more taxa are sampled within the superfamily Plagiorchioidea it is possible a more closely related outgroup could be found. If we take a conservative approach and do not designate the character state of *M. simplex*, then the ancestral state of Node A becomes equivocal.

In relation to the evolution of the cyst, the molecular-based phylogeny again leads to a different interpretation compared to the morphological-based phylogeny. In particular, there is no reversal where there was a loss then regain of the cyst as inferred from Fig. 2. Rather, there appears to be two independent losses: one leading to the *A. renale-A. greeri* clade within Clade II and the other leading to Clade I (*A. cardicolum* and *A. dolandi*).

4.3. Caveats and incongruencies

As the biggest caveat, we need to recognize that the life history data for most of the species within the genus *Alloglossidium* is incomplete, especially for cryptic species found herein (which will be described in future manuscripts). For example, the designation of a 3-host life cycle is given to all species that have been found in a catfish host. However, first and second intermediate hosts have only been described for *A. corti* (McCoy, 1928; McMullen, 1935; Crawford, 1937) and this was prior to the discovery of cryptic species in the genus (Tkach and Mills, 2011,

results herein). Hence, it is not known what species these older studies were working with. It may be that some of the species we designated as 3-host in fact have facultative 2-host patterns as well. Likewise, only one of the leech species, *A. macrobdellensis*, has had its life cycle elucidated (Corkum and Beckerdite, 1975). Future survey work could alter the life cycle or cyst patterns we outlined above.

We also note that Node E does not have strong support (Fig. 3) and that there were incongruencies between the rDNA and mtDNA phylogenies in the sister taxa of *A. anomophagis*. These issues lead to uncertainty in inferring the number of transitions from an obligate 3-host state to an obligate 2-host crustacean state. However, after accounting for uncertainty in tree estimation, the average number of times this transition was observed to occur was greater than 1 (1.7 or 2.6 depending on method of estimation). In the case of the *A. anomophagis*, the sister species had a 3-host pattern in both the rDNA or mtDNA datasets. Importantly, results of the phylogenetic constraint tests strongly favored a hypothesis that *A. anomophagis* was not sister taxa to Clades I and II further supporting the inference of an independent loss of a catfish host. Hence, all additional analyses that take into account the lack of resolution of the terminal relationships within Clade X (observed in the phylogenetic analysis itself, Fig. 3) lead to the inference of more than one transition from an obligate 3-host state to an obligate 2-host crustacean state.

The other incongruencies of note occurred within Clade Y. The rDNA (Fig. S1) showed *A. turnbulli* to be very divergent, but yet the mtDNA showed close similarity to *A. hamrumi* (Fig. S2). In general, the tip species in Clade Y had incongruent relationships between rDNA and mtDNA. Divergence among the species using a 2-host cycle and leech final host may have been rapid leading to incomplete lineage sorting patterns. However, we also note that many of these species can co-occur within the same host individuals (Vande Vusse et al., 1981, personal observations). Thus, hybridization that leads to mtDNA introgression might also be a plausible explanation.

4.4. Hypotheses for the evolution of precocious life cycles

Here, we relate possible hypotheses of precocious life cycle evolution to our study system. The most common hypothesis (logical, not theoretical) put forth for why precocious life cycles have evolved is that there are low rates of transmission from the second host (i.e. prey) to a predatory third host (Lefebvre and Poulin, 2005a). This argument may work for some but certainly not all species of *Alloglossidium*. We do not know the historical ecological conditions in which the lineage of 2-host species using leeches arose (Fig. 3). Nevertheless, in current day habitats such as lakes of the upper Midwest and swamps, leeches are syntopic with catfishes, which prey upon leeches. In the streams with the facultative precocious populations of *A. progeneticum*, catfishes, which we commonly observe to have crayfish parts in their guts, are syntopic with crayfish infected with gravid trematodes (Kasl et al., 2015). Thus, among the species with a 2-host cycle, leech final host and among *A. progeneticum* populations, the absence of or reduced exposure to a predator serving as a third host does not provide a reasonable explanation of the evolution of precociousness. Such an explanation could apply to *A. anomophagis*, especially since the populations we sampled are from ephemeral habitats, i.e., fish are absent, and may apply to Clade I because these species (*A. cardicolum* and *A. dolandi*) infect procambarid “ditch” crayfish where again the aquatic habitat is ephemeral and not conducive to supporting catfishes.

Recent models by Parker and colleagues focused on adding hosts via trophic interactions to a parasite’s life cycle (Parker et al., 2003; Parker et al., 2015a). Reversal of the conditions in their models highlights factors that may represent costs to life cycle truncation. They proposed two models (downward and upward incorporation), both of which include a cost to generalism, i.e., the ability to survive in more than one host. In the downward incorporation model, which is similar to a model by Choisy et al. (2003), an intermediate host is added to facilitate survivorship to the final host (e.g., the intermediate host fills a trophic gap between the parasite infective stage and the final host; Benesh et al., 2014). Reversal of the conditions of the model (e.g., high generalism cost and low survivorship benefits using an intermediate host) could result in the loss of the intermediate host. However, in the precocious life cycles of *Alloglossidium*, the final host is lost. Thus, the downward incorporation model is not relevant to the evolution of precocious life cycles where the predator (the previous final host) is lost, as is the case for *Alloglossidium*.

In the upward incorporation model, the existing final host becomes the prey of a predator, which in turn, becomes the new final host. The driving force behind this addition is increased parasite fecundity as a higher trophic level host is assumed to enable greater parasite body size and survivorship. We note that body size is often the only means to infer helminth parasite fecundity (Poulin, 2007). A study by Lefebvre and Poulin (2005b) showed that precocious trematode species had no differences in body size from those species with 3-host life cycles. Qualitatively, the body size patterns in the genus *Alloglossidium* are consistent with this pattern: the body sizes for species in leeches (Carney and Brooks, 1991) overlap those species in catfishes (Kasl, unpublished) and there is no difference in body size of gravid worms of *A. progeneticum* from catfish or crayfish (Kasl et al., 2015). Therefore, body size does not appear to be a constraint to the evolution of precocious life cycles within the genus *Alloglossidium* or trematodes in general (Lefebvre and Poulin, 2005b).

Given that basic life history traits do not appear to be constraints to evolving a precocious life cycle, Lefebvre and Poulin (2005b) raise the important question, why have precocious life cycles not evolved more often? Parker et al. (2015a) pose a similar question, why suppress reproduction in an earlier host? A logical argument historically put forth to answer these questions is that precocious life cycles might promote inbreeding (e.g., encysted *A. progeneticum* are forced to self-mate) and thus, are susceptible to inbreeding depression (Grabda-Kazubska, 1976; Font, 1980). This argument now has theoretical support in the mating system model of Brown et al. (2001), which argues that a complex life cycle provides a concentration effect to ensure that potential mates meet up. A complex life cycle may also allow more mixing of offspring before they meet back in the final host, thereby reducing inbreeding via sib-mating (Criscione and Blouin, 2006; Gorton et al., 2012). In the case of trematodes, a complex life cycle may also allow mixing of clones coming from the intermediate host as mating between identical clones is equal to self-mating (Rauch et al., 2005; Criscione et al., 2005). With regard to *A. progeneticum*, the precocious individuals are forced to self-mate (Kasl et al., 2015), thus what remains to be tested is if there is any cost to this inbreeding. Coupled with no change in body size and if there is no inbreeding depression, the shorter life cycle may win out due to the cost of generalism. In the case of the species using a 2-host cycle with a leech final host, there may still be plenty of chances to outcross (i.e., there is no inbreeding at all), so again the shorter life cycle would be favored with generalism costs. Our future work will be measuring both inbreeding depression and degrees of sibling or clone mixing to test these hypotheses.

Acknowledgements

The authors are thankful to J. Detwiler, M. Gorton Janecka, K. Peart, A. Ince, K. Bass, N. Stokes, A. Sakla, and B. Trejo for their assistance with conducting field surveys. The authors would also like to acknowledge D.J. Marcogliese and S.A. Locke for providing specimens. We also thank H. Blackmon for advice on trait evolution and topology test analyses. This work was supported by the National Science Foundation, DEB #1145508 (CDC and WFF) and #1655147 (CDC).

Data accessibility

GenBank accessions are given in Table 1. A nexus file with the

MrBayes command blocks is given in the supplemental material. This file contains the alignment of the sequences and shows the corresponding alignment positions given in Table S2.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ympev.2018.04.027.

References

Auld, S.K.J.R., Tinsley, M.C., 2015. The evolutionary ecology of complex lifecycle parasites: linking phenomena with mechanisms. Heredity 114, 125–132.

Benesh, D.P., Chubb, J.C., Parker, G.A., 2013. Complex life cycles: why refrain from growth before reproduction in the adult niche? Am. Nat. 181, 39–51.

Benesh, D.P., Chubb, J.C., Parker, G.A., 2014. The trophic vacuum and the evolution of complex life cycles in trophically transmitted helminths. Proc. Roy. Soc. B 281, 20141462.

Brabec, J., Kostadinova, A., Scholz, T., Littlewood, D.T.J., 2015. Complete mitochondrial genomes and nuclear ribosomal RNA operons of two species of *Diplostomum* (Platyhelminthes: Trematoda): a molecular resource for taxonomy and molecular epidemiology of important fish pathogens. Parasit. Vectors. 8, 336.

Bray, R.A., Gibson, D., Jones, A. (Eds.), 2008. Keys to the Trematoda,, vol 3 CAB International and Natural History Museum, London.

Briscoe, A.G., Bray, R.A., Brabec, J., Littlewood, D.T.J., 2016. The mitochondrial genome and ribosomal operon of *Brachycladium* goliath (Digenea: Brachycladiidae) recovered from a stranded minke whale. Parasitol. Int. 65, 271–275.

Brown, J.M., Thompson, R.C., 2016. Bayes factors unmask highly variable information content, bias, and extreme influence in phylogenomic analyses. Syst. Biol. 66, 517–530.

Brown, S.P., Renaud, F., Guégan, J.-F., Thomas, F., 2001. Evolution of trophic transmission in parasites: the need to reach a mating place? J. Evol. Biol. 14, 815–820.

Brooks, D.R., 2003. Lessons from a quiet classic. J. Parasitol. 89, 878–885.

Carney, J.P., Brooks, D.R., 1991. Phylogenetic analysis of *Alloglossidium* Simer, 1929 (Digenea: Plagiorchiiformes: Macroderoididae) with discussion of the origin of truncated life cycle patterns in the genus. J. Parasitol. 77, 890–900.

Carvalho-Sobrinho, J.G., Alverson, W.S., Alcantara, S., Queiroz, L.P., Mota, A.C., Baum, D.A., 2016. Revisiting the phylogeny of Bombacoideae (Malvaceae): Novel relationships, morphologically cohesive clades, and a new tribal classification based on multilocus phylogenetic analyses. Mol. Phylogenet. Evol. 101, 56–74.

Chauhan, V., Pandey, A.K., 2014. Structure and evolution of the pod in *Indigofera* (Fabaceae) reveals a trend towards small thin indehiscent pods. Bot. J. Linn. Soc. 176, 260–276.

Choisy, M., Brown, S.P., Lafferty, K.D., Thomas, F., 2003. Evolution of tropic transmission in parasites: why add intermediate hosts? Am. Nat. 162, 172–181.

Corkum, K.C., Beckerdite, F.W., 1975. Observations on the life history of *Alloglossidium macrobdellensis* (Trematoda: Macroderoididae) from *Macrobdella ditetra* (Hirudinea: Hirudinidae). Am. Midl. Nat. 93, 484–491.

Crawford, W.W., 1937. A further contribution to the life history of *Alloglossidium corti* (Lamont), with especial reference to dragonfly naiads as second intermediate hosts. J. Parasitol. 23, 389–399.

Cribb, T.H., Bray, R.A., Olson, P.D., Littlewood, D.T.J., 2003. Life cycle evolution in the Digenea: a new perspective from phylogeny. Adv. Parasitol. 54, 197–254.

Criscione, C.D., Blouin, M.S., 2004. Life cycles shape parasite evolution: comparative population genetics of salmon trematodes. Evolution 58, 198–202.

Criscione, C.D., Blouin, M.S., 2006. Minimal selfing, few clones, and no among-host genetic structure in a hermaphroditic parasite with asexual larval propagation. Evolution 60, 553–562.

Criscione, C.D., Poulin, R., Blouin, M.S., 2005. Molecular ecology of parasites: elucidating ecological and microevolutionary processes. Mol. Ecol. 14 (8), 2247–2257.

Font, W.F., 1980. The effect of progenesis on the evolution of *Alloglossidium* (Trematoda, Plagiorchiida, Macroderoididae). Acta Parasitol. Pol. 27, 173–183.

Feutry, P., Castelin, M., Ovenden, J.R., Dettai, A., Robinet, T., Cruaud, C., Keith, P., 2013. Evolution of diadromy in fish: Insights from a tropical genus (*Kuhlia* species). Am. Nat. 181, 52–63.

Grabda-Kazubska, B., 1976. Abbreviation of the life cycles in plagiorchid trematodes. General remarks. Acta Parasitol. Pol. 24, 125–141.

Gorton, M.J., Kasl, E.L., Detwiler, J.T., Criscione, C.D., 2012. Testing local scale panmixia provides insights into the cryptic ecology, evolution, and epidemiology of metazoan animal parasites. Parasitology 139, 981–997.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Sympos. Ser. 4, 95–98.

Hammerschmidt, K., Koch, K., Milinski, M., Chubb, J.C., Parker, G.A., 2009. When to go: optimization of host switching in parasites with complex life cycles. Evolution 63, 1976–1986.

Hasegawa, M., Iida, Y., Yano, T., Takaiwa, F., Iwabuchi, M., 1985. Phylogenetic relationships among eukaryotic kingdoms inferred from ribosomal RNA sequences. J. Mol. Evol. 22 (1), 32–38.

Hernández-Mena, D.I., Mendoza-Garfias, B., Ornelas-García, C.P., Perez-Ponce de León, G., 2016. Phylogenetic position of *Magnivitellinum* Kloss, 1966 and *Perezitrema* Baruš & Moravec, 1967 (Trematoda: Plagiorchioidea: Macroderoididae) inferred from partial 28S rDNA sequences, with the establishment of Alloglossidiidae n. fam. Syst. Parasitol. 93, 525–538.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.

Jenner, R.A., 2004. Accepting partnership by submission? Morphological phylogenetics in a molecular millennium. Syst. Biol. 53, 333–342.

Kasl, E.L., Fayton, T.J., Font, W.F., Criscione, C.D., 2014. *Alloglossidium floridense* n. sp. (Digenea: Macroderoididae) from a Spring Run in North Central Florida. J. Parasitol. 100, 121–126.

Kasl, E.L., McAllister, C.T., Robison, H.W., Connior, M.B., Font, W.F., Criscione, C.D., 2015. Evolutionary consequence of a change in life cycle complexity: a link between precocious development and evolution towards female-biased sex allocation in a hermaphroditic parasite. Evolution 69, 3156–3170.

Kass, R.E., Rafferty, A.E., 1995. Bayes factors. J Am. Stat. Assoc. 90, 773–795.

Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

Lafferty, K.D., Kuris, A.M., 2002. Trophic strategies, animal diversity and body size. Trends Ecol. Evol. 17, 507–513.

Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Bio. Evo. 29, 1695–1701.

Lefebvre, F., Poulin, R., 2005a. Progenesis in digenean trematodes: a taxonomic and synthetic overview of species reproducing in their second intermediate hosts. Parasitology 130, 587–605.

Lefebvre, F., Poulin, R., 2005b. Life history constraints on the evolution of abbreviated life cycles in parasitic trematodes. J. Helminth. 79, 47–53.

Lyons, J., 1989. Changes in the abundance of small littoral-zone fishes in Lake Mendota. Wisconsin. Can. J. Zoo. 67, 2910–2916.

Maddison, W. P. and D.R. Maddison. 2010. Mesquite: a modular system for evolutionary analysis. Version 3.40 http://mesquiteproject.org.

McAllister, C.T., Kasl, E.L., Robison, H.W., Connior, M.B., Font, W.F., Trauth, S.E., Criscione, C.D., 2016. New host records for *Alloglossidium progeneticum* (Digenea: Macroderoididae) in crayfishes (Decapoda: Cambaridae) from Arkansas and Oklahoma, USA. Comp. Parasitol. 83, 255–259.

McCoy, O.R., 1928. Life history studies on trematodes from Missouri. J. Parasitol. 14, 207–228.

McMullen, D.B., 1935. The life histories and classification of two allocreadiid-like plagiorchids from fish, *Macroderoides typicus* and *Alloglossidium corti* (Lamont). J. Parasitol. 21, 369–380.

Olson, P.D., Cribb, T.H., Tkach, V.V., Bray, R.A., Littlewood, D.T.J., 2003. Phylogeny and classification of the Digenea (Platyhelminthes: Trematoda). Int. J. Parasitol. 33, 733–755.

Parker, G.A., Chubb, J.C., Ball, M.A., Roberts, G.N., 2003. Evolution of complex life cycles in helminth parasites. Nature 425, 480–484.

Parker, G.A., Ball, M.A., Chubb, J.C., 2015a. Evolution of complex life cycles in trophically transmitted helminths. I. Host incorporation and trophic ascent. J. Evol. Biol. 28, 267–291.

Parker, G.A., Ball, M.A., Chubb, J.C., 2015b. Evolution of complex life cycles in trophically transmitted helminths. II. How do life-history stages adapt to their hosts. J. Evol. Biol. 28, 292–304.

Poinar Jr., G.O., Schwarts, S.S., Cameron, G., 1995. *Alloglossidium anomophagis* sp. n. (Trematoda: Plagiorchiidae) exhibiting progenesis in water fleas (Anomopoda: Daphniidae). Experientia 51, 388–390.

Poulin, R., 2007. Evolutionary Ecology of Parasites, second ed. Princeton University Press, Princeton, NJ.

Poulin, R., Randhawa, H.S., 2015. Evolution of parasitism along convergent lines: from ecology to genomics. Parasitology 142, S6–S15.

Rauch, G., Kalbe, M., Reusch, T.B.H., 2005. How a complex life cycle can improve a parasite's sex life. J. Evol. Biol. 18, 1069–1075.

Riggs, M., Ulmer, M.J., 1983. Host-parasite relationships of helminth parasites in leeches of the genus Haemopis. I. Associations at the individual host level. Trans. Amer. Micro. Soc. 102(3), 213–226.

Scotland, R.W., Olmstead, R.G., Bennett, J.R., 2003. Phylogeny reconstruction: the role of morphology. Syst. Biol. 52, 539–548.

Simer, P.H., 1929. Fish trematodes from the lower Tallahatchie River. Am. Mid. Nat. 11, 563–588.

Smythe, A.B., Font, W.F., 2001. Phylogenetic analysis of *Alloglossidium* (Digenea: Macroderoididae) and related genera: life-cycle evolution and taxonomic revision. J. Parasitol. 87, 386–391.

Sokolov, S.G., Shchenkov, S.V., 2017. Phylogenetic position of the family Orientocreadiidae within the superfamily Plagiorchioidea (Trematoda) based on partial 28S rDNA sequence. Parasitol. Res. 116, 2831–2844.

Vande Vusse, F.J., Fish, T.D., Neumann, M.P., 1981. Adult digenea from upper midwest hirudinid leeches. J. Parasitol. 67, 717–720.

Tkach, V.V., Mills, A.M., 2011. *Alloglossidium fonti* sp. nov. (Digenea: Macroderoididae) from black bullheads in Minnesota with molecular differentiation from congeners and resurrection of *Alloglossidium kenti*. Acta Parasitol. 56, 154–162.

Tkach, V.V., Greiman, S.E., Steffes, K.R., 2013. *Alloglossidium demshini* sp. nov. (Digenea: Macroderoididae) from leeches in Minnesota. Acta Parasitol. 58, 434–440.

Turner, H.M., McKeever, S., 1993. *Alloglossoides dolandi* n. sp. (Trematoda: Macroderoididae) from the crayfish *Procambarus epicyrtus* in Georgia. J. Parasitol. 79, 353–355.

Xie, W., Lewis, P.O., Fan, Y., Kuo, L., Chen, M., 2011. Improving marginal likelihood estimation for bayesian phylogenetic model selection. Syst. Biol. 60 (2), 150–160.