An investigation into the evolution of Chirality in Gastropod Molluscs

Submitted May 2023, in fulfilment of the conditions for the award of the degree MRes Molecular Genetics

Eleanor Bull

Supervised by Prof. Angus Davison

School of Life Sciences

University of Nottingham

Impact of the duplication of diaphanous-related formin on the evolution of chirality within Pulmonata

1.1 Abstract

Creating new genetic material via gene duplications plays a fundamental role in the emergence of morphological diversity within the animal kingdom. Two forms of duplication of an actin nucleator diaphanous-related formin (diaph) are associated with the establishment of left/right polarity in the chirally variable phyla of gastropoda. In the freshwater clade, Hygrophila, the emergence of dextral Lymnaeidae from a sinistral ancestor has been attributed to tandem gene duplication. In the terrestrial clade, Stylommatophora, however, an ancestral whole genome duplication is associated with the emergence of chiral variation. In the dextral bradybaena similaris, individuals who produced both sinistral and dextral offspring showed a reduced expression of the duplicated diaph gene (diaph-b) during embryonic development. This prompted the hypothesis that chiral direction in Stylommatophora is dictated by regulating diaph-b gene expression. We aimed to perform a phylogenetic study on species within each clade first to confirm the location of duplication events and then evaluate the association with chirality. Here, we report that diaphanous-related formin has undergone two separate duplications throughout Pulmonate evolution. In Hygrophila, the single gene tandem-duplication occurred in an ancestor of Lymnaeidae, whereas in Stylommatophora, an ancestral whole-genome duplication resulted in the two paralogs present. In the sole sinistral species studied, Euhadra quaesita, we failed to resolve a diaph-b paralog suggesting down-regulation of formin may be associated with chirality in Stylommatophora. However, the pathological nature of the mutation in both species suggests additional gene pairs may be involved in the determination of chirality.

1.2 Introduction

A longstanding ambition of biologists has been to elucidate the precise mechanism governing the establishment and patterning of the three body axes and the extent to which they are conserved across phyla (Brown and Wolpert, 1990; Boorman and Shimeld, 2002; Blum *et al.*, 2014). This ambition aided our understanding of evolution and congenital disease, as a broad spectrum of defects results from errors in axis formation (Blum *et al.*, 2014).

Several signalling pathways play critical roles in patterning embryonic axes. The nodal cascade involved in L/R patterning was long thought to be a vertebrate invention (Schier, 2003), yet evidence from snails (Grande and Patel, 2009) and sea urchins (Luo and Su, 2012) suggest it is an ancestral feature of Bilateria.

Complete gene-regulatory networks are, however, only a single piece of the puzzle. Upstream physical prompts are required to trigger polarised gene expression, as cells are not innately spatially aware. External cues determine the polarity of the first two axes (DV, AP). In the nematode, the anterior-posterior axis of the zygote is associated with the entry point of the sperm. The location of the sperm-donated pronucleus and associated centrosomes specifies the posterior pole. Gravity has been linked to the establishment of the dorsal-ventral axis (Mercola and Levin, 2001; Vandenberg and Levin, 2013; McDowell *et al.*, 2016).

The definition and orientation of the third axis are harder to conceive, as nothing in the universe acts to distinguish left from right. An eloquent analogy lies in the Ozma problem – trying to teach left and right to an alien via radio (Gardener, 1970). Left and right on Earth may be distinguishable by describing the relative placement of the heart and liver

to the mid-line; however, reversed L/R polarity on the other planet would render confusion. A longstanding concept for symmetry breaking is Wolpert's chiral F molecule. Two of the arms align with the DV and PA axis. The direction of the third dictates polarity (Brown and Wolpert, 1990).

The notion that the regulation of cytoskeletal dynamics plays a conserved role in establishing L/R polarity is gaining traction (McDowell *et al.*, 2016). The actomyosin cytoskeleton is a dynamic scaffold formed of helical actin filaments, actin-binding proteins, and Myosin motors. Interaction between elements of the actin-myosin cytoskeleton is fundamental to multiple cellular functions. Actin-based unconventional motor protein Myosin has been shown to drive the twisting of the gut and genitalia in *Drosophila melanogaster* (Spéder and Noselli, 2007). Handedness is dictated by the motor protein domain. The presence of Myo1D compared to Myo1C resulted in dextral rather than sinistral twisting (Lebreton *et al.*, 2018; Chougule *et al.*, 2020). Myosin has also been implicated in establishing the L/R axis in the nematode *C. elegans* (Wood, 1991; Naganathan *et al.*, 2014; Middelkoop *et al.*, 2021) and vertebrate *Xenopus laevis* (Tingler *et al.*, 2018).

In addition, the regulation of actin polymerization plays a role in establishing L/R asymmetry. Several actin-nucleating proteins essential for promoting single-cell and entire organismal chirality have been identified (Tee *et al.*, 2021). One such protein is formin. In conjunction with other signalling molecules and proteins, formins are conserved multi-domain proteins that govern cytoskeletal remodelling (Heimsath and Higgs, 2012). A conserved feature of the family is a 400bp formin-homology domain (FH2), which binds to the barbed end of actin filaments stabilizing and permitting the steady elongation of the filament (Higgs and Peterson, 2005; Courtemanche *et al.*, 2013). Based on phylogenetic analysis of the FH2 domain, metazoan formins divide into seven groups: DIA, FMN, FHOD, defilin, INF, FRL, and DAAM (Higgs and Peterson, 2005).

Diaphanous-related formins (DIA) are a sub-family, distinguishable by the presence of an additional formin homology domain (FH3) along with amino-terminal GTPase-binding and carboxy-terminal Dia-autoregulatory domains (referred to as GBD and DAD respectively). The integration of the GBD, DAD and g-protein modulates actin filament polymerization (Davison, 2020). GBD and DAD regions interact, forming a complex preventing C-terminus FH domains from polymerizing actin. The presence of G-protein Rho GTPase at the GBD disrupts this interaction. It has been shown that pharmacological inhibition of diaphanous-related formin results in a chiral reversal in both snails and frogs(Davison *et al.*, 2016).

Snails are remarkable as they not only display outward asymmetry by way of the spiral shell but also remain the only group to vary in their asymmetry ordinarily. Dextral (right-handed) coiling snails are the most common; however, sinistral (left-handed) snails occur across all taxonomic levels (Schilthuizen and Davison, 2005). The handedness of the adult is determined early in development and is intrinsically linked to the embryonic spiral cleavage pattern (Lambert, 2010). During spiral cleavage, the blastomeres divide initially symmetrically and later asymmetrically to generate a highly stereotypic, spiral-like arrangement towards the animal pole. A 45-degree shift in the mitotic spindle twists the emerging micromeres clockwise or anticlockwise relative to the macromeres. The direction of the micromeres after the third cleave dictates L/R polarity (Martín-Durán, Vellutini and Hejnol, 2016). The dextral coiling shell and corresponding body result from a clockwise twist; a sinistral snail, on the other hand, results in a sinistral snail. A single maternally inherited locus controls this event.

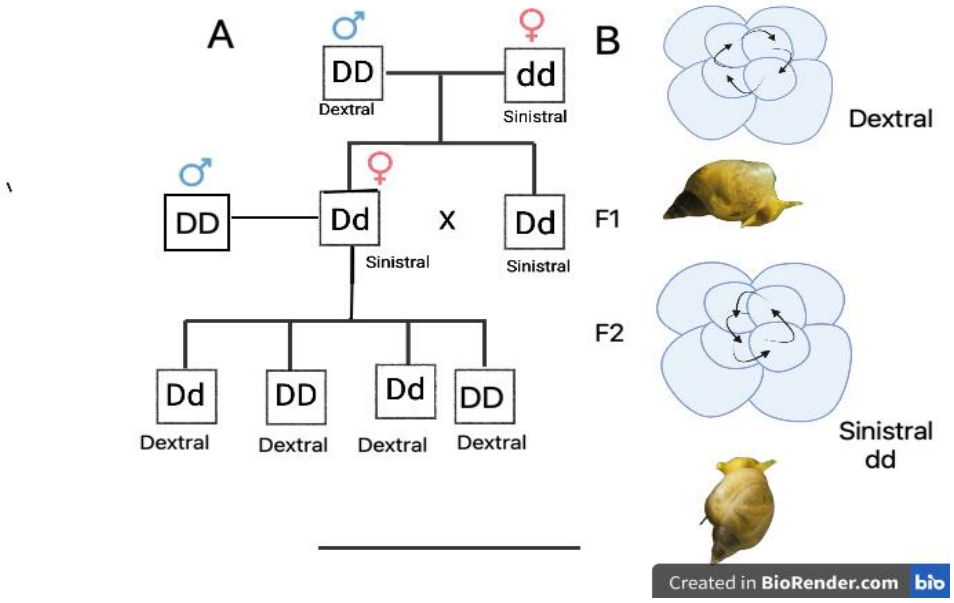
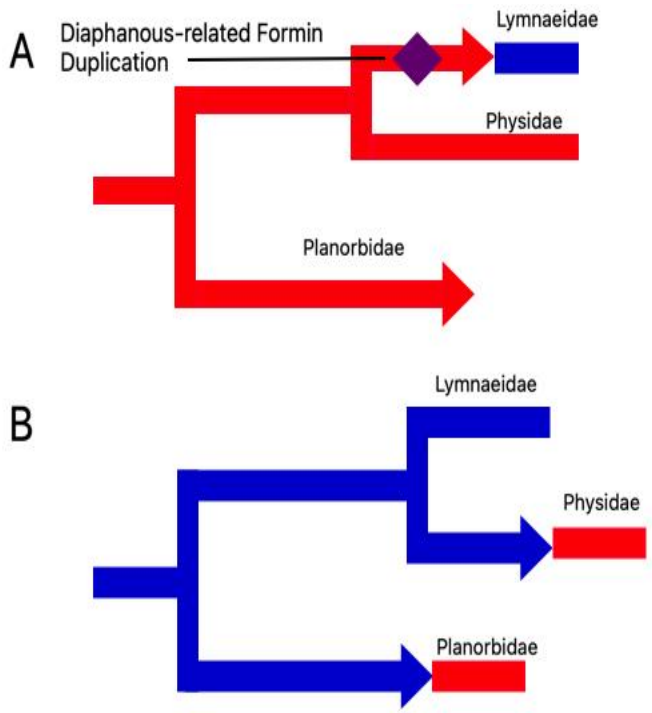


Fig. 1.
A. Cross between homozygous dextral and sinistral snails demonstrating the maternal inheritance of chirality.
B. Spiral cleavage mechanism which results in the sinistral and dextral adult

Boycott and Diver (1923) performed extensive crosses of the chirally variable snail species *Lymnaea peregra*. The progeny of a sinistral mother (dd) and dextral father (DD) were phenotypically sinistral yet genotypically dextral (Dd), inconsistent with a Mendelian pattern of inheritance (Fig.1a). This pattern of maternal or delayed inheritance has since been observed in all other snail species studied, the dominant allele, however, does vary (Sturtevant, 1923; Schilthuizen and Davison, 2005).

A century on from the elucidation of the mechanism of chiral inheritance, a duplication of diaphanous-related formin has been implicated in L/R asymmetry determination of two different superorders within Panpulmonata, Hygrophila and Stylommatophora (Davison *et al.*, 2016; Kuroda *et al.*, 2016; Abe and Kuroda, 2019; Noda, Satoh and Asami, 2019).



A schematic depicting the two hypotheses for the evolution of chirality in three Hygrophilan families. (sinistral = red; dextral = blue) (inspired by Davison et al, 2016)

(A) Either, sinistraltiy evolved once from a dextral ancestor with a reversal to dextrality occuring in ancestor of Lymnaidae, faciltated by a tandem-duplication in diaphanous-related formin
(B) The alternative is sinistrality evolved from a dextral ancester on two separate occasions.

Hygrophila is a superorder of air-breathing aquatic snails and limpets found to inhabit a range of freshwater habitats worldwide. Morphologically, they are similar to other pulmonates in that they have a single lung and lack an operculum; however, their eyes are located at the base of their tentacles, as opposed to the tips (Ponder, W.F and Lindberg, David. R, 2008*).*

Molecular studies have led to the classification of the clade into two superfamilies: Lymnaeoidea, which consists of Acroloxidae, Burnupiidae, Lymnaeidae, Physidae and Planorbidae and the Chilinoidea, which consist of families Chilinidae and Latiidae (Saadi, Davison and Wade, 2020).

Chirality is variable across the superorder. Chilinoidea is an entirely dextral superfamily, whereas two of the five families within Lymnaeoidea are sinistral. To understand the evolution of the chirality within Hydrophilia, Saadi, Wade and Davison (2020) mapped the coiling direction onto the phylogeny. The branching pattern suggests two parsimonious hypotheses for the origin of sinistrality in the group (Fig.2).

Either sinistrality evolved from a dextral ancestor on two separate occasions, in Planorbidae and Physidae, or there is a single origin with a reversal to dextrality occurring in an ancestor of Lymnaeaidea (Davison *et al.*, 2016; Saadi, Davison and Wade, 2020).

Molecular genetic studies have provided insight into the possible evolutionary path, concentrating on two species of Lymnaeidae, *Lymnaea stagnalis* and *Lymnaea peregra*. In both species, a pair of tandemly-duplicated diaphanous-related formin (*diaph*) genes were linked with the chiral locus. Analysis of the amino-acid sequences revealed an 89.4% similarity between the two proteins and conservation of diaphanous-related formin domains (FH 2,3 and GBD, DID, DAD) (Kuroda *et al.*, 2016). The two genes were named differently by the two research groups. *Ldia1* and *Ldia2* were chosen by Davison et al. (Davison *et al.*, 2016). Kuroda et al. (2016), chose *Lsdia2* and *Lsdia1*. In this paper, the genes are called according to Davison *et al.* (2016), with the first gene (Ldia1) referring to the ancestral copy. A single base-pair deletion in exon 3 one copy was found in rare sinistral morphs. The deletion triggered a frameshift mutation resulting in a severely truncated protein. Homozygous knock-outs of Ldia2 in dextral morphs resulted in sinistral coiling offspring, supporting the link between the diaph and chirality (Abe and Kuroda, 2019).

To explore the links between diaphanous-related formin and the evolution of chirality within the clade, the genomes and transcriptomes of other species were studied (Davison *et al.*, 2016a). *Physa acuta* and *Biomphararia glabrata*, belonging to sinistral clades Physidae and Planorbidae, resolved a single copy more genetically similar to Ldia1. Two copies of diaph (Ldia1, Ldia2) were determined from the other Lymnaeaidae species studied, *Lymnaea trunculata*. In addition to the molecular evidence, the presence of only a Ldia1 homolog in the sinistral species suggests a single origin of sinistrality. The reversal to dextrality was facilitated by the duplication of Diaph in an ancestor of Lymnaeaidae. The duplication is seemly required for dextrality in Lymnaediae. The link between the diaphanous-related formin gene duplication and the handedness determination of dextral Lymnaeae prompted whether this mechanism of chiral regulation was exclusive to Hygrophila or familiar to all pulmonates.

The most diverse pulmonate clade is Stylommatophora consisting of approximately 20,000 species of air-breathing snails and slugs. A morphological feature common to all is the presence of two pairs of head tentacles, one pair with an eye at the tip. Adaptation to a selection of terrestrial habitats, including tropical rainforests to deserts and temperate hedgerows, has led to morphological diversity. Striking variation occurs in body size, shell shape, colour and coiling direction. Most species within Stylommatophora are dextral; however, sinistral morphs are present at all taxonomic levels. Excluding the entirely sinistral family Clausiliidae, commonly known as door snails, most sinistral individuals belong to the diverse family Camaenidae. Mechanisms by which rare sinistral alleles can overcome strong positive dependent selection to become fixed in a population are increasingly well-researched. However, the precise genetic mechanisms underlying the chiral variation are still unknown.

To shed light on the mechanism in Stylommatophora, Noda et al., (2019) studied *Bradybaena similaris,* a small invasive species of land snails with light brown dextral coiling globular shells known as the Asian tramp snail. Previous research has revealed a racemic mutant capable of producing an even split of sinistral and dextral offspring (Utsuno and Asami, 2010). RNA-Seq analysis of wild-type and racemic mutant at various

stages of development was carried out to understand if the chiral reversal was also linked to a frameshift mutation in diaph (Noda, Satoh and Asami, 2019).

Analogous to dextral *Lymnaea stagnalis*, *Diaph* was found to be associated with chiral determination. Transcripts of two *diaph* genes (*diaph-a* and *diaph-b*) were identified. The proteins share the same conserved domains yet differ in precise amino-acid composition – only 70% of the amino-acid sequence is conserved (Noda, Satoh and Asami, 2019). Molecular conservation between FH2 regions of *diaph-a* and the other genes established it as the ancestral copy.

Transcript alignment revealed no frameshift mutation in the duplication copy, *diaph-b*, of the racemic mutants, which could explain the reversal; however, quantitative gene expression analysis revealed a reduction in *diaph-b* expression in embryos of racemic mutants compared to wild-types. The expression of the ancestral gene remained consistent between the two morphs.

Noda, Satoh and Asami (2019hypothesizeded that the low levels of *diaph-b* expression in racemic mutations could explain the presence of dextral offspring. Entire sinistral broods may result if the expression reduces further. The results suggested that gene duplication associated with a regulatory mechanism is responsible for the stable chiral variation. A single copy is likely lethal if mutated during development, as evidenced in *C. elegans* (Castrillon and Wasserman, 1994). In addition, there is a low hatch rate in sinistral *Lymnaea stagnalis* (Schilthuizen and Davison, 2005; Davison, Barton and Clarke, 2009).

The first step to evaluate if the mechanism specified in *Bradybaeana similaris* is ultimately conserved within Stylommatophora is to identify the origin of the gene duplication. Examination of transcriptomes of sinistral snails can be used to examine if, indeed, a reduction in expression diaph-b paralog is a feature of sinistrality, hypothesized by Noda.

Chromosomal level genome assemblies of two giant African snails, *Lissachatina immaculata* (Liu *et al.*, 2021) and *Lissachatina fulica* (Guo *et al.*, 2019), revealed evidence of an ancestral whole-genome duplication (WGD). A WGD was not evident from the genome analysis of Hygrophilan species –*Biomphalaria glabrata* (Liu *et al.*, 2021), suggesting two separate events leading to duplication of diaphanous-related formin in the two clades, a single gene tandem duplication in the ancestor of Lymnaeadae within Hygrophila and a WGD at the base of Stylommatophora. Expanding the pool of species studied to include additional sinistral species and members of the non-achatinoid clade Helicina is required to confirm these hypotheses.

Over the last decade, there has been a rapid increase in the number of assembled genomes. Over 5500 metazoan genomes were published due to the following technological advances. Next-generation short read platforms, e.g. Illumina and Ion Torrent, allowed for the generation of large numbers of reads for a fraction of the cost. Second, the advent and rapid improvement of read length agnostic sequencing (longest reads ~4mb) from Oxford Nanopore has enabled the generation of highly contiguous assemblies. Long reads stretch large regions of the genome, proven unamenable to short-read assembly.

However, the speed at which sequences are being generated has led to a knowledge gap regarding the number and quality of genomic resources available for each species. Hotaling et al. (2021) provide a phyla-level evaluation of the genomes published genomes, citing specifically the lack of representation in certain groups such as Mollusca. 67 of 50'000 species have assembled genomes. The species represented is not mentioned. Confusingly, across the two newly curated comprehensive molluscan genomic databases (MolluscanDB), only 24 species are represented, only two of which are gastropoda.

Considering this, we aimed to curate our own genomic and transcriptomic database for Gastropoda – focusing on Stylommatophora, Hygrophila and related clades. Databases. We will include taxonomic details, information pertaining to the quality of genomic

resources and chirality for each species. After collating the data, paralogs present in each species are extracted and assembled using the known diaphanous-related formin sequences from *Bradybaenae similaris* and *Lymnaea stagnalis* as references. Next, we uncover the location of duplication event(s) resulting in the two paralogs by using the alignments to generate two phylogenetic trees. This data will allow the association between chirality and the presence of the derived paralog Ldia2 in transcriptomic or genomic data to be investigated.

1.2 Methods

Data Collection

First, a systematic search of scientific literature using appropriate keywords was carried out to find evidence of species with an assembled transcriptome or genome. The resources were then located and downloaded from the sequence repositories linked in the paper. In addition, for all species, details of the taxa, references, assembly type, size and tissue type were assembled.

Gene Extraction

A tblastn search (matrix Blosum62, expectation value 1.0E-10) was performed on each transcriptome (Table1) using duplicated diaphanous-related formin protein sequences from *Lymnaea stagnalis* (Ldia1,2-nomenclature from Davison *et al.*, 2016) or *Bradybaenae similaris* as a query. Extracted sequences were then manually assembled in BioEdit (Hall, 1999). For the genome sequences, blastn (Ldia1,2 gene query and genome subject) was performed using a blast-python script written by Peter Mulhair (University of Leeds), enabling the extraction of similar sequences by using biopython to remove introns. The script was executed on the command line.

Alignment

Bioedit 7.2.6.1 was used to manually assemble using the diaphanous formin genes *Bradybaena* diaphanous-related paralogs *diaph-a* and *diaph-b* and *L.stagnalis Ldia1, two* as references(Hall, 1999). The differences in the Diaphanous GTPase binding domain 120-150 aa) and the Diaphanous auto-regulatory domain (DAD) were used to distinguish the paralogs. A highly variable region mid-gene, along with regions of poor-quality alignment at the beginning and end of the sequences, were removed (Fig 3,4). X regions were used in the final phylogeny as a compromise between the number of species and poorly resolved regions.

Phylogenetic Analysis

Maximum likelihood (ML) phylogenetic analyses were performed using IQ-TREE 2 on each genomic data set, compromising diaphanous-related formin genes in Stylommatophora and Hygrophila (Felsenstein, 1981; Minh *et al.*, 2020). IQ-TREE2 was run on the command line using iqtree -s <FASTA FILE> -o <"OUTGROUP"> -bb 1000. ModelFinder was implemented to determine the optimal model of nucleotide substitution. The chosen model was TIM + F + I + G4. The maximum-likelihood tree was found by conducting searches from an initial maximum-parsimony starting point, selected by log-likelihood scores (Kalyaanamoorthy *et al.*, 2017). Branch support was calculated by performing 1000 ultrafast bootstrap replicates (Minh, Nguyen and von Haeseler, 2013; Hoang *et al.*, 2018). An additional algorithm, hill-climbing nearest neighbour interchange (NNI), prevented over-estimation of branch supports. The trees were rooted using the R package, ape (version 5.6.2). ggTree (version 1.14.6) package in R was used for annotation and visualization (Yu *et al.*, 2017).

1.3 Results

A total of 52 species from the Panpulmonata clades: Stylommatophora, Hygrophila, and Systellommatophora were discovered to have assembled genomes or transcriptomes (detailed in Table 1). Most of the sequences were found uploaded to the following genomic databases: NCBI Genbank, DDBJ and Dryad.

17 of the 52 species found to have genomic resources belong to the clade of freshwater snails Hygrophila. The clade is formed of two superfamilies, Lymnaeoidea and Chilinoidea; however, genomic resources were only discovered for four families within Lymnaeoidea: Lymnaeidae, Planorbidae, Physidae, and Acroloxidae. However, differences were evident in the resources available. An assembled transcriptome was found for each species; however, a large size range was evident. Transcriptomes of two species of Planorbidae, *Biom-pferreri* and *Indoplanorbis exustus,* differ in size by almost two orders of magnitude. Scaffold-level assembled genomes were found for two sinistral species, *Biomphalaria glabrata* and *Physa acuta.* Contig and scaffold-level genomes were uncovered for two species of dextral Lymnaeidae, *Lymnaea stagnalis* and *Lymnaea auricularis*.

Stylommatophora is comprised of three clades Helicina ("non-achatinoid clade"), Achatina ("achatinoid clade") and Scolodontina (Wade, Mordan and Clarke, 2001; Saadi and Wade, 2019). Genomic resources were found for 27 dextral and one sinistral species within the sub-order of Helicina. Two species, Cepaea *nemoralis* and *Candidula unifasciata,* were found to have published scaffold-level genomes. A chromosome-level genome is now available for another Helicina species, *Arion vulgaris*; however, the public release of the data was too late for it to be included in the phylogenetic element of the research. The remaining 25 species are represented in the study by assembled transcriptomes. In addition, chromosomal level assemblies were found of two members of the clade, Achatina. The third clade Scolodontina is not represented due to a lack of available genomic resources.

Table 1 Overview of genomic resources available for Panpulmonate species. ** denotes species not included in the phylogenetic study.

Classification	Species	Chirality	Transcriptome	Size (mb)	Tissue Type	Genome	Size (Gb)	Level of Assembly	Reference
Clade Panpulmonata									
Order Hygrophila									
Superfamily Lymnaeoidea									
Family Lymnaeidae Subfamily Lymnaeinae	*Lymnaea stagnalis*	Dextral	✓	148	whole-body adult	✓	0.882	Contig	(Davison *et al.*, 2016; Seppälä *et al.*, 2021)
	Radix peregra	Dextral	✓						(Kuroda *et al.*, 2016; Abe and Kuroda, 2019)
	Radix auricularis	Dextral	✓	78	Adult	✓	0.909	Scaffold	(Feldmeyer *et al.*, 2015; Schell *et al.*, 2017)
	Radix balthica	Dextral	✓	19,70	Embryonic				(Tills, Truebano and Rundle, 2015; Tills *et al.*, 2018)
	Radix sp. 3	Dextral	✓	78	Adult				(Feldmeyer *et al.*, 2015)
	Radix sp. 5	Dextral	✓	78	Adult				(Feldmeyer *et al.*, 2015)
	Galba cubensis	Dextral	✓	22	Adult				(Burgarella *et al.*, 2015)
	Galba schirazensis	Dextral	✓	33	Adult				
	Galba trunculata	Dextral	✓	63	Adult				(Burgarella *et al.*, 2015)
	Pseudosuccinea columnella	Dextral	✓	252	Juvenile foot				(Alba, Tetreau, Chaparro, Sánchez, A. Vázquez, *et al.*, 2019)
Family Physidae **Subfamily** Physinae	*Physella acuta*	Sinistral	✓	279		✓	0.804	Scaffold	(Schultz, Adema and Kamel, 2018; Schultz, Bu and Adema, 2018)
Family Planorbidae **Subfamily** Planorbinae	*Biomphalaria glabrata*	Sinistral	✓	92		✓	0.976	Scaffold	(Tennessen *et al.*, 2020)
	Indoplanorbis exustus	Sinistral	✓	316					

	Planorbarius corneus	Sinistral	✓	87					
	Isidorella newcombi	Sinistral	✓	66					(Ubrihien *et al.*, 2017)
	Biomphalaria pfefferi	Sinistral	✓	4					(Buddenborg *et al.*, 2017)
Family Acroloxidae	*Acroloxus*	Dextral(internal)	✓	43					(Wade and Davison, no date)
Order Stylommatophora									
Clade Achatinoid									
Superfamily Achatinoidea									
Family Achatinidae	*Lissachatina fulica*	Dextral				✓	1.89	Chromosome	(Guo *et al.*, 2019b)
	Lissachatina immaculata	Dextral				✓	1.67	Chromosome	(Liu *et al.*, 2021b)
Suborder Helicina (Non-Achatinoid Clade)									
Infraorder Rhytidoidei									
Superfamily Rhytidoidea									
Family Rhytididae **Subfamily** Rhytidinae	*Austrorhytida capillacea*	Dextral	✓	52	Whole body-adult				(Teasdale *et al.*, 2016)
	Terrycarlessia turbinata	Dextral	✓	75	Whole body-adult				(Teasdale *et al.*, 2016)
	Victaphanta atramentaria	Dextral	✓	53	Whole body-adult				(Teasdale *et al.*, 2016)
Infraorder Pupilloidei									
Superfamily Pupilloidea									

Family Enidae **Subfamily** Eninae	*Apoecus apertus*	Dextral	✓	85	Whole body-adult				(Teasdale *et al.*, 2016)
Family Cochlicopae **Subfamily** Cochlicopinae	*Cochlicopa lubrica*	Dextral	✓	74	Whole body-adult				(Teasdale *et al.*, 2016)
Family Cerastidae	*Amimopina macleayi*	Dextral	✓	59	Whole body-adult				(Teasdale *et al.*, 2016)
Superfamily Helicoidea									
Family Helicidae	*Cepaea nemoralis*	Dextral	✓	129	Adult foot	✓	3.54	Scaffold	(Kerkvliet *et al.*, 2017; Saenko *et al.*, 2021a)
	Cornu aspersum	Dextral	✓	96	Whole body-adult				(Teasdale *et al.*, 2016)
Family Camaenidae	*Austrochloritis kosciuszkoensis*	Dextral	✓	60	Whole body-adult				(Teasdale *et al.*, 2016)
	Chloritobadistes victoriae	Dextral	✓	89	Whole body-adult				(Teasdale *et al.*, 2016)
	Ramogenia challengeri	Dextral	✓	58	Whole body-adult				(Teasdale *et al.*, 2016)
	Sphaerospira fraseri	Dextral	✓	121	Whole body-adult				(Teasdale *et al.*, 2016)
Subfamily Bradybaeninae	*Satsuma myomphala*	Dextral	✓	65	Adult visceral mass				(Kang *et al.*, 2017)
	Aegista quelpartensis	Dextral	✓	181	Adult				(Kang, Patnaik, Hwang, Park, Wang, *et al.*, 2016)
	Aegista chejuensis	Dextral	✓	163	Adult				(Kang, Patnaik, Hwang, Park, Wang, *et al.*, 2016)
	Koreanohadra kurodana	Dextral	✓	147	Adult visceral mass				(Kang, Patnaik, Hwang, Park, Chung, *et al.*, 2016)
	Bradybaena similaris	Dextral/Sinistral	✓	3500	Embryonic				(Noda, Satoh and Asami, 2019b)

	Euhadra quaesita	Sinistral	✓	1800	Mantel				(Shimizu *et al.*, 2019a)
	Thersites novaehollandiae	Dextral	✓	67	Whole body-adult				(Teasdale *et al.*, 2016)
Family Geomitridae	*Cochlicella acuta*	Dextral	✓	49	Adult -CNS				(Zhao *et al.*, 2018)
	Candidula unifasciata	Dextral			Adult foot	✓	1.356	Scaffold	(Chueca, Schell and Pfenninger, 2021)
Infraorder Limacoidea									
Superfamily Trochomorphoidea									
Family Dyakiidae	*Asperitas stuartiae*	Dextral	✓	63	Whole body-adult				(Teasdale *et al.*, 2016)
Family Microcystidae	*Lamprocystis* sp.	Dextral	✓	89	Whole body-adult				(Teasdale *et al.*, 2016)
Superfamily Limacoidea									
Family Milacidae	*Milax gigates*	Dextral	✓	51	Whole body-adult				(Teasdale *et al.*, 2016)
Family Oxychilidae	*Oxychilus alliarius*	Dextral	✓	93	Whole body-adult				(Teasdale *et al.*, 2016)
Family Limacidae	*Limax flavus*	Dextral	✓	71					(Teasdale *et al.*, 2016)
Family Helicarionidae	*Fastosarion virens*	Dextral	✓	78	Whole body-adult				(Teasdale *et al.*, 2016)
Superfamily Arionoidea									
Family Arionidae	*Arion vulgaris***				Whole body-adult	✓	1.5	Chromosome	(Chen *et al.*, 2022)**
Order Systellommatophora									
Superfamily Ellobioidea									

Family Ellobiidae	*Cassidula angulifera*	Dextral	✓	88	Whole body-adult	(Teasdale *et al.*, 2016)
Superfamily Otinoidea						
Family Smeagolidae	*Smeagol phillipensis*		✓	72	Whole body-adult	(Teasdale *et al.*, 2016)
Superfamily Veronicellidea						
Family Veronicellidae	*Semperula maculata*		✓	90	Whole body-adult	(Teasdale *et al.*, 2016)

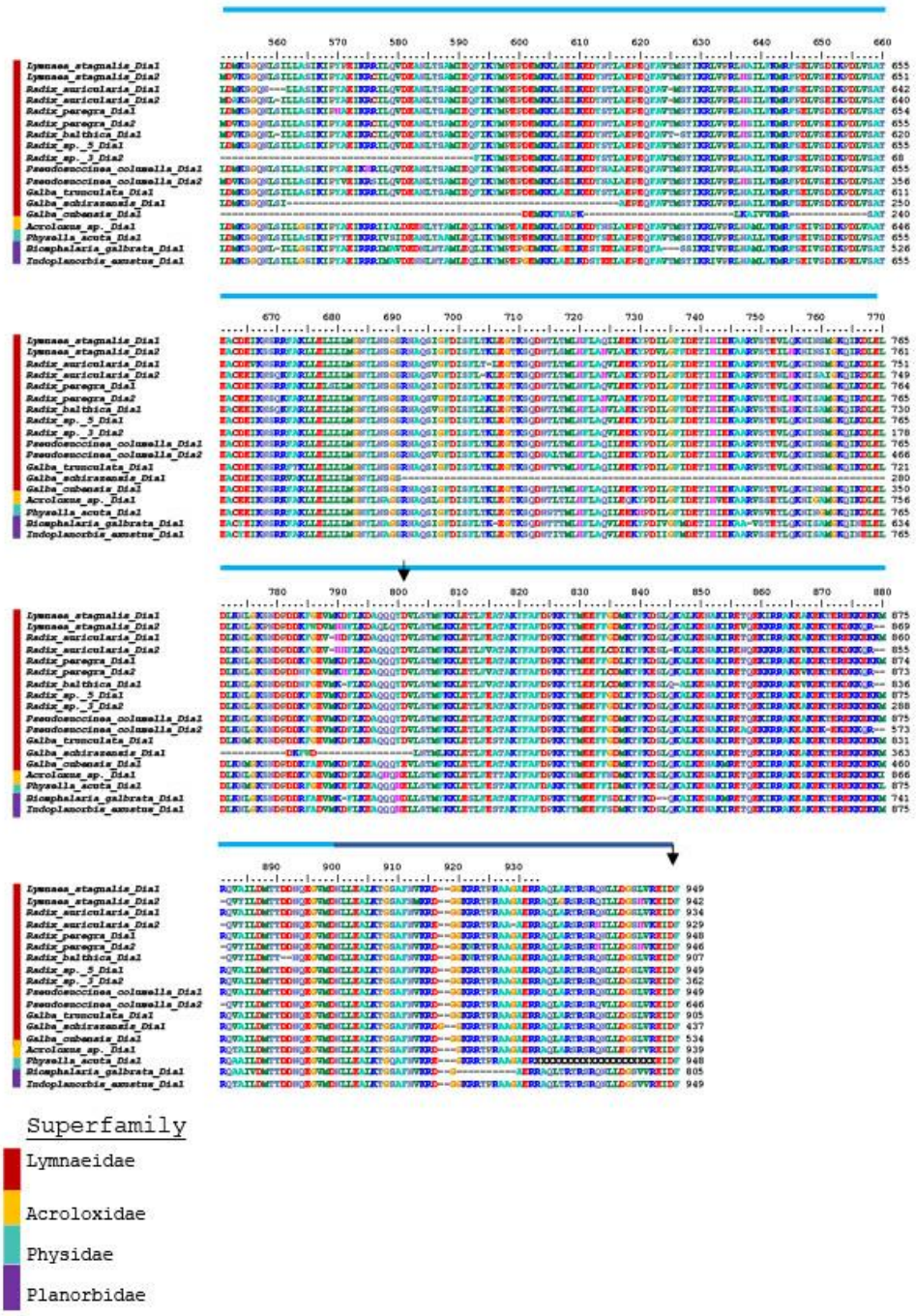
Diaphanous-related formin genes in Hygrophila

Two copies of Diaphanous-related formin were extracted in three of the Lymnaeidae genera studied- *Radix* Montfort 1810, *Lymnaea* Lamark 1799 and *Pseudosuccinea* Baker, 1908. The paralogs are distinguishable by the base-pair differences in DAD and GBD regions. However, only a single copy in alignment with Ldia1 was resolved from the three *Galba* species studied. Moreover, the resolution of the Galba paralogs was low in contrast to the remaining Lymnaeidae genes recovered.

Diaphanous-formin genes are tandemly duplicated in *Radix auricularia* (contig M UZC01001993.1 scaffold 289), albeit only the ancestral paralog *Ldia1* was resolved from the transcriptome data. The transcriptome of *Pseudosuccinea columella* exhibited transcripts of paralogs of the ancestral Ldia1 and duplicated *Ldia2.* However, gene expression data from 3 of the 5 *Radix* species studied resolved only *Ldia1* homologs.

Single gene copies were recovered from transcriptome and genomic resources of individuals of sinistral families *Planorbidae*, *Physidae* and dextral *Acroloxidae. Planorbidae corneus* resolved a single copy of a *Ldia1* homolog, albeit the low quality of the assembly resulted in omittance from the phylogeny (Fig. 3). All display more significant similarity to the ancestral paralog- *Ldia1* than to the perceivably derived *Ldia2*. Diaphanous formin gene copies could not be resolved from *Planorbidae* species *Biomphalaria pfeifferi and Isidorella newcombi* transcriptomes.


A


Superfamily
Lymnaeidae
Acroloxidae
Physidae
Planorbidae

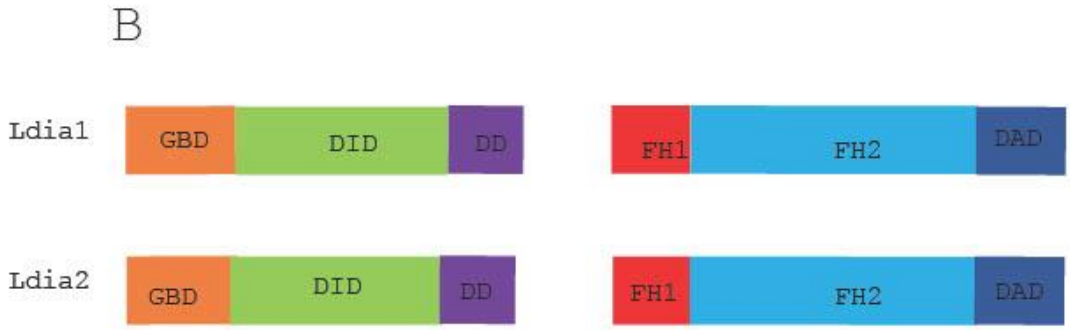


Figure 3.

A) An alignment of amino-acid sequences of Dia obtained from 14 Hygrophilan species. The black arrows show the regions used for assembling the phylogenetic trees. The location of the truncation in Ldia2 is represented by the blue arrow. Horizontal bars link the AA sequences to the corresponding Ldia protein domains shown in section B.

B) Dia protein domain structures. GBD; Diaphanous GTPase binding domain DAD; Diaphanous autoregulatory domain DD; Dimerization domain FH1; Formin homology domain 1 FH2; Formin homology domain and Diaphanous auto-regulatory domain DAD.

Molecular Phylogeny of Order Hygrophila

The ML phylogeny (Fig.4) displays with strong bootstrap clustering of species within Hygrophila families: *Lymnaeaidae, Planorbidae, and Physidae* in agreement with previous phylogenies (Davison *et al.*, 2016b; Ahmed J. Saadi, Davison and Wade, 2020). In addition, within Lymnaeidae, there is a strong division of the two paralogs (Ldia1, Ldia2). The three Galba species branch off prior to the remaining two genera; however, this is not well supported.

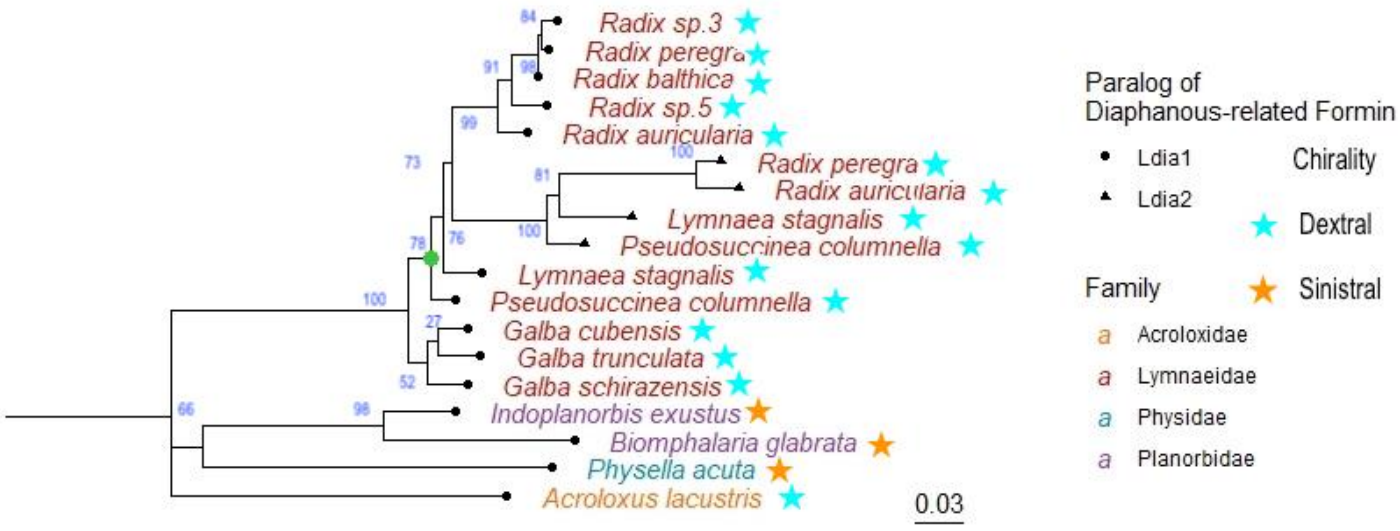


Fig. 4. A ML phylogeny for diaphanous-related formin genes from 14 species across the dextral Lymnaeidae, sinistral Physidae, and Planorbidae families. A Lime green dot indicates the origin of gene duplication. The phylogeny is based on 462 nucleotide sites, and genetic distances are estimated using a transition model (TIM + F + I + G4). Based on 1000 replicates, bootstrap values pinpoint the percentage of support for individual branches. The scale bar represents the number of changes per 100 nucleotides. A lack of suitable outgroups contributed to the decision to leave the phylogeny unrooted.

Diaphanous-related formin genes in Stylommatophora

Achatinina (Achatinoid Clade)

Diaphanous-related formin is duplicated in both members of the invasive genus Lissachatina studied. Each copy is located on a different chromosome. The ancestral copy Diaph-a resides on chromosome 17 in *Lissachatina fulica* and chromosome 18 in *Lissachatina immaculata*, yet the believed to-be derived copy resides on chromosomes 29 and 20, respectively.

Helicina (Non-achatinoid clade)

Helicina is a suborder of Stylommatophora, sub-divided into 11 morphologically diverse infra-orders (Saadi, Davison and Wade, 2020). Helicoidea is a morphologically and ecologically diverse superfamily of terrestrial gastropod molluscs that belongs to the Helicina infra-order Helicoidei. Helicoidea is made up of 5,600 species divided into 19 distinct families, many of which are both economically and evolutionarily important. However, despite this, only 15 species from 4 distinct families have published genomes or transcriptomes, none of which are of a chromosomal level (Table 1).

Geomitridae is a family of small to medium land snails and is characterized by the presence of a free right ommatophoral retractor and double stimulatory apparatus associated with adaption to shrubland habitats. Both paralogs of diaphanous-related formin (*Diaph-a*, *Diaph-b*) were found on different contigs in the genome of dextral *Candidula unifasicata*. The scaffold locations of Cuni 708 and Cuni 1141 were unavailable. The second family member studied, *Cochlicella acuta*– pointed snail, resolved only the ancestral *Diaph-a* gene. However, the low quality of the gene restricted it from being included in the phylogeny.

Another family within Helicoidea studied was Helicidae. Helicidae, are a large, diverse family of air-breathing land snails native to Eurasia. Helicidae, or typical snails, have soft bodies capable of fully retracting into patterned globular shells. Both paralogs of Diaphanous related formin are expressed in the Helicidae species studied: dextral *Cornu aspersum* (garden snail) and dextral *Cepaea nemoralis* (grove snail). Genome sequencing of the *Cepaea nemoralis* revealed contigs of the paralogs, Diaph-a on ctg407 and Diaph-b on tig00402505 on unplaced scaffolds. The relative location of the contigs is unknown.

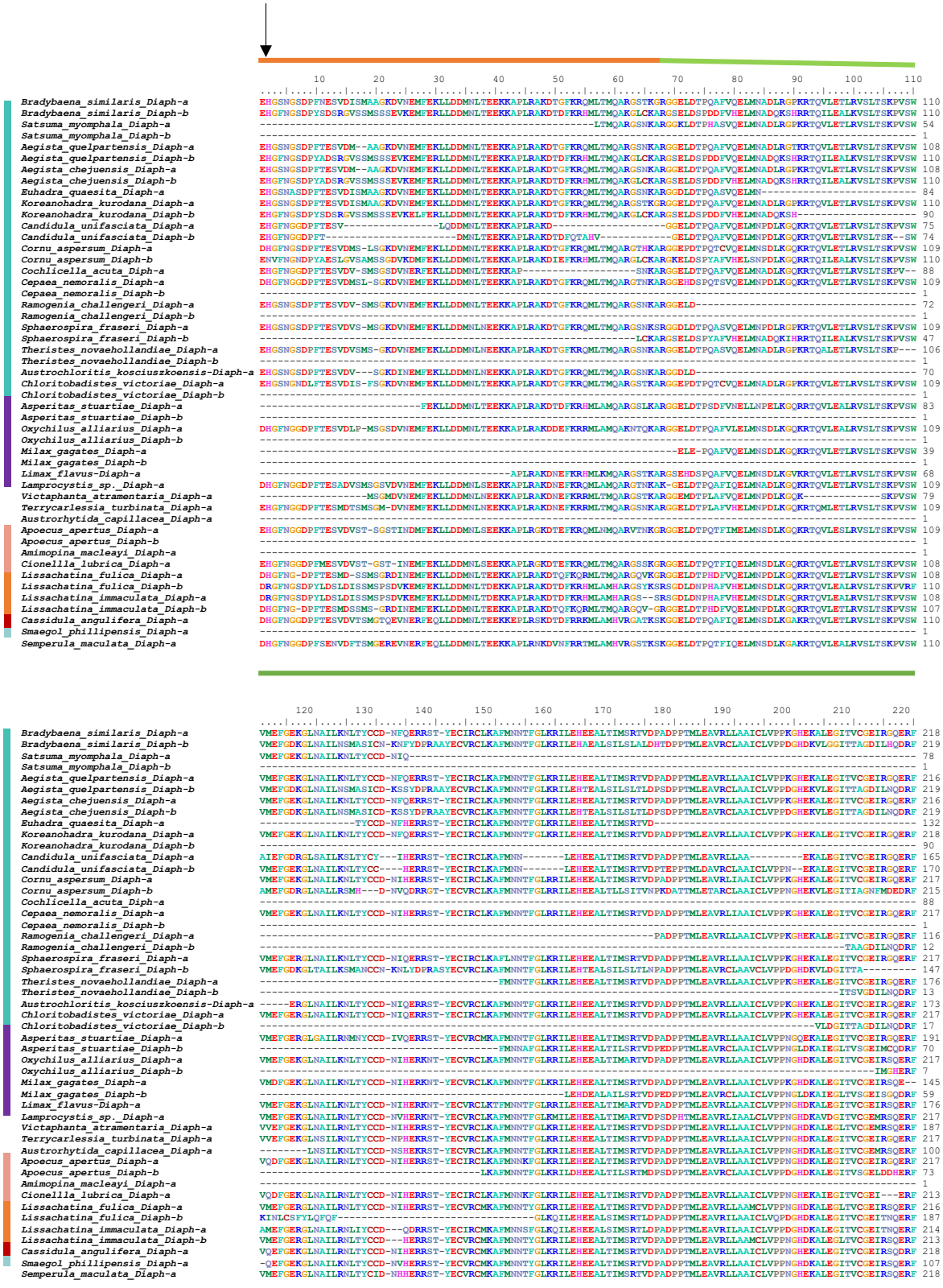
Another family within Helicidae studied is Camaenidae. Camaenidae is one of the most diverse families in Stylommatophora, exhibiting a variety of shell sizes, colours and chirality. The uppermost branch represents the Australian camaenids. Both paralogs of the diaphanous related-formin genes were resolved from transcriptomes of three species: *Ramogenia challengeri, Sphaerospira fraseri*, *Chloritobadistes victoriae*, *Thersites novaehollandiae*. However, only the ancestral Diaph gene was resolved from the *Austrochloritis kosciuszkoensis* transcriptome. The lower branch represents a selection of species from the sub-family Bradybaeninae, globular snails predominantly found in Asia. Both Diaph paralogs are resolved from four primarily dextral species endemic to Korea, clustering with the genes from *Bradybaena similaris* isolated by Noda et al., 2019. In contrast, the single primarily sinistral species in Stylommatophora analysis, *Euhadra quaesita,* appeared to express a copy of the ancestral *diaphanous formin* gene.

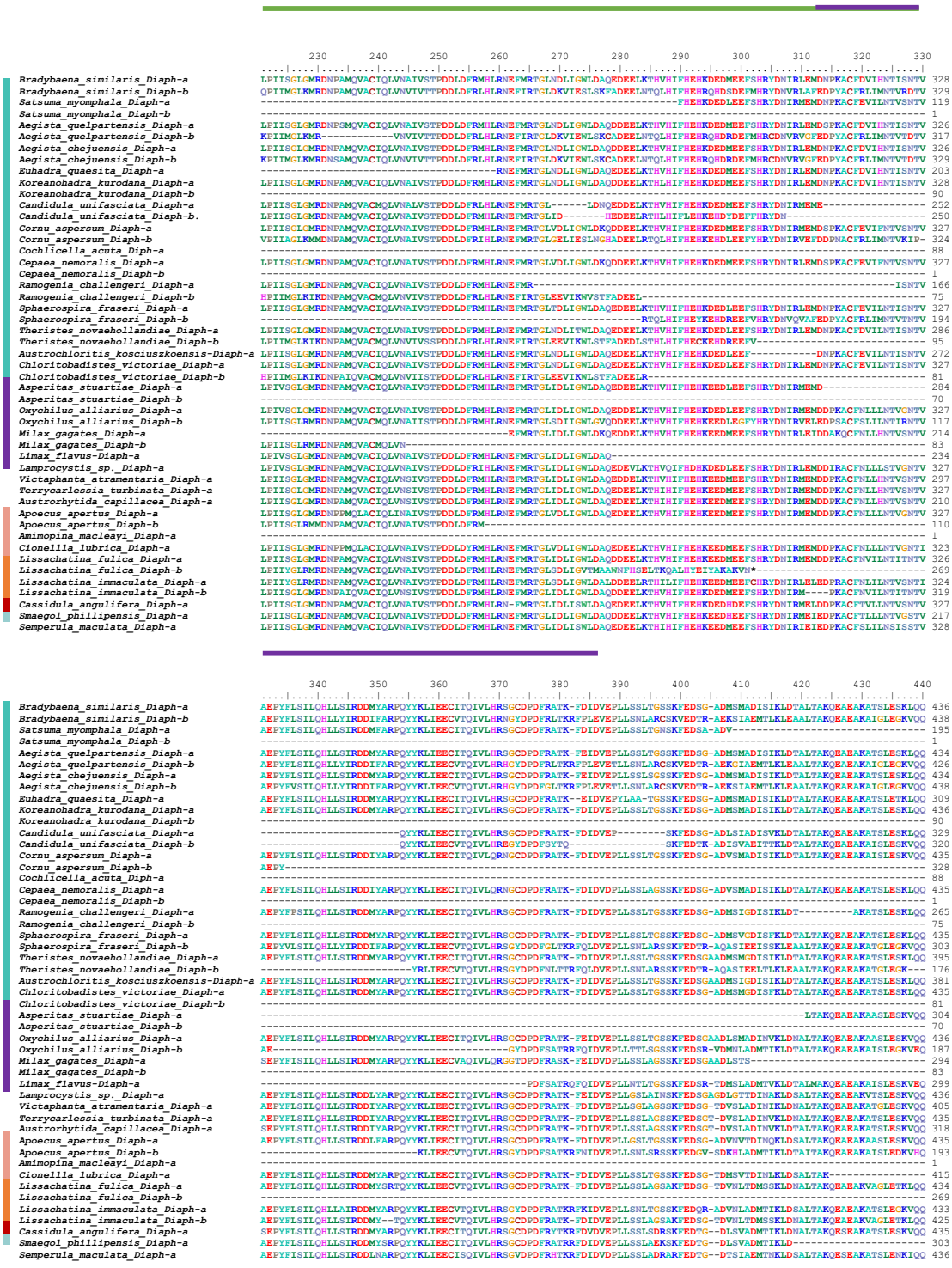
An additional superfamily studied is, Pupilloidea. The derived *Diaph-b* gene was only extracted from *Apoecus apertus*– the ancestral copy is the single copy in the other species.

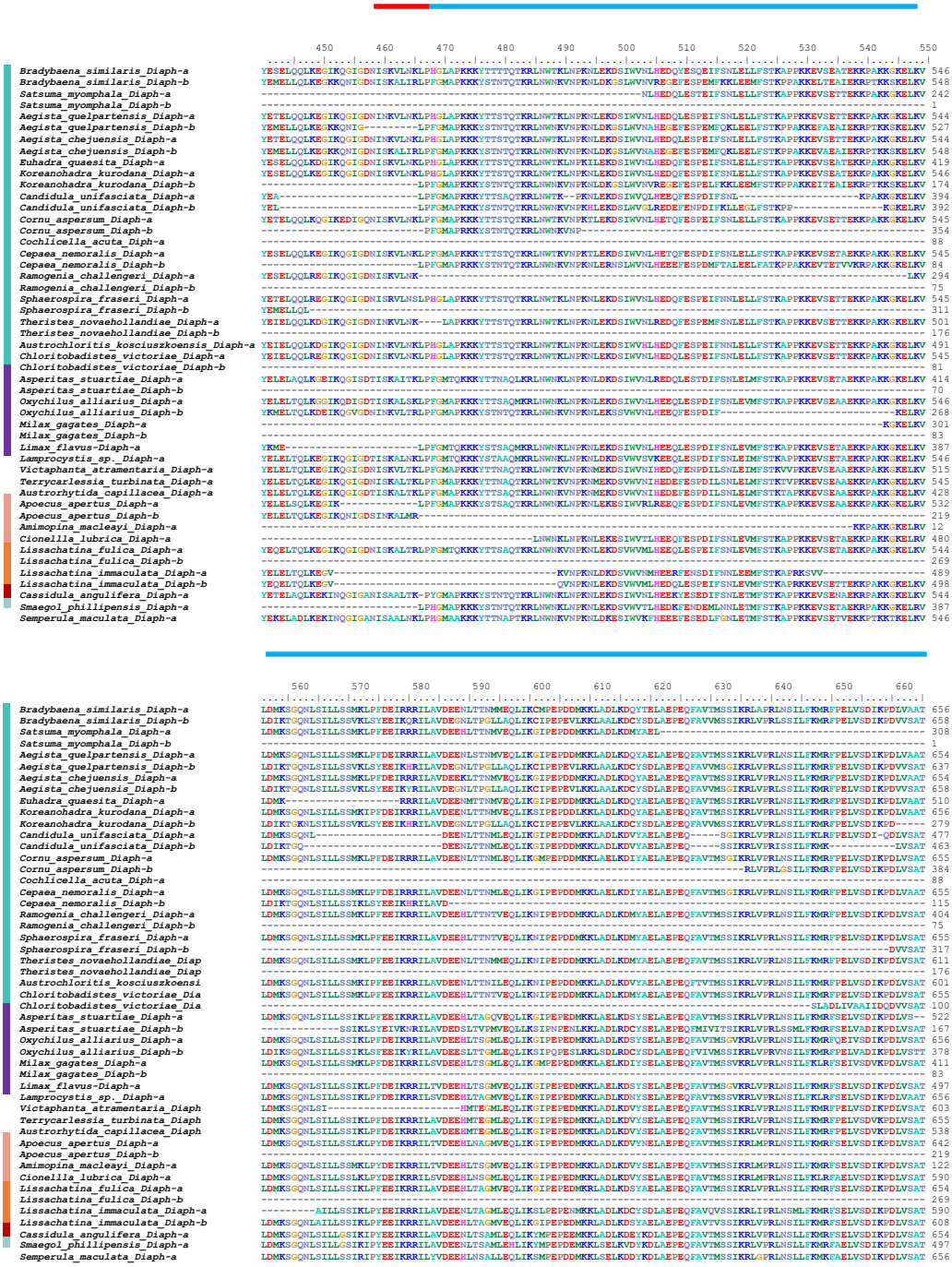
Rhytididae is a family of medium-sized dextral-coiling carnivorous land snails within the superfamily Rhytidoidea, native to Australia and New Zealand. Analysis of transcriptomes of four species within the family revealed only an ancestral Diaph paralog (*diaph-a*).

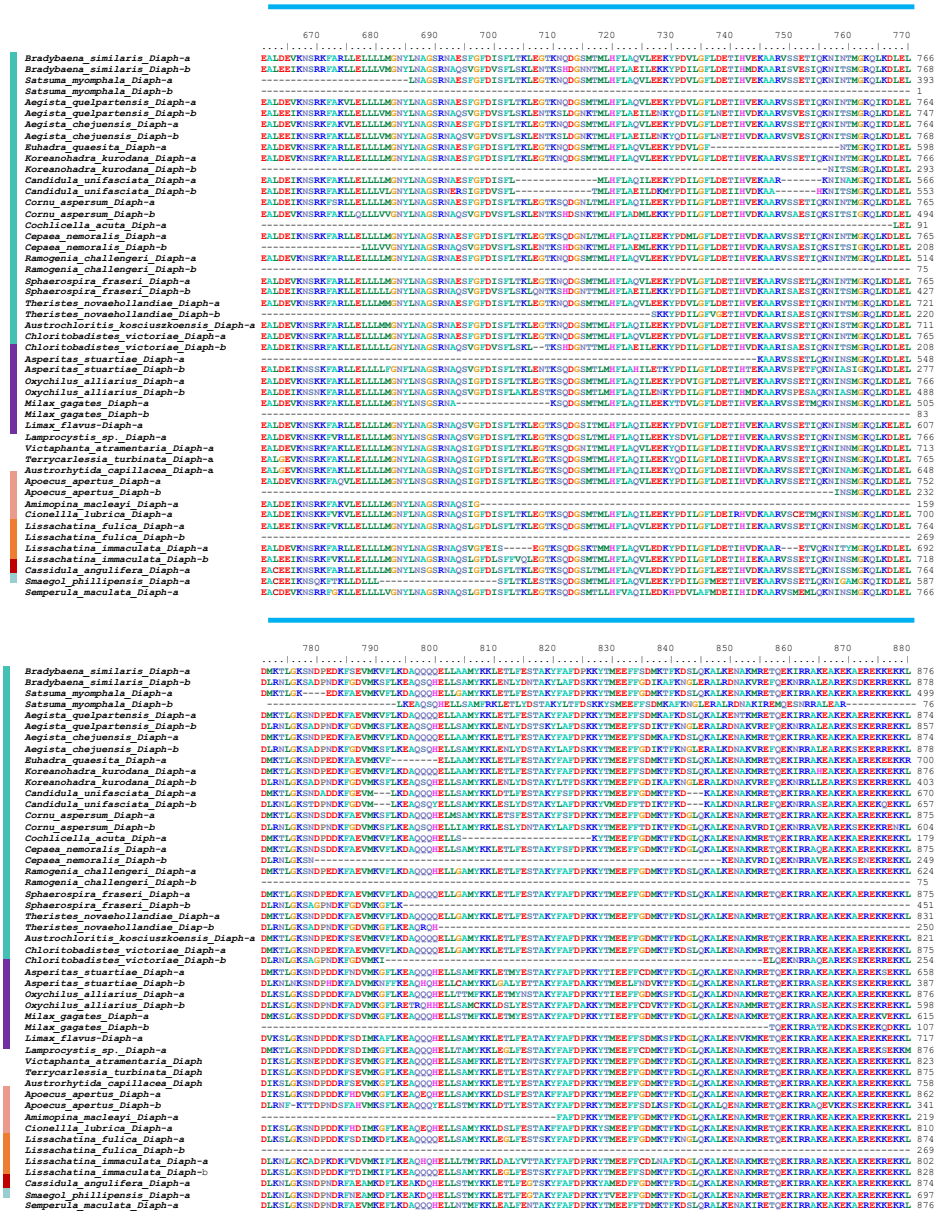
Limacoideia, a term derived from "limacoid", meaning slug–like, is an infraorder of medium–large terrestrial land snails and slugs inhabiting the Palearctic region. Analysis of transcriptomes revealed that three species; the slug *Milax gagantes*; garlic snail *Oxychilus alliarius*, and the snail *Asperitas stuartiae,* express both diaphanous-related formin paralogs. An additional two species, the slug *Limax flavus* and the small land snail *Lamprocytis sp.* expressed a single diaphanous-related formin gene (*Diaph-a*). A gene copy could not be recovered from the transcriptome of the semi-slug *Fastosarion virens.* A single copy of diaph-a was resolved from the three non-stylommatophoran species, *Semperula maculata*, *Cassidula angulifera* and *Smaegol phillipensis.*

A


10 20 30 40 50 60 70 80 90 100 110
Bradybaena_similaris_Diaph-a
Bradybaena_similaris_Diaph-b
Satsuma_myomphala_Diaph-a
Satsuma_myomphala_Diaph-b
Aegista_quelpartensis_Diaph-a
Aegista_quelpartensis_Diaph-b
Aegista_chejuensis_Diaph-a
Aegista_chejuensis_Diaph-b
Euhadra_quaesita_Diaph-a
Koreanohadra_kurodana_Diaph-a
Koreanohadra_kurodana_Diaph-b
Candidula_unifasciata_Diaph-a
Candidula_unifasciata_Diaph-b
Cornu_aspersum_Diaph-a
Cornu_aspersum_Diaph-b
Cochlicella_acuta_Diph-a
Cepaea_nemoralis_Diaph-a
Cepaea_nemoralis_Diaph-b
Ramogenia_challengeri_Diaph-a
Ramogenia_challengeri_Diaph-b
Sphaerospira_fraseri_Diaph-a
Sphaerospira_fraseri_Diaph-b
Theristes_novaehollandiae_Diaph-a
Theristes_novaehollandiae_Diaph-b
Austrochloritis_kosciusskoensis-Diaph-a
Chloritobadistes_victoriae_Diaph-a
Chloritobadistes_victoriae_Diaph-b
Asperitas_stuartiae_Diaph-a
Asperitas_stuartiae_Diaph-b
Oxychilus_alliarius_Diaph-a
Oxychilus_alliarius_Diaph-b
Milax_gagates_Diaph-a
Milax_gagates_Diaph-b
Limax_flavus-Diaph-a
Lamprocystis_sp._Diaph-a
Victaphanta_atramentaria_Diaph-a
Terrycarlessia_turbinata_Diaph-a
Austrorhytida_capillacea_Diaph-a
Apoecus_apertus_Diaph-a
Apoecus_apertus_Diaph-b
Amimopina_macleayi_Diaph-a
Cionella_lubrica_Diaph-a
Lissachatina_fulica_Diaph-a
Lissachatina_fulica_Diaph-b
Lissachatina_immaculata_Diaph-a
Lissachatina_immaculata_Diaph-b
Cassidula_angulifera_Diaph-a
Smaegol_phillipensis_Diaph-a
Semperula_maculata_Diaph-a
120 130 140 150 160 170 180 190 200 210 220
Bradybaena_similaris_Diaph-a
Bradybaena_similaris_Diaph-b
Satsuma_myomphala_Diaph-a
Satsuma_myomphala_Diaph-b
Aegista_quelpartensis_Diaph-a
Aegista_quelpartensis_Diaph-b
Aegista_chejuensis_Diaph-a
Aegista_chejuensis_Diaph-b
Euhadra_quaesita_Diaph-a
Koreanohadra_kurodana_Diaph-a
Koreanohadra_kurodana_Diaph-b
Candidula_unifasciata_Diaph-a
Candidula_unifasciata_Diaph-b
Cornu_aspersum_Diaph-a
Cornu_aspersum_Diaph-b
Cochlicella_acuta_Diph-a
Cepaea_nemoralis_Diaph-a
Cepaea_nemoralis_Diaph-b
Ramogenia_challengeri_Diaph-a
Ramogenia_challengeri_Diaph-b
Sphaerospira_fraseri_Diaph-a
Sphaerospira_fraseri_Diaph-b
Theristes_novaehollandiae_Diaph-a
Theristes_novaehollandiae_Diaph-b
Austrochloritis_kosciusskoensis-Diaph-a
Chloritobadistes_victoriae_Diaph-a
Chloritobadistes_victoriae_Diaph-b
Asperitas_stuartiae_Diaph-a
Asperitas_stuartiae_Diaph-b
Oxychilus_alliarius_Diaph-a
Oxychilus_alliarius_Diaph-b
Milax_gagates_Diaph-a
Milax_gagates_Diaph-b
Limax_flavus-Diaph-a
Lamprocystis_sp._Diaph-a
Victaphanta_atramentaria_Diaph-a
Terrycarlessia_turbinata_Diaph-a
Austrorhytida_capillacea_Diaph-a
Apoecus_apertus_Diaph-a
Apoecus_apertus_Diaph-b
Amimopina_macleayi_Diaph-a
Cionella_lubrica_Diaph-a
Lissachatina_fulica_Diaph-a
Lissachatina_fulica_Diaph-b
Lissachatina_immaculata_Diaph-a
Lissachatina_immaculata_Diaph-b
Cassidula_angulifera_Diaph-a
Smaegol_phillipensis_Diaph-a
Semperula_maculata_Diaph-a


Bradybaena_similaris_Diapha-a
Bradybaena_similaris_Diapha-b
Satsuma_myomphala_Diapha-a
Satsuma_myomphala_Diapha-b
Aegista_quelpartensis_Diapha-a
Aegista_quelpartensis_Diapha-b
Aegista_chejuensis_Diapha-a
Aegista_chejuensis_Diapha-b
Euhadra_quaesita_Diapha-a
Koreanohadra_kurodana_Diapha-a
Koreanohadra_kurodana_Diapha-b
Candidula_unifasciata_Diapha-a
Candidula_unifasciata_Diapha-b.
Cornu_aspersum_Diapha-a
Cornu_aspersum_Diapha-b
Cochlicella_acuta_Diapha-a
Cepaea_nemoralis_Diapha-a
Cepaea_nemoralis_Diapha-b
Ramogenia_challengeri_Diapha-a
Ramogenia_challengeri_Diapha-b
Sphaerospira_fraseri_Diapha-a
Sphaerospira_fraseri_Diapha-b
Theristes_novaehollandiae_Diapha-a
Theristes_novaehollandiae_Diapha-b
Austrochloritis_kosciuszkoensis-Diapha-a
Chloritobadistes_victoriae_Diapha-a
Chloritobadistes_victoriae_Diapha-b
Asperitas_stuartiae_Diapha-a
Asperitas_stuartiae_Diapha-b
Oxychilus_alliarius_Diapha-a
Oxychilus_alliarius_Diapha-b
Milax_gagates_Diapha-a
Milax_gagates_Diapha-b
Limax_flavus-Diapha-a
Lamprocystis_sp._Diapha-a
Victaphanta_atramentaria_Diapha-a
Terrycarlessia_turbinata_Diapha-a
Austrorhytida_capillacea_Diapha-a
Apoecus_apertus_Diapha-a
Apoecus_apertus_Diapha-b
Amimopina_macleayi_Diapha-a
Cionellla_lubrica_Diapha-a
Lissachatina_fulica_Diapha-a
Lissachatina_fulica_Diapha-b
Lissachatina_immaculata_Diapha-a
Lissachatina_immaculata_Diapha-b
Cassidula_angulifera_Diapha-a
Smaegol_phillipensis_Diapha-a
Semperula_maculata_Diapha-a


Bradybaena_similaris_Diaph-a
Bradybaena_similaris_Diaph-b
Satsuma_myomphala_Diaph-a
Satsuma_myomphala_Diaph-b
Aegista_quelpartensis_Diaph-a
Aegista_quelpartensis_Diaph-b
Aegista_chejuensis_Diaph-a
Aegista_chejuensis_Diaph-b
Euhadra_quaesita_Diaph-a
Koreanohadra_kurodana_Diaph-a
Koreanohadra_kurodana_Diaph-b
Candidula_unifasciata_Diaph-a
Candidula_unifasciata_Diaph-b
Cornu_aspersum_Diaph-a
Cornu_aspersum_Diaph-b
Cochlicella_acuta_Diph-a
Cepaea_nemoralis_Diaph-a
Cepaea_nemoralis_Diaph-b
Ramogenia_challengeri_Diaph-a
Ramogenia_challengeri_Diaph-b
Sphaerospira_fraseri_Diaph-a
Sphaerospira_fraseri_Diaph-b
Theristes_novaehollandiae_Diaph-a
Theristes_novaehollandiae_Diaph-b
Austrochloritis_kosciuszkoensis_Diaph-a
Chloritobadistes_victoriae_Diaph-a
Chloritobadistes_victoriae_Diaph-b
Asperitas_stuartiae_Diaph-a
Asperitas_stuartiae_Diaph-b
Oxychilus_alliarius_Diaph-a
Oxychilus_alliarius_Diaph-b
Milax_gagates_Diaph-a
Milax_gagates_Diaph-b
Limax_flavus-Diaph-a
Lamprocystis_sp._Diaph-a
Victaphanta_atramentaria_Diaph-a
Terrycarlessia_turbinata_Diaph-a
Austrorhytida_capillacea_Diaph-a
Apoecus_apertus_Diaph-a
Apoecus_apertus_Diaph-b
Amimopina_macleayi_Diaph-a
Cionellla_lubrica_Diaph-a
Lissachatina_fulica_Diaph-a
Lissachatina_fulica_Diaph-b
Lissachatina_immaculata_Diaph-a
Lissachatina_immaculata_Diaph-b
Cassidula_angulifera_Diaph-a
Smaegol_phillipensis_Diaph-a
Semperula_maculata_Diaph-a



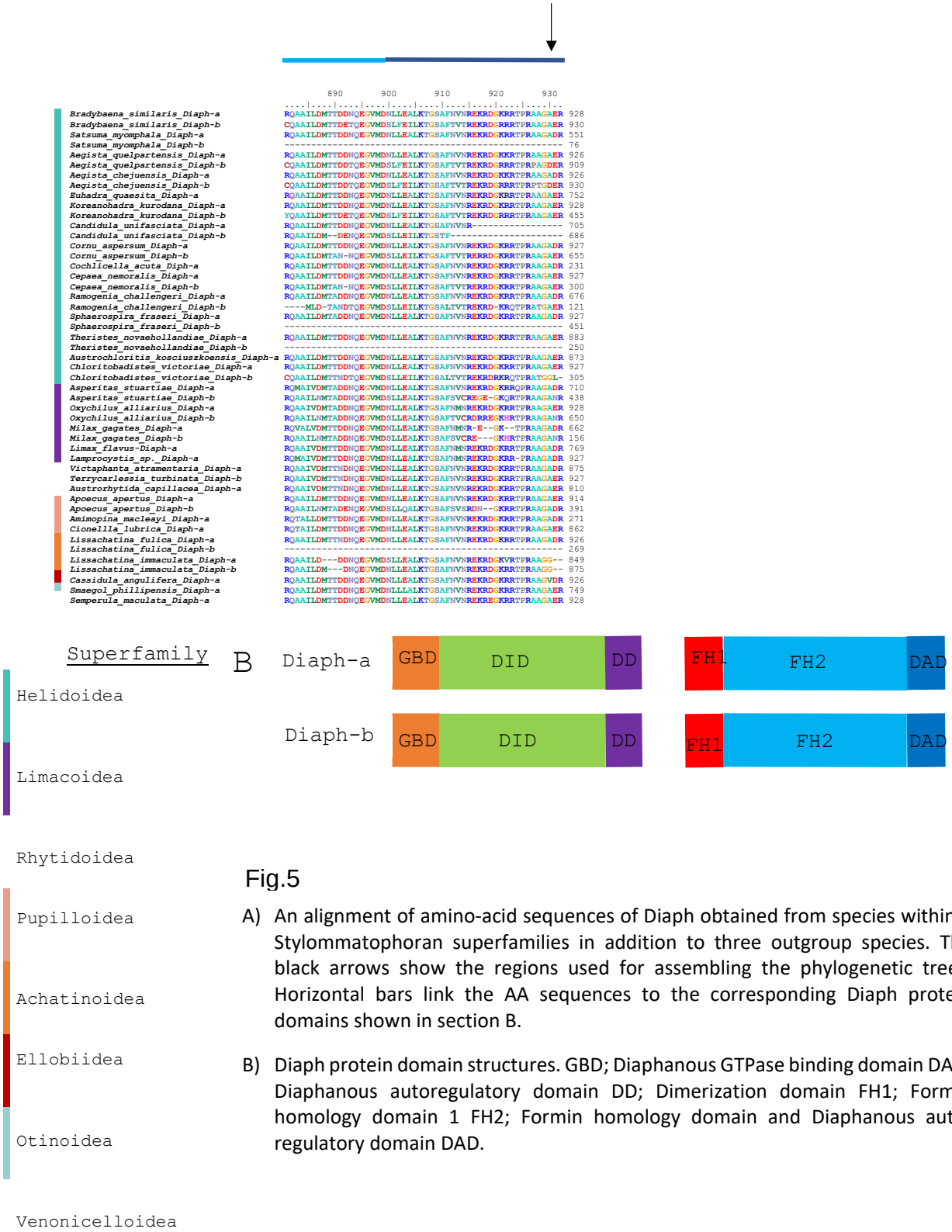


Fig.5

A) An alignment of amino-acid sequences of Diaph obtained from species within 5 Stylommatophoran superfamilies in addition to three outgroup species. The black arrows show the regions used for assembling the phylogenetic trees. Horizontal bars link the AA sequences to the corresponding Diaph protein domains shown in section B.

B) Diaph protein domain structures. GBD; Diaphanous GTPase binding domain DAD; Diaphanous autoregulatory domain DD; Dimerization domain FH1; Formin homology domain 1 FH2; Formin homology domain and Diaphanous auto-regulatory domain DAD.

Molecular Phylogeny of order Stylommatophora

According to the most recent molecular phylogenetic study, Stylommatophora is divided up into three clades: Achatinina, Helicina and Scolodontina. The most profound split is the one between Achatinoidea and the other Stylommatophora – Helicina; however, according to the evolution of diaphanous-related formin across the clade, the defined split is linked to the diaph copies (Fig. 6). A clear separation of the diaph paralogs is fully supported at the base of Stylommatophora indicating the duplication occurred in an ancestor of the clade.

The *diaph-a* clade splits into two groups; the Helicoidea superfamily forms a monophyletic group, away from the remaining members of Stylommatophora. Three non-stylommatophoran species from two separate Orders, Ellobidae and Systellommatophora, act as outgroups, with the slug *Semperula maculata* rooting the tree.

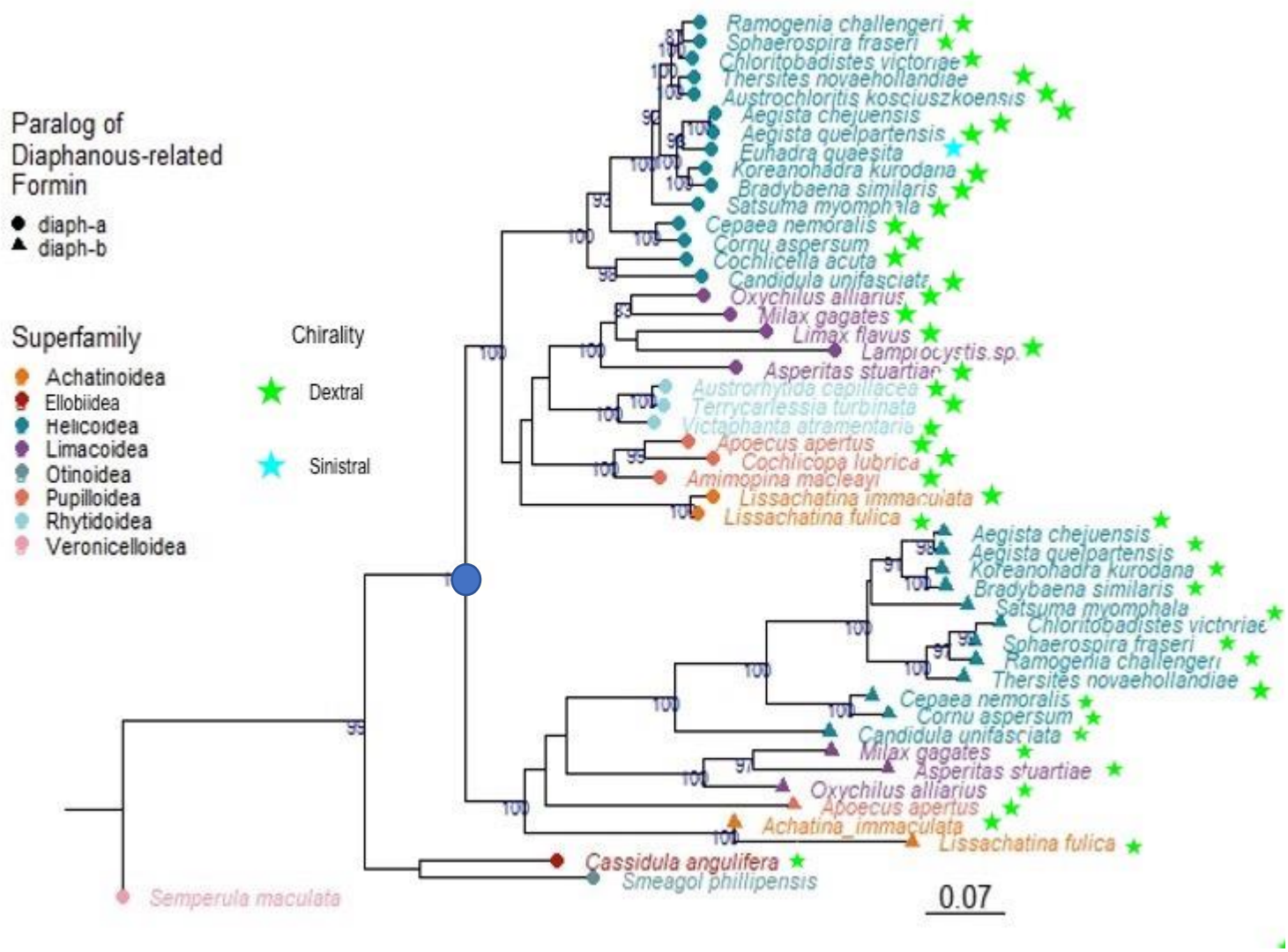


Figure 6. A Phylogenetic tree was constructed using diaphanous-related formin gene paralogs (diaph-a, b) from 28 species belonging to the following stylommatophoran families: Helicoidae, Limacoidea, Rhytidoidea, Pupilloidea, and Achatinoidea. Species from the superfamily Ellobiiea –*Cassidula angulifera*, Otinoidea –*Smeagol phillipensis* and Venonicelloidea –Semperula maculata were included as outgroups. The phylogeny is based on 2850 nucleotide sites, and genetic distances are estimated using the best-fit model – TIM + F + I + G4, including rate variation between sites. Bootstrap values are based on 1000 replicates, pinpoint support for individual branches. The scale bar represents 3 changes per 100 nucleotides. The tree is rooted on *Semperula maculata*. The proposed gene duplication event at the base of Stylommatophora is indicated by the blue dot.

1.3 Discussion

Bilateria, are externally symmetric yet display consistent internal asymmetries in organ patterning and arrangement. Elements of the mechanism controlling these asymmetries are well understood; however, the genes controlling the initial symmetry-breaking process are yet to be defined (Levin, Klar and Ramsdell, 2016). Snails are a crucial model for this research. The direction of the coiled shell and body, defined at the 4-8 cell stage by the action of a single locus, is variable (Okumura *et al.*, 2008; Davison, 2020).

Duplication of actin nucleator diaphanous-related formin (diaph) has been linked to the evolution of chirality in species within two pulmonate groups, Hygrophila (Davison *et al.*, 2016; Kuroda *et al.*, 2016) and Stylommatophora (Noda, Satoh and Asami, 2019). The slow pace of molluscan genomics has limited the number of species studied.

Genomic resources for Panpulmonata

We found a total of 52 species from the Panpulmonata clades: Stylommatophora (29), Hygrophila(17), and Systellommatophora (3) with assembled genomes or transcriptomes. Eighty per cent of the assemblies have been published in the last seven years. Three of those published in the last two years were of chromosomal level. This ties in with the trend that improving sequencing technology is leading to a rapid increase in genome assemblies. However, whilst we report an upward trajectory with regard to molluscan genetics, there is still a long way to go. Gastropods are one of the most diverse groups on Earth with more than 62'000 documented species, yet we only have chromosomal level assemblies for 79. Furthermore, none of the genome assemblies are of snails, which naturally vary in their chirality.

Origin of the Duplication Events

We resolved both *diaph* homologs in three out of the four Lymnaeidae genera. The copies lie in tandem on the same contig within the genome of *Radix auricularis*. A single copy was resolved from each species within Planorbidae, in addition to the lone representative of Physidae (*Physella acuta*). A single copy of the ancestral gene was also resolved in the species of dextral Acroloxidae.

The phylogeny indicates two potential locations for the tandem-duplication event of *diaph* in Hygrophila. Either a tandem-duplication event occurred after the divergence of the Galba genus, or tandem duplication occurred in an ancestor of Lymnaeidae with insufficient transcriptomic depth accounting for the lack of derived paralog resolution across the three *Galba* species.

We suggest the latter being the most plausible location for the duplication event for the following reasons. The only two transcriptomes to resolve both copies, *Pseudosuccinea columella* (Teasdale *et al.*, 2016) and *Lymnaea stagnalis* (Seppälä *et al.*, 2021) are over 200mb, at least twice as large as other transcriptome sampled. This indicates that the Ldia2 transcript is present at a far lower level than the other paralog in adult tissue. Asymmetric gene expression biasing the ancestral gene is commonplace after duplication events (Conant and Wagner, 2003; Guschanski *et al.,* 2017). Davison et al. (2016) reported high levels of Ldia2 expression in embryonic Lymnaea stagnalis, which may suggest that the expression levels tail off as the snail reaches adulthood. Finally, as evidenced in the alignment and subsequent phylogeny, the quality of the assembly of the Ldia1 genes is poor. This firmly suggests the transcriptomes are too shallow for resolving the mRNA transcript of Ldia2.

We identified both paralogs of diaph from at least one member of four of the five Stylommatophora superfamilies studied. A single copy of the ancestral paralog *diaph-a* was resolved from an outgroup member – Systellommatophora – studied. Only the

ancestral *diaph-a* paralog was identified in the three species of Rhytidoidea in addition to two members of Limacoidea (*Limax flavus*, *Lamprocytis sp) a*nd three members of the helicoidea family studied (*Cochlicella acuta, Euhadra quaesita and Austrochloritis kosciuszkoensis).* RNA-seq experiments on the Helicoidea species *Bradybaena* revealed that the expression of *diaph-b* is significantly reduced after the embryonic stage (Noda, Satoh and Asami, 2019). Therefore, the absence of the *diaph-b* paralog could be explained by the tissue type and modest transcriptome size.

The molecular phylogeny of the diaphanous-related formin genes in Stylommatophora supports the theory proposed by Noda and colleagues that the gene duplication observed in *Bradybaena similaris* resulted from a single ancient gene duplication event at the base of Stylommatophora after the split from Hygrophila (Noda, Satoh and Asami, 2019a). The discovery of the gene paralogs (*diaph-a* and *diaph-b*) on orthologous chromosomes in both Lissachatina species supports ancient WGD in an ancestor of Stylommatophora (Liu *et al.*, 2021; Chen *et al.*, 2022).

Our results reveal that diaph has undergone two distinct duplications during the evolution of the pan-pulmonata, a single gene tandem-duplication in Lymnaeaid ancestor, –Hygrophila and as part of a WGD in a stylommatophoran ancestor.

Association between duplication of diaphanous related formin gene and the evolution of sinistrality

The clade of freshwater snails and Limpets, Hygrophila, comprises dextral and sinistral families. How the sinistral species evolved has been the subject of debate. A phylogenetic study of the Hygrophilan clade by Saadi, Wade and Davison (2020) highlighted two parsimonious hypotheses for the evolution of sinistrality. Either sinistrality evolved from a dextral ancestor on two separate occasions, in Planorbidae and Physidae, or there is a single origin with a reversal to dextrality occurring in an ancestor of Lymnaeaidea (Davison *et al.*, 2016; Ahmed J Saadi, Davison and Wade, 2020). Molecular evidence from two research groups led by Davison and Kuroda (2016; 2019) revealed that Ldia2 was required to establish dextrality, which in addition

to the absence of the duplication in sinistral species studied, suggested a single origin of sinistrality. The tandem duplication of *diaph* facilitates the reversal of dextrality.

Our results support the conclusion made by Davison et al.l, (2016) that the tandem duplication diaphanous-related formin occurred in an ancestor of Lymnaeid. Th, is combined with the molecular evidence does suggest that the duplication permitted reversal from sinistrality in the clade. However, pathology and mixed broods are associated with the loss of one copy.

A duplication of the gene coding for diaphanous-related formin was linked to determining chirality in stylommatophoran species *Bradybaeanae similaris*. Expression levels of the derived copy, *diaph-b,* were reduced in embryos of the racemic mutants, compared to the wild-type. Noda, Satoh and Asami (2019) theorized that the varying numbers of sinistral offspring could be a result of mutations within the diaph-b enhancer. Our research indicated that gene duplication was common to all members of Stylommatophora, partially supporting the hypothesis. *Euhadra quaesita* is a sinistral species within the chiral variable genus *Euhadra*—the transcriptome generated from immature snail tissue resolved only the ancestral paralog. The presence of two copies in the other dextral *Bradybaenidae*, *Aegista chejuensis and Aegista quelpartensis,* could indicate that the sinistral phenotype in *E.quaesita* is linked to the downregulation of *diaph-b*. An alternative explanation is that there is a lower level of diaph-b expression in the tissue type used for the euhadra transcriptome, preventing diaph-b resolution. The transcriptome was of mantle tissue as opposed to the entire snail.

If the diaph-b is downregulated in Euhadra quaesita it sparks the question if the duplication is what permits chiral reversal, why is sinistrality so rare? One explanation is there is a clear between the mutation and pathology. A proportion of offspring produced by a mother with the mutant racemic all fail to hatch (Noda et al., 2019)., This combined with the findings in pond snails and other model organisms, suggests that other gene pairs are produced through the duplication involved in chiral determination.

Whole genome duplications are key sources of evolutionary and genetic novelty. Duplicate genes can accumulate mutations leading to novel phenotypes without leading to pathology. In Stylommatophora, the WGD is linked to the transition from aquatic to terrestrial life. The expansion provided additional genes linked to muscle, nerve and kidney development. It is likely that the regulation of multiple gene pairs resulted in the evolution of chirality in Stylommatophoran species.

1.4. Conclusion

In conclusion, the study confirms the hypothesis of Noda and colleagues (2019) that diaphanous-related formin underwent two distinct duplications during the evolution of the Pan-Pulmonata. In addition, a potential link between the regulation of *diaph-b* expression and sinistral coiling in the variable land snail genus *Euhadra* was revealed.

However, additional evidence gathered from single cells (Tee *et al.*, 2021), fruit flies (Chougule *et al.*, 2020), and nematode (McDowell *et al.*, 2016) indicate that several different cytoskeletal proteins likely determine chirality. For instance, gut coiling in fruitflies requires formin DAAM in addition to MyosinD (Chougule *et al.*, 2020). The evidence for the functionality of diaph in chiral determination is based on species, where variation is linked to pathology (Davison, 2020).

The next step would be evaluating the genes responsible for stable chiral determination in the helicoid land snail genus *Euhadra*. The initial stage would be to generate a highly contiguous reference genome. A template protocol can be found in the research paper documenting the assembly of the Spanish slug *A.vulgaris.* The team used

Chapter 2:

Evaluation of *de novo* RAD-seq assemblers for investigating population structure of land snail genus *Euhadra.*

2.1 Abstract

Snails display outward asymmetry by way of their coiled shell, the direction of which is dictated by a single maternally inherited locus. Although, the majority of snails are right-handed, however rare mirror-image morphs persist at all taxonomic levels. The evolution of mirror-image morphs is a curiosity and provides a great opportunity for studying single gene speciation as inter-chiral mating was deemed impossible due to genital location. *De novo* RAD-Seq analysis has since revealed ongoing gene flow between dextral and sinistral species of Japanese *Euhadra* suggesting that more than a single gene is required to promote speciation. In addition, one dextral individual clustered with sinistral individuals across all analyses. This pattern may suggest that the genetic divergence between these Individuals is the locus controlling chirality. However, the placement of the individual may also be due to errors associated with denovo RAD seq assembly. Multiple pipelines are used if a reference is unavailable to ensure the validity of inferences regarding population structure. We used two pipelines designed for De novo analysis to analyse the *Euhadra* RAD-seq data confirm the location of the anomalous individual. Our results showed confirmed the strong regional divergence and confirmed the placement of the dextral individual sinistral species. This is a key finding in the chase for the loci determining chiral variation.

2.2 Introduction

Chiral, from the Greek word χέρι meaning hand, defines an object which cannot be super-imposed onto its mirror image. An object and its mirror image are referred to as enantiomers. The left and right human hands exemplify a pair of enantiomers. Every time bilateral symmetry is broken, each enantiomer has a seemingly equal chance of being generated. However, in nature, one enantiomer is chosen preferentially over the other (Wood, 2005). This phenomenon of handedness bias is reflected across all levels. On a molecular level, all amino acids are left-handed, yet the double helix of DNA is right-handed. This bias is translated into gross anatomy. Animals are externally symmetric yet exhibit internal L/R asymmetry in their organ placement and patterning. The animal body plan can take one of two forms depending on the polarity of the L/R axis; however, all individuals of a species will exhibit the same handedness. There is, however, one exception to the rule. Snails are the only group to vary in their asymmetry. The direction of the shell and internal body coiling is evident by the eight-cell stage and is determined by a single maternally inherited locus (Boycott and Diver, 1923; Sturtevant, 1923).

Research on the pond snail *Lymnaea stagnalis*- yielded a potential identity for the chiral locus, a pair of tandemly duplicated diaphanous-related formin genes (*Ldia1, Ldia2*). Sinistral coiling phenotypes in this ordinarily dextral species have been linked to a frameshift mutation within the derived paralog (*Ldia2*) (Davison *et al.*, 2016; Kuroda *et al.*, 2016; Abe and Kuroda, 2019). The discovery of a single paralog, homologous to Ldia1 in sinistral species, confirmed the link and suggested the tandem duplication facilitated a reversal to dextrality in the clade (Davison *et al.*, 2016). Diaphanous-related duplication has been linked with chiral variation in the primarily dextral stylommatophoran species *Bradybaenae similaris* (Noda, Satoh and Asami, 2019). Embryos of racemic mutants, which could produce broods of sinistral and dextral offspring, exhibited reduced expression of the *diaph-b* (derived paralog)

compared to the wild-type. In both species' chiral reversal is associated with pathology (Davison, 2020).

Dextral (right-handed) coiling snails are the most common; however, L/R reversal has frequently evolved, resulting in sinistral (left-handed) snails occurring across all taxonomic levels (Schilthuizen and Davison, 2005). An allele for reversal must become fixed in a population for a reversed chiral morph to be established. The evolutionary mechanisms by which chiral reversal alleles are driven to fixation, despite strong positive-frequency dependent selection, are gradually being uncovered (Johnson, 1982; Schilthuizen *et al.*, 2007; Danaisawadi *et al.*, 2016; Richards *et al.*, 2017a).

An essential factor impacting chiral reversal is the ability to promote gene flow through successfully inter-chiral mating. The probability of inter-chiral mating is greatly influenced by shell morphology and mating behaviour (Cain and Sheppard, 1977). Snails with high-spired shells prefer to mate using a non-reciprocal shell mounting technique. The shells' alignment counteracts the chiral variation, permitting successful inter-chiral mating (Schilthuizen, 2003). Prominent examples of high-spired snails are the brightly coloured tree snail *Amphidromus inversus* and the sinistral snail *Partula sulturalis. Amphidromus sensu stricto*, the sub-genus of brightly coloured South Asian tree snails, has, over the years, drawn much attention from evolutionary biologists as it remains a rare example of persistent di-morphism. Within populations sinistral and dextral morphs occur at approximately equal frequencies.

Schilthuizen and colleagues (2007) revealed that in *Amphidromus*, there is a high degree of sexual selection for the opposite morph, which acts to select for a dimorphism. The now-extinct Polynesian tree snail *Partula sulturalis* is another example of persistent chiral dimorphism. Narrow clines existed where sinistral and dextral morphs of the primarily sinistral species lived alongside one another. The presence of dextral morphs alongside other sinistral species of the same genus indicates reproductive character displacement is supporting the persistence of rare morphs in these zones (Johnson, 1982).

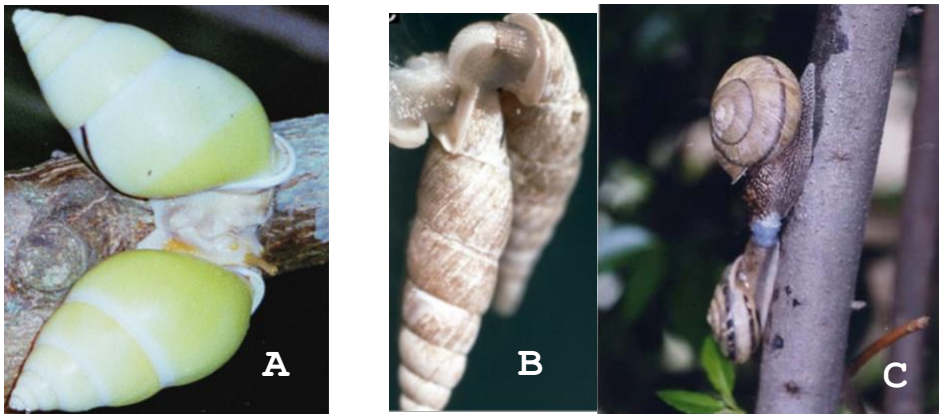


Fig 1. Copulating strategies of three pulomanate snail speies

A. Recirpcoal, face to face mating of two high spired Amphidromus atricallosus (Sutcharit et al., 2007)
B. Non-reciprocal copulation by shell mounting of sinistral clausiliids credit: T Asami
C. Face to face mating between sinistral *Euhadra quaesita* and dextral *Euhadra peliomphala*, credit: Kentaro Nakao & Seiichi Takase

On the other hand, reciprocal, face-to-face mating is performed by snails with flat, globular shells. The relative placement of the genitalia prevents inter-chiral mating without additional adjustment (Asami, Cowie and Ohbayashi, 1998; Richards *et al.*, 2017). If inter-chiral mating is impossible or incredibly difficult, how has a fixation of chiral reversal occurred in flat-shelled species?

A predator–prey interaction has been implicated in the frequent evolution to sinistrality in the dextral flat-shelled genus *Satsuma*. The jaws of the snail eating snake *Pareas iwasakii* display consistent L/R asymmetry, a specialization for grasping the soft part of snails. The snake morphology and predation behaviour show a preference for dextral snails. Sinistral morphs are more likely to evade capture (Danaisawadi *et al.*, 2016).

An additional genus where fixation for the opposite coiling morph has repeatedly occurred is the camaenid *Euhadra*. The genus is native to Japan and is made up of 22 species, five sinistral and 17 dextral. The flat-shells and face-to-face mating

technique, coupled with chiral mating rituals, had led to the belief that inter-chiral mating was impossible. Therefore, speciation could occur upon reversal.

Various methods have been employed to test chiral reversal in *Euhadra* as a candidate system for single-gene speciation. Mathematical models suggested that certain factors, including reproductive character displacement, and predation the rare chiral morph, could become fixed in a population, potentially resulting in a new species (Gittenberger, 1988; Orr, 1991). Single-gene mitochondrial phylogenies (Ueshima and Asami ,2003) revealed that chiral reversal resulting in speciation had occurred multiple times in *Euhadra*.

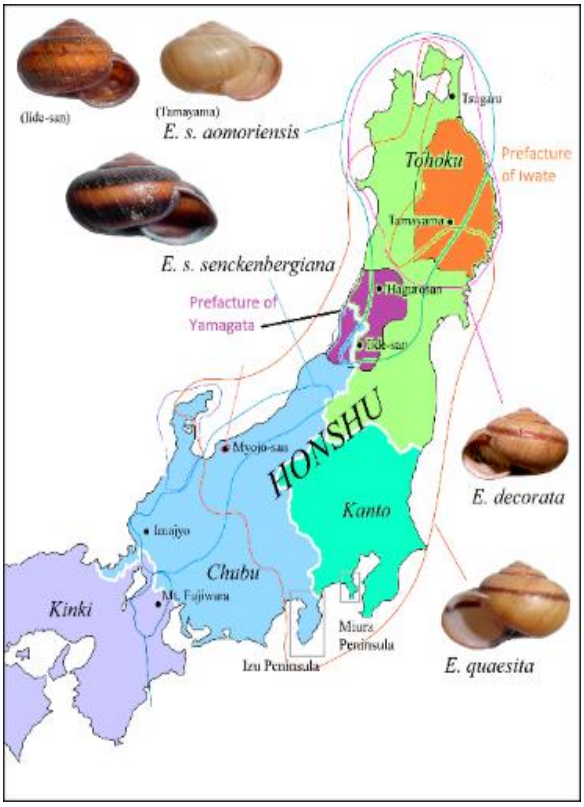


Figure 1. Visual of the Japanese island of Honshu, with the regions of Iwate and Yamagata which in which the inter-chiral contact zones were found, highlighted in Orange and Purple. Distribution of species sampled from each region are labelled. *(Adapted from* (Davison *et al.*, 2005)

However, evidence from a theoretical model from Davison et al. (2005) showed that gene-flow occurred between chiral morphs as a result of maternal delayed inheritance, further, countering the of speciation by chiral reversal induced reproductive isolation. However, additional population-level genetic data is required to confirm the potential for single-gene speciation, as distinguishing between low levels of gene flow, introgressive hybridization and speciation is difficult using mtDNA alone (Taylor, Anderson and Friesen, 2013; Richards *et al.*, 2017).

To evaluate if speciation as a direct result of chiral reversal has occurred in *Euhadra*, Richards et al. (2017) used restriction site-associated DNA sequencing (RAD-seq) to gather empirical evidence about the potential for gene flow between chiral morphs. To evaluate the effect of geographic and chiral isolation, sinistral *E. quaesita* and dextral E. *aomoriensis* living in sympatry were sampled from two geographically isolated regions of Japan, Iwate and Yamagata (Fig 2.). Sinistral *E. decorata* and dextral *E. senckenbergiana*, often found in sympatry with the *E. aomoriensis* and *E. quaesita,* were added for comparison.

Analysis of genome-wide SNP markers revealed that there had been recent and ongoing gene flow between the two chiral morphs within each region. The separation of *E. quaesita* and *E. aomoriensis* from the other species, sinistral *E. decorata* and dextral *E. senckenbergiana,* even suggests they could be the same species. In addition, there is a significant genetic distinction between the sinistral and dextral individuals from Yamagata and those from Iwate, indicating that geography plays a more substantial role in establishing genetic isolation than chiral reversal.

In addition to ruling out the land snail genus *Euhadra* as a candidate for single gene speciation, the research also revealed a clue for identifying the single locus. The phylogenetic analysis revealed members of the dextral species *E. aomoriensis,* which clustered with members of the sinistral species *E. quaesita.* The individual in question labelled E102_4, would presumably be genetically similar to the *E. quaesita*, apart from the locus determining chirality.

However, the individual's position may also be attributed to bias introduced through the de novo RAD locus assembly and SNP calling. It is common practice to re-affirm population structure information by using different types of de novo pipelines (Shafer et al., 2017).

This research aims to re-analyze the RAD-Seq data using two comparatively different pipelines for de-novo RAD-Seq analysis. We chose Stacks2 (Rochette, Rivera-Colón and Catchen, 2019, p. 2) as it is the updated pipeline version of Richards et al. (2017a). Ipyrad (Eaton and Overcast, 2020) is a Python-based tool offering a simple workflow and built-in visualization tools. The recent publication of the genetic resources, including the transcriptome of *Euhadra quaesita* (Shimizu *et al.*, 2019) and genome of *Cepaea nemoralis* (Saenko *et al.*, 2021) offered the opportunity to perform referenced-based variant calling using the *Euhadra* RAD-seq data set. Incorporating a reference genome avoids the bias introduced when assembling SNPs *de novo* (Galla *et al.*, 2018).

2.2. Methods

Quality Filtering

The RAD-Seq library containing barcoded single-ended reads of 16 *Euhadra* individuals generated by Richards et. al (2017) was processed using Trimmomatic v0.32 (Bolger, Lohse and Usadel, 2014), removing Illumina tru-seq adaptors, low-quality bases and reads of less than 50 bases. FastQC v0.11.8 was subsequently used to establish whether the trimming was successful.

De novo loci assembly and variant calling

RAD-Seq analysis involves assembling trimmed demultiplexed reads into putative alleles and loci. Loci are subsequently aligned between individuals with single-nucleotide polymorphisms called within homologous loci. Analysis of the SNP variation permits the inference of population structure and phylogenetics.

Stacks2 consists of a four-step pathway implemented using the command line either as individual steps or using a Perl wrapper (denovo_map.pl). Initially, process_radtags trims reads to equal length before demultiplexing, a degree of nucleotide mismatch permitted to account for the same barcode. Ustacks assembles reads into stacks which are merged into loci– a specified number of nucleotide mismatches are permitted between stacks, and required stack read depth. The module cstacks acts to merge loci across individuals to form a catalogue. Within Sstacks the variation between individuals is recovered as loci from each individual are matched to the cstack catalogue. Using population data given in the form of a tsv file, the module *Populations* indexes the variation within and between populations (Chirality, Species, Location).

Ipyrad is a user-friendly, API-compatible complete pipeline with built-in Python functions for implementing common phylogenetic algorithms (Eaton and Overcast, 2020b). Implementation of the seven-step modular pipeline through Jupyter-notebooks improves reproducibility and enables parallelization across multiple cores,

decreasing the analytical time taken on a local machine or cloud server. The seven steps include: 1) demultiplexing the RAD-seq library using given barcodes and a given degree of mismatch permitted per barcode; 2) read filtering and trimming; 3) clustering of reads with over 85% similarity; 4) calculation of sequencing error and heterozygosity; 5) generation of consensus allele sequences and filtered by several underdetermined sites per locus; 6) Clustering of aligned consensus sequences across individuals; 7) filtering and outputting of alignments.

The filtered RAD-seq libraries were assembled using Ipyrad (default parameters for steps 2-7) and separately with stacks2 (default parameters) cited in (Richards *et al.*, 2017) used for comparison. We altered the barcode-mismatch parameter to assess whether the location of the *E. aomoriensis* individual E102_4 is due to an erroneous barcode, Reviewing the even number of reads corresponding to each individual, and initial phylogenetic analysis, a parameter of 1 base mismatch per barcode was implemented in our analysis. All the analysis was carried out on a CPU-powered Linux server.

SNP Filtering

Raw SNP panels generated from Stacks2 and Ipyrad pipelines were filtered using vcftools (v0.1.16). The same parameters were used for both datasets to ensure consistency, enabling direct pipeline comparison after phylogenetic analysis (Danecek *et al.*, 2011; Casanova *et al.*, 2021). 6432 and 6990 SNPs resulted after filtering out loci/alleles of insufficient depth and excess missing data (>5%). After the removal of singletons, datasets containing 2946 bi-allelic SNPs (Stacks) and 3420 bi-allelic SNPs (Ipyrad) remained. Filtered SNPs were converted into concatenated matrices using Python script vcf2phylip (v2.6) –a requirement for investigating evolutionary relationships between individuals (Ortiz, 2019).

Phylogenomic Analysis

One of the selling points of Ipyrad is the built-in functions for visualizing relationships between species/individuals within populations and between populations. Functions have been written parsing VCF files of SNPs generated by RAD sequencing to phylogenetic tools, ie. Treemix, Structure, RaxML and subsequently to the python visualization packages Toytree and Toyplot, permitting phylogenetic analysis without knowledge of scripting. However, the rigidity of the functions resulted in the decision to re-run the analysis for comparison with Richards et al., 2017 using packages written in R version 3.6.3 and 4.1.0. With the exception of ML tree generation, all analyses required VCFR (v.1.13.0) for the importation of VCF files and ADEGENET (v.2.13) (Jombart, 2008; Jombart and Ahmed, 2011) for conversion to mutable genlight object permitting the addition of population data.

ML trees were built using concatenated SNP matrices imported into IQTREE (Minh *et al.*, 2020). The transversion model (TVM) was used to calculate the rate of nucleotide substitution and subsequent estimation of evolutionary distance between individuals (Posada and Crandall, 2001; Kalyaanamoorthy *et al.*, 2017). Trees were visualized using ggTree (Dray and Dufour, 2007, p. 4) and associated packages (Yu *et al.*, 2017). Although phylogenies represent an insight into the evolutionary relationships conflicting signals due to degrees of recombination and hybridization within a population, minimum spanning networks were constructed using Poppr (v.2.92) (Kamvar, Tabima and Grünwald, 2014; Knaus and Grünwald, 2016, 2017) to visualize the relationships between individuals, within four populations: Iwate Sinistral, Iwate Dextral, Yamagata Sinistral and Yamagata Dextral, in addition to the outgroups *E.decorata* and *E.senckenbergiana* from Yamagata. Genetic distances are calculated between multi-locus genotypes (individuals), represented by edges on the chart. Minimum genetic distance dictates the connections between the nodes with the multiple connections permitted if equally genetically distant.

Principle component analyses exclusively of *E. quaesita* and *E. aomoriensis* to resolve relationships between individuals were executed using ADEGENET (Jombart and Ahmed, 2011) and ADE4 (v.1.7) (Dray and Dufour, 2007) with ggplot2 utilized for

plotting. Genetic loci contributing to the variation were extracted from single PCA (ADEGENET) and gene pop files in R.

Generation of SNP markers from Whole Genome Sequencing

To evaluate where the phylogenetic patterns can be seen at a whole genome level or a result of bias induced by *de novo* assembly methods from reduced representation data, an attempt was made to align whole genome sequences from corresponding individuals to an assembled reference before running a reference-based variant calling pipeline As a reference genome for either *Euhadra quaesita* or *Euhadra aomoriensis* is yet to be published, a reduced depth *E. quaesita* transcriptome and a genome of *Cepaea nemoralis* was used. All analysis was carried out on a UBUNTU server with a 256GB hard drive. A constant deletion of intermediate files was required to permit further analysis.

Quality Filtering

The raw paired-end reads were processed using fastp *v0.20.1* (Chen *et al.*, 2018) to remove illumina adapter sequences and sequences which fall below default quality thresholds. To ensure complete adapter removal, sequences were passed through fastqc v0.11.9 (Andrews, 2019).

Alignment

BWA-MEM v0.7.17 (Burrow-Wheeler Aligner) was used to align the sequencing reads from all individuals against each assembled reference (Li and Durbin, 2009). Deduplication of the aligned reads was carried out by mark duplicates-Picard, (Broad Institute, 2019).

Variant Calling

A variant calling pipeline, bcftools, was implemented on both sets of sequences outputting over 1GB of VCF files for filtering. Unfortunately, upon analysis of the

outputted files aligned to either reference sequence, it became apparent after filtering that the quality of the initial alignments was questionable, potentially due to the genetic divergence between *Euhadra* and *Cepaea nemoralis*. Further analysis using less stringent filter parameters was halted by time constraints inflicted by the reduced computational power of the server 'Mandarina' with multiple users.

2.3 Results

Both phylogenies (Fig 3. A, B) based on whole genome RAD-seq data analysed both display with strong bootstrap a separation of dextral *E. aomoriensis* and sinistral *E.quaesita* from *E.senckenbergiana* and *E. decorata*. In addition, both phylogenies display strong bootstrap support, that *E.aomoriensis* and *E.quaesita* individuals first branch off by region. The phylogenies also show strong support for the separation of the *E. aomoriensis* E102_1 and *E. aomoriensis* E102_2 within the Iwate branch. However, the phylogenies deviate with the placement of E102_4 relative to the sinistral *E.quaesita* individuals in Iwate. Ipyrad, places the *E.aomoriensis* individual branching off between two *E.quaesita* individuals, however in Fig 3. B, there is strong support for the separation of *E.aomoriensis* away from the remaining *E.quaesita* individuals.

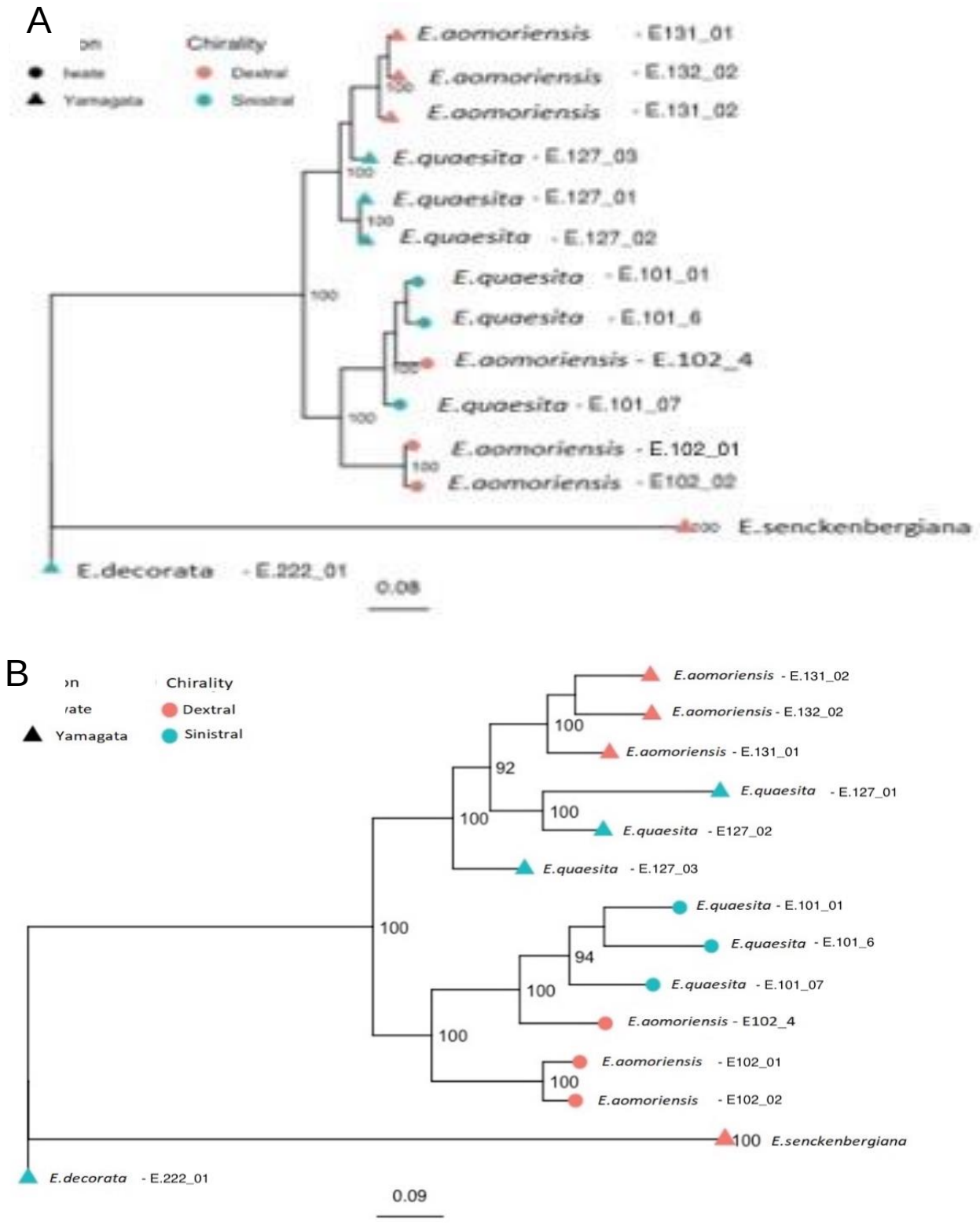


Figure 3.
Phylogenomic analysis of 2946 and 3420 RAD-Seq derived bi-allelic loci assembled using Ipyrad (A) and Stacks2 (B) respectively. A Maximum likelihood phylogeny with strong bootstrap support generated using PMB+F+R2 and TPM3+F models of IQTREE. E102_4 clusters with the sinistral individuals, albeit the precise evolutionary relationship differs.

To further evaluate the relationships between individuals, an MSN for each dataset was generated. Both networks show there is a high degree of genetic distance between *E. decorata*, *E.senckenbergiana*. However, stacks (B) places E.decorata closer to *E.amoriensis* as opposed to *E. quaesita* as in the Ipyrad plot (A). A degree of distance separates the individuals by region, however the distance is calculated as larger from the Stacks dataset. Finally, the placement of E102_4 differs slightly between plots. There is a degree of distance between the *E.aomoriensis* individual (E102_4) and the remaining *E.aomoriensis* in the Ipyrad plot, however STACKS places the individual more closely with *E.quaesita*.

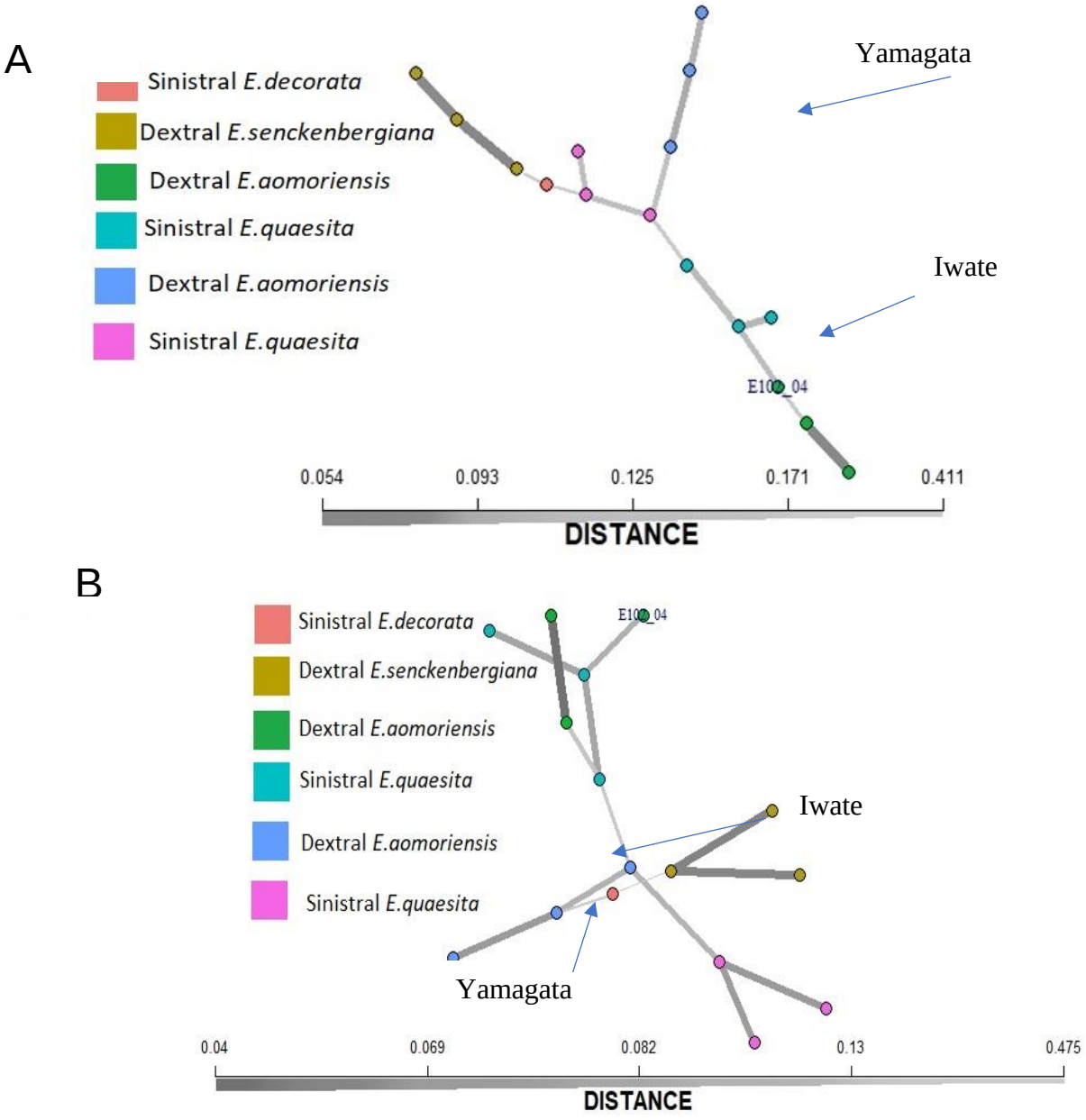


Figure.4 Phylogenomic analysis of 2946 and 3420 RAD-Seq derived bi-allelic loci assembled using Ipyrad (A) and Stacks2 (B) respectively.
Minimum Spanning Network(MSN) shows greater genetic variance between individuals in response to geographic distance as opposed to chirality, in concordance with the ML Trees.

The PCA (fig.5) excluding the outlying groups *E. senckenbergiana* and *E. decorata* concurred with the phylogenetic pattern in both datasets, however the percentage of variation differed depending on pipeline used. The first axis separating the individuals by region accounted for 41.7% (Stacks) and 62% (Ipyrad, B) of the variation respectively. The second axis separated dextral *E. aomoriensis* and sinistral *E.quaesita* within each region explaining 19.6% and 12.5% of the variation, respectively. In both cases the dextral E102_4 is clustered with the three E.quaesita individuals, however, the cluster is strongest in plot Fig 5. A.

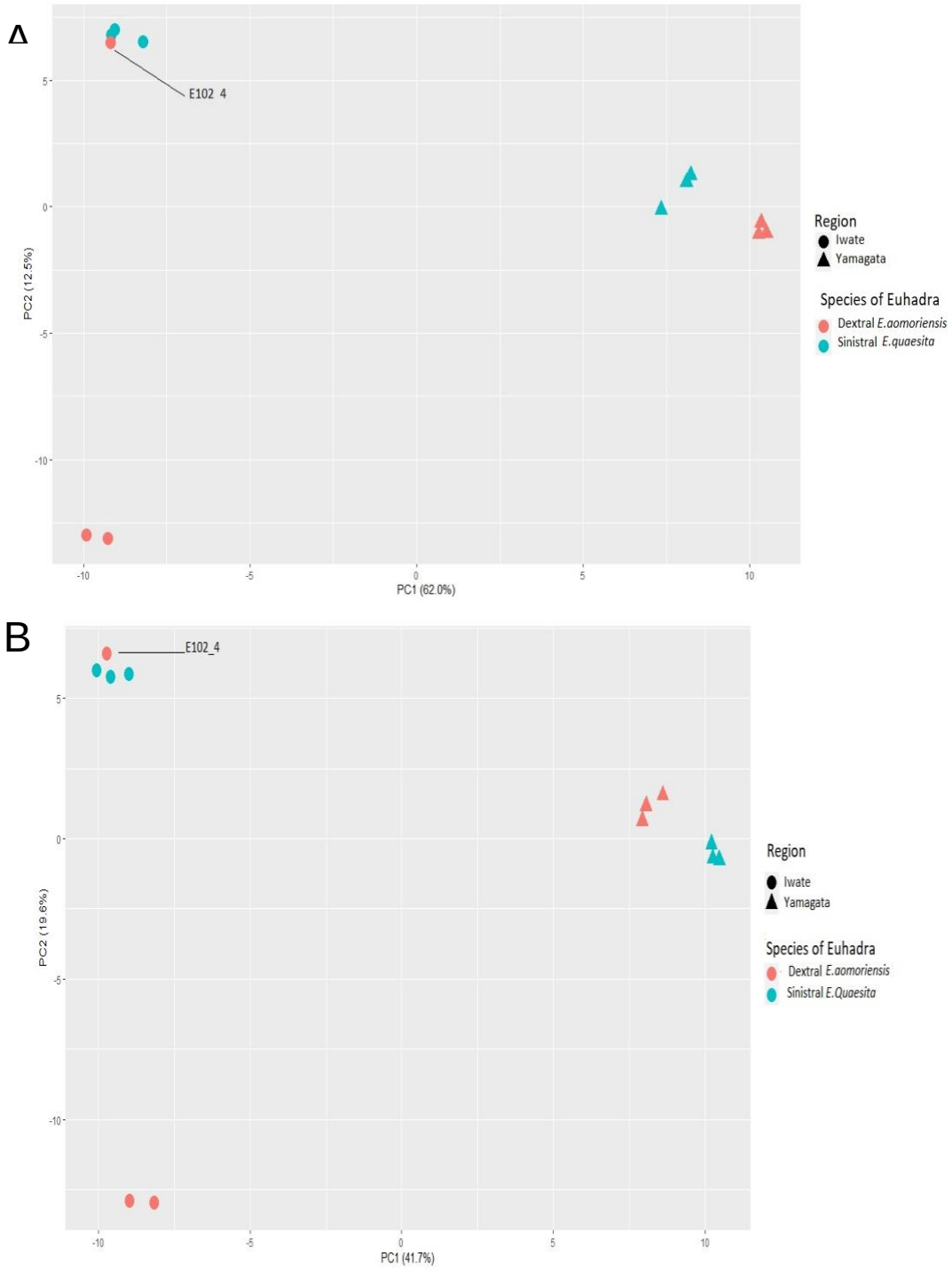


Figure. 5. Phylogenomic analysis of 2946 and 3420 RAD-Seq derived bi-allelic loci assembled using Ipyrad (A) and Stacks2 (B) respectively. Principle components analysis splits the individuals first by region then by species within regions.

To further evaluate the results from each pipeline the genetic divergence between species and region was calculated (Fig,6). The results followed the same pattern as with the previous analyses, a large genetic divergence between *E.quaesita*, *E.aomoriensis* and their respective chiral outgroup, *E.decorata*, and *E. senckenbergiana*. The second largest being between groups located in Iwate and Yamagata. The pattern is consistent between the two plots, however the precise FST values differ, with a greater degree of divergence visible from Stacks dataset (Fig.6B).

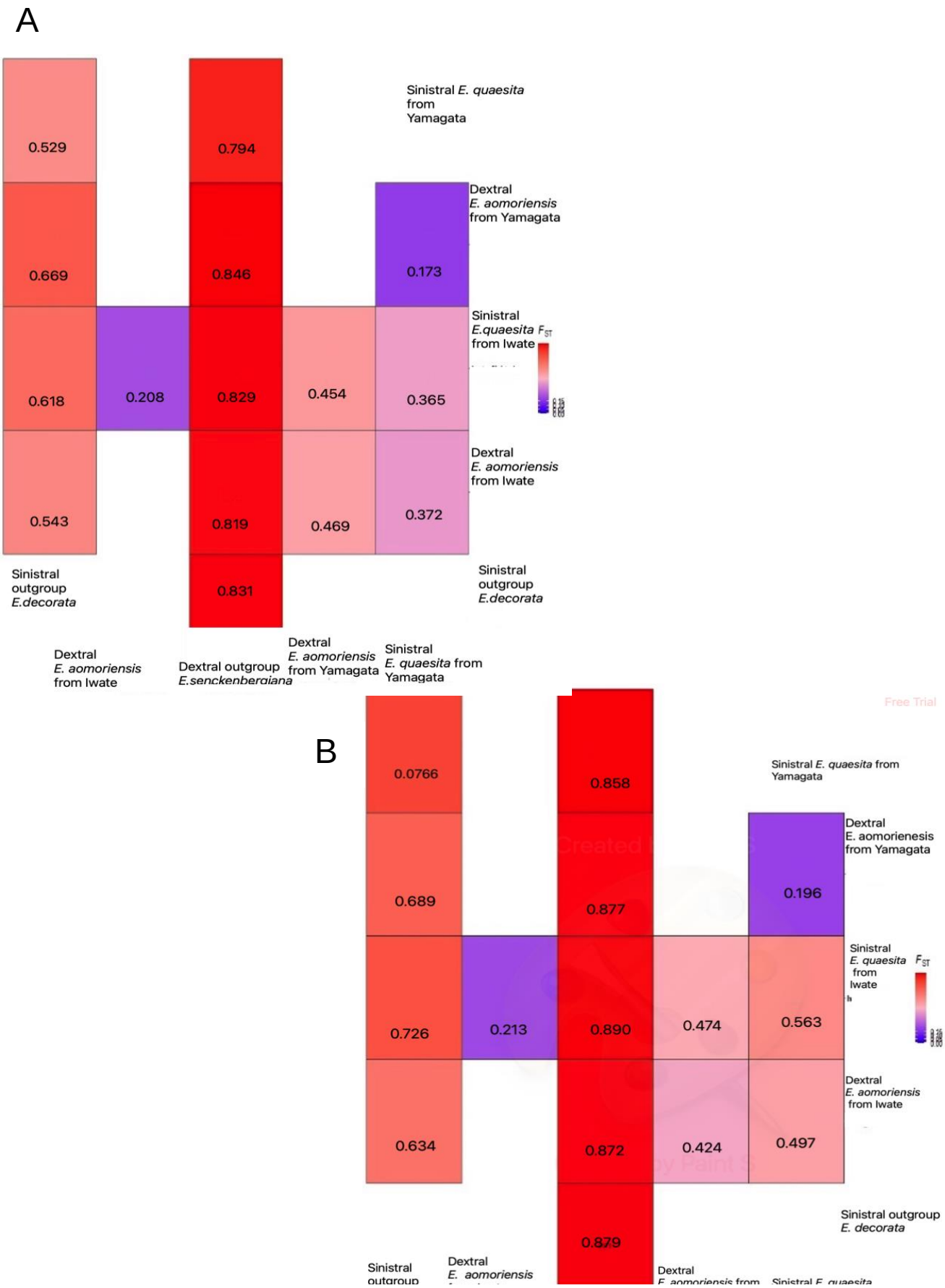


Figure 6.
Visualization of genetic divergence (*Fst* values) between the five populations (Yamagata Dextral, Yamagata Sinistral, Iwate Dextral and Iwate Sinistral) from both datasets. A) Ipyrad B) Stacks

With respect to *E. quaesita* and *E. aomoriensis* interregional genetic divergence is greater than between chiral variants within a particular region, with values of ~ 0.4, ~ 0.2 respectively. The genetic divergence is high between species of the same chirality, sinistral *E. quaesita* and *E. decorata* and dextral *E. aomoriensis* and *E. Senckenbergiana*.

Elucidating Potential Chiral Loci.

To discover putative loci which may contribute to chirality, a PCA was performed on a sinistral *E. quaesita* individual E101_01 alongside the outlying dextral *E.aomoriensis* E102_4. Variation between the individuals was extracted. Genotypes at the divergent loci indicated by the PCA were then elucidated. Two putative loci were discovered where the SNPs in at least three places displayed homozygous divergence from the other individual.

However, the variation within the two loci of the single dextral and sinistral individuals was deemed irrelevant for understanding chirality because the SNP pattern exhibited in E102_4 was not mirrored within other dextral individuals.

Loc 19020	14	18	24	39	40
E. quaesita	AA	AA	AA	AA	AA
E.aomoriensis	TT	TT	TT	TT	TT

Loc 36176	1	2	22	26
E. quaesita	AA/AT	AA	AA	AA
E.aomoriensis	AA/TT	TT	TT	TT

2.4 Discussion

Previous analysis of the genome-wide SNP data generated from the *de novo* assembly and SNP-calling of RAD-seq data using the Stacks pipeline revealed that sinistral *E. quaesita* and dextral *E. aomoriensis* from Iwate display a greater degree of genetic similarity to each other than their counterparts in Yamagata (Richards *et al.*, 2017). Genetic divergence was also high between species of the same chirality within Yamagata (*E. quaesita* and *E. decorata*; *E. aomoriensis* and *E. senckenbergiana*). However, a single dextral individual from Iwate clustered strongly with the sinistral *E. quaesita* snails. Re-assembly of the RAD-seq library using the original pipeline STACKS (Catchen *et al.*, 2013; Rochette, Rivera-Colón and Catchen, 2019) and Ipyrad (Eaton and Overcast, 2020b), and subsequent phylogenetic analysis revealed clustering patterns within the regions and species analogous to each other and to Richards et al. (2017). Divergence was present between pipelines, as shown by the inconsistent placement of E102_4 in the MSN and ML phylogeny, proving the importance of using multiple pipelines for *de novo* analysis.

The highlighted dextral *E. aomoriensis* individual E102_4 repeatedly clustered with members of *E. quaesita* rather than with members of its species. This indicates that the individual is indeed genetically similar to *E. quaesita* and thus a potential resource for understanding the genes determining chirality, as opposed to being the product of an assembly or barcoding error. The variation does not appear to be consistent across all dextral individuals.

One explanation for the variation across RAD loci assembly methods is likely to be attributed to errors in generating the maximum-likelihood trees using IQTREE (Minh *et al.*, 2020). However, research explicitly examining the robustness of phylogenetic tree generation using RADseq data indicates the variability of topology due to assembly methods and parameters (Blanco-Pastor *et al.*, 2019).

The next step to identify the locus or loci responsible for determining chirality in land snails would be to generate a chromosomal-level reference genome of *Euhadra quaesita*.

However, there are several issues to contend with when sequencing gastropod genomes. From the perspective of assembling the sequences, difficulties arise in compiling many repetitive regions within the genome using short reads. This can be alleviated by using long-read sequencing technology, which routinely delivers read lengths >10kb, enabling the entire region to be captured in a single read (Jung *et al.*, 2020). Hybrid assemblies taking advantage of both technologies to improve genome resolution are becoming commonplace. A long-read scaffold structure is assembled initially and then "polished" by short-reads to correct induced errors by the sequencing speed in third-generation technology (Miller *et al.*, 2017; Amarasinghe *et al.*, 2020).

A large drawback with long-read technology such as Oxford Nanopore is the requirement for at least 400ng of high-molecular-weight DNA (> 50kb) (Oxford Nanopore Technologies, 2020). Many biological properties of molluscan taxa, in particular the excretion of a protective mucus layer, act to inhibit the clean extraction of long-read high molecular weight DNA. Pore blocking and flow cell membrane rupture occur if the DNA library loaded is contaminated with proteins such as phenol and salts. However, a recent publication from Chen et al.(2022) provides a solution. To extract the DNA from foot tissue of the Spanish slug *A.vulgaris,* they used a combination of MagAttract HMW DNA Kit and the CTAB method (Arseneau et al., 2016). After extraction the best method, would be to use the ligation sequencing kit from Oxford Nanopore, LSK114 to prepare the library for sequencing on a PromethION. This kit is the upgraded kit to that used by chen et al, (2022), and provides a raw read with higher accuracy. Base-calling with the super-high accuracy base caller will be the next step, followed by denovo assembly following the method outlined in Chen et al, (2022). This will enable a chromosomal-based reference genome to be generated. The reference can be used in combination with the RAD seq data as in this paper investigating colour in bumble bees (Rahman *et al*, 2021).

Citations

Abe, M. and Kuroda, R. (2019) 'The development of CRISPR for a mollusc establishes the formin Lsdia1 as the long-sought gene for snail dextral/sinistral coiling', *Development*, 146(9), p. dev175976. Available at: https://doi.org/10.1242/dev.175976.

Alba, A., Tetreau, G., Chaparro, C., Sánchez, J., Vázquez, A., *et al.* (2019) 'Natural resistance to Fasciola hepatica (Trematoda) in Pseudosuccinea columella snails: A review from literature and insights from comparative "omic" analyses', *Developmental and Comparative Immunology*, 101, p. 103463. Available at: https://doi.org/10.1016/j.dci.2019.103463.

Amarasinghe, S.L. *et al.* (2020) 'Opportunities and challenges in long-read sequencing data analysis', *Genome Biology*, 21(1), p. 30. Available at: https://doi.org/10.1186/s13059-020-1935-5.

Andrews, K.R. *et al.* (2016) 'Harnessing the power of RADseq for ecological and evolutionary genomics', *Nature Reviews Genetics*, 17(2), pp. 81–92. Available at: https://doi.org/10.1038/nrg.2015.28.

Andrews, S. (2019) *FastQC: A quality control tool for high throughput sequence data.* [Java Runtime Environment]. Babraham Bioinformatics.

Arseneau, J.-R., Steeves, R. and Laflamme, M. (2017), Modified low-salt CTAB extraction of high-quality DNA from contaminant-rich tissues. Mol Ecol Resour, 17: 686-693. https://doi.org/10.1111/1755-0998.12616

Asami, T., Cowie, R.H. and Ohbayashi, K. (1998) 'Evolution of mirror images by sexually asymmetric mating behavior in hermaphroditic snails', *The American Naturalist*, 152(2), pp. 225–236. Available at: https://doi.org/10.1086/286163.

Blanco-Pastor, J.L. *et al.* (2019)' Evolutionary networks from RADseq loci point to hybrid origins of Medicago carstiensis and Medicago cretacea', *American Journal of Botany*, 106(9), pp. 1219–1228. Available at: https://doi.org/10.1002/ajb2.1352.

Blum, M. *et al.* (2014) 'The evolution and conservation of left-right patterning mechanisms', *Development*, 141(8), pp. 1603–1613. Available at: https://doi.org/10.1242/dev.100560.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) 'Trimmomatic: a flexible trimmer for Illumina sequence data', *Bioinformatics*, 30(15), pp. 2114–2120. Available at: https://doi.org/10.1093/bioinformatics/btu170.

Boorman, C.J. and Shimeld, S.M. (2002) 'The evolution of left–right asymmetry in chordates', *BioEssays*, 24(11), pp. 1004–1011. Available at: https://doi.org/10.1002/bies.10171.

Boycott, A.E. and Diver, C. (1923) 'On the inheritance of sinistrality in Limnæa peregra', *Proceedings of the Royal Society of London. Series B, Containing Papers*

of a Biological Character, 95(666), pp. 207–213. Available at: https://doi.org/10.1098/rspb.1923.0033.

Broad Institute (2019) 'Picard toolkit', *Broad Institute, GitHub repository* [Preprint]. Available at: http://broadinstitute.github.io/picard/ (Accessed: 25 July 2021).

Brown, N.A. and Wolpert, L. (1990) 'The development of handedness in left/right asymmetry', *Development*, 109(1), pp. 1–9. Available at: https://doi.org/10.1242/dev.109.1.1.

Buddenborg, S.K. *et al.* (2017) 'Transcriptomic responses of Biomphalaria pfeifferi to Schistosoma mansoni: Investigation of a neglected African snail that supports more S. mansoni transmission than any other snail species', *PLOS Neglected Tropical Diseases*, 11(10), p. e0005984. Available at: https://doi.org/10.1371/journal.pntd.0005984.

Burgarella, C. *et al.* (2015) 'Molecular Evolution of Freshwater Snails with Contrasting Mating Systems', *Molecular Biology and Evolution*, 32(9), p. :2403-2416. Available at: https://doi.org/10.1093/molbev/msv121.

Cain, A.J. and Sheppard, P.M. (1977) 'Variation in the spire index of some coiled gastropod shells, and its evolutionary significance', *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 277(956), pp. 377–428. Available at: https://doi.org/10.1098/rstb.1977.0026.

Casanova, A. *et al.* (2021) 'Low impact of different SNP panels from two building-loci pipelines on RAD-Seq population genomic metrics: case study on five diverse aquatic species', *BMC Genomics*, 22(1), p. 150. Available at: https://doi.org/10.1186/s12864-021-07465-w.

Castrillon, D.H. and Wasserman, S.A. (1994) 'Diaphanous is required for cytokinesis in Drosophila and shares domains of similarity with the products of the limb deformity gene', *Development*, 120(12), pp. 3367–3377. Available at: https://doi.org/10.1242/dev.120.12.3367.

Catchen, J. *et al.* (2013) 'Stacks: an analysis tool set for population genomics', *Molecular Ecology*, 22(11), pp. 3124–3140. Available at: https://doi.org/10.1111/mec.12354.

Chen, S. *et al.* (2018) 'fastp: an ultra-fast all-in-one FASTQ preprocessor', *Bioinformatics*, 34(17), pp. i884–i890. Available at: https://doi.org/10.1093/bioinformatics/bty560.

Chougule, A. *et al.* (2020) 'The Drosophila actin nucleator DAAM is essential for left-right asymmetry', *PLoS Genetics*, 16(4), p. e1008758. Available at: https://doi.org/10.1371/journal.pgen.1008758.

Chueca, L.J., Schell, T. and Pfenninger, M. (2021a) 'De novo genome assembly of the land snail Candidula unifasciata (Mollusca: Gastropoda)', *G3 Genes|Genomes|Genetics,*11(8). jkab180, Available at: https://doi.org/10.1093/g3journal/jkab180

Chueca, L.J., Schell, T. and Pfenninger, M. (2021b) 'Whole-genome re-sequencing data to infer historical demography and speciation processes in land snails: the study of two Candidula sister species', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1825), p. 20200156. Available at: https://doi.org/10.1098/rstb.2020.0156.

Conant, G.C., and Wagner, A (2003). 'Asymmetry sequence divergence of duplicate genes.' *Genome research* 13(9) pp.2052-2058. Available at: https://doi.org/10.1101/gr1252603

Courtemanche, N. *et al.* (2013) 'Tension modulates actin filament polymerization mediated by formin and profilin', *Proceedings of the National Academy of Sciences*, 110(24), pp. 9752–9757. Available at: https://doi.org/10.1073/pnas.1308257110.

Danaisawadi, P. *et al.* (2016) 'A snail-eating snake recognizes prey handedness', *Scientific Reports*, 6, p. 23832. Available at: https://doi.org/10.1038/srep23832.

Danecek, P. *et al.* (2011) 'The variant call format and VCFtools', *Bioinformatics*, 27(15), pp. 2156–2158. Available at: https://doi.org/10.1093/bioinformatics/btr330.

Davison, A. *et al.* (2005) 'Speciation and Gene Flow between Snails of Opposite Chirality', *PLOS Biology*, 3(9), p. e282. Available at: https://doi.org/10.1371/journal.pbio.0030282.

Davison, A. *et al.* (2016) 'Formin Is Associated with Left-Right Asymmetry in the Pond Snail and the Frog', *Current Biology*, 26(5), pp. 654–660. Available at: https://doi.org/10.1016/j.cub.2015.12.071.

Davison, A. (2020) 'Flipping Shells! Unwinding LR Asymmetry in Mirror-Image Molluscs', *Trends in Genetics*, 36(3), pp. 189–202. Available at: https://doi.org/10.1016/j.tig.2019.12.003.

Davison, A., Barton, N.H. and Clarke, B. (2009) 'The effect of coil phenotypes and genotypes on the fecundity and viability of Partula suturalis and Lymnaea stagnalis: implications for the evolution of sinistral snails', *Journal of Evolutionary Biology*, 22(8), pp. 1624–1635. Available at: https://doi.org/10.1111/j.1420-9101.2009.01770.x.

Dray, S. and Dufour, A.-B. (2007) 'The ade4 Package: Implementing the Duality Diagram for Ecologists', *Journal of Statistical Software*, 22(4). Available at: https://doi.org/10.18637/jss.v022.i04.

Eaton, D.A.R. and Overcast, I. (2020) 'ipyrad: Interactive assembly and analysis of RADseq datasets', *Bioinformatics*. Edited by R. Schwartz, 36(8), pp. 2592–2594. Available at: https://doi.org/10.1093/bioinformatics/btz966.

Feldmeyer, B. *et al.* (2015) 'Positive selection in development and growth rate regulation genes involved in species divergence of the genus Radix', *BMC Evolutionary Biology*, 15(1), p. 164. Available at: https://doi.org/10.1186/s12862-015-0434-x.

Felsenstein, J. (1981) 'Evolutionary trees from DNA sequences: A maximum likelihood approach', *Journal of Molecular Evolution*, 17(6), pp. 368–376. Available at: https://doi.org/10.1007/BF01734359.

Galla, S.J. *et al.* (2018) 'Reference Genomes from Distantly Related Species Can Be Used for Discovery of Single Nucleotide Polymorphisms to Inform Conservation Management', *Genes*, 10(1), p. 9. Available at: https://doi.org/10.3390/genes10010009.

Gardener, M. (1970) *The Ambidextrous Universe Left Right and the Fall of Parity*. 1st edn. Penguin.

Gittenberger, E. (1988) 'Sympatric Speciation in Snails; a Largely Neglected Model', *evolution*, 42(4), pp. 826–828. Available at: https://doi.org/10.2307/2408875.

Grande, C. and Patel, N.H. (2009) 'Nodal signalling is involved in left-right asymmetry in snails', *Nature*, 457(7232), pp. 1007–1011. Available at: https://doi.org/10.1038/nature07603.

Guo, Y. *et al.* (2019) 'A chromosomal-level genome assembly for the giant African snail Achatina fulica', *GigaScience*, 8(10), p. giz124. Available at: https://doi.org/10.1093/gigascience/giz124.

Hall, T. (1999) 'BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/N', *Nucl. Acids. Symp. Ser.*, 41(1), pp. 95–98.

Heimsath, E.G. and Higgs, H.N. (2012) 'The C terminus of formin FMNL3 accelerates actin polymerization and contains a WH2 domain-like sequence that binds both monomers and filament barbed ends', *The Journal of Biological Chemistry*, 287(5), pp. 3087–3098. Available at: https://doi.org/10.1074/jbc.M111.312207.

Higgs, H.N. and Peterson, K.J. (2005) 'Phylogenetic Analysis of the Formin Homology 2 Domain', *Molecular Biology of the Cell*, 16(1), pp. 1–13. Available at: https://doi.org/10.1091/mbc.e04-07-0565.

Hoang, D.T. *et al.* (2018) 'UFBoot2: Improving the Ultrafast Bootstrap Approximation', *Molecular Biology and Evolution*, 35(2), pp. 518–522. Available at: https://doi.org/10.1093/molbev/msx281.

Johnson, M.S. (1982) 'Polymorphism for direction of coil in Partula suturalis: Behavioural isolation and positive frequency dependent selection', *Heredity*, 49(2), pp. 145–151. Available at: https://doi.org/10.1038/hdy.1982.80.

Jombart, T. (2008) 'adegenet: a R package for the multivariate analysis of genetic markers', *Bioinformatics*, 24(11), pp. 1403–1405. Available at: https://doi.org/10.1093/bioinformatics/btn129.

Jombart, T. and Ahmed, I. (2011) 'adegenet 1.3-1: new tools for the analysis of genome-wide SNP data', *Bioinformatics*, 27(21), pp. 3070–3071. Available at: https://doi.org/10.1093/bioinformatics/btr521.

Jung, H. *et al.* (2020) 'Twelve quick steps for genome assembly and annotation in the classroom', *PLOS Computational Biology*, 16(11), p. e1008325. Available at: https://doi.org/10.1371/journal.pcbi.1008325.

Kalyaanamoorthy, S. *et al.* (2017) 'ModelFinder: fast model selection for accurate phylogenetic estimates', *Nature Methods*, 14(6), pp. 587–589. Available at: https://doi.org/10.1038/nmeth.4285.

Kang, S.W., Patnaik, B.B., Hwang, H.-J., Park, S.Y., Wang, T.H., *et al.* (2016) 'De novo Transcriptome Generation and Annotation for Two Korean Endemic Land Snails, Aegista chejuensis and Aegista quelpartensis, Using Illumina Paired-End Sequencing Technology', *International Journal of Molecular Sciences*, 17(3). Available at: https://doi.org/10.3390/ijms17030379.

Kang, S.W., Patnaik, B.B., Hwang, H.-J., Park, S.Y., Chung, J.M., *et al.* (2016) 'Transcriptome sequencing and de novo characterization of Korean endemic land snail, Koreanohadra kurodana for functional transcripts and SSR markers', *Molecular genetics and genomics: MGG*, 291(5), pp. 1999–2014. Available at: https://doi.org/10.1007/s00438-016-1233-9.

Kang, S.W. *et al.* (2017) 'Sequencing and de novo assembly of visceral mass transcriptome of the critically endangered land snail Satsuma myomphala: Annotation and SSR discovery', *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 21, pp. 77–89. Available at: https://doi.org/10.1016/j.cbd.2016.10.004.

Kerkvliet, J. *et al.* (2017) 'Candidate genes for shell colour polymorphism in Cepaea nemoralis', *PeerJ*, 5, p. e3715. Available at: https://doi.org/10.7717/peerj.3715.

Kuroda, R. *et al.* (2016) 'Diaphanous gene mutation affects spiral cleavage and chirality in snails', *Scientific Reports*, 6(1), p. 34809. Available at: https://doi.org/10.1038/srep34809.

Lambert, J.D. (2010)' Developmental Patterns in Spiralian Embryos', *Current Biology*, 20(2), pp. R72–R77. Available at: https://doi.org/10.1016/j.cub.2009.11.041.

Lebreton, G. *et al.* (2018) 'Molecular to organismal chirality is induced by the conserved myosin 1D.', *Science (New York, N.Y.)*, 362(6417), pp. 949–952. Available at: https://doi.org/10.1126/science.aat8642.

Levin, M., Klar, A.J.S. and Ramsdell, A.F. (2016b) 'Introduction to provocative questions in left–right asymmetry', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1710), p. 20150399. Available at: https://doi.org/10.1098/rstb.2015.0399.

Li, H. and Durbin, R. (2009) 'Fast and accurate short read alignment with Burrows-Wheeler transform', *Bioinformatics*, 25(14), pp. 1754–1760. Available at: https://doi.org/10.1093/bioinformatics/btp324.

Liu, C. *et al.* (2021a) 'Giant African snail genomes provide insights into molluscan whole-genome duplication and aquatic–terrestrial transition', *Molecular Ecology*

Resources, 21(2), pp. 478–494. Available at: https://doi.org/10.1111/1755-0998.13261.

Luo, Y.-J. and Su, Y.-H. (2012) 'Opposing Nodal and BMP Signals Regulate Left–Right Asymmetry in the Sea Urchin Larva', *PLOS Biology*, 10(10), p. e1001402. Available at: https://doi.org/10.1371/journal.pbio.1001402.

Martín-Durán, J.M., Vellutini, B.C. and Hejnol, A. (2016) 'Embryonic chirality and the evolution of spiralian left–right asymmetries', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1710), p. 20150411. Available at: https://doi.org/10.1098/rstb.2015.0411.

McDowell, G.S. *et al.* (2016) 'Conserved roles for cytoskeletal components in dertermining laterality', *Integrative biology : quantitative biosciences from nano to macro*, 8(3), pp. 267–286. Available at: https://doi.org/10.1039/c5ib00281h.

Mercola, M. and Levin, M. (2001) 'Left-Right Asymmetry Determination in Vertebrates', *Annual Review of Cell and Developmental Biology*, 17(1), pp. 779–805. Available at: https://doi.org/10.1146/annurev.cellbio.17.1.779.

Middelkoop, T.C. *et al.* (2021) 'CYK-1/Formin activation in cortical RhoA signaling centers promotes organismal left–right symmetry breaking', *Proceedings of the National Academy of Sciences*, 118(20), p. e2021814118. Available at: https://doi.org/10.1073/pnas.2021814118.

Miller, J.R. *et al.* (2017) 'Hybrid assembly with long and short reads improves discovery of gene family expansions', *BMC Genomics*, 18(1), p. 541. Available at: https://doi.org/10.1186/s12864-017-3927-8.

Minh, B.Q. *et al.* (2020) 'IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era', *Molecular Biology and Evolution*, 37(5), pp. 1530–1534. Available at: https://doi.org/10.1093/molbev/msaa015.

Minh, B.Q., Nguyen, M.A.T. and von Haeseler, A. (2013) 'Ultrafast Approximation for Phylogenetic Bootstrap', *Molecular Biology and Evolution*, 30(5), pp. 1188–1195. Available at: https://doi.org/10.1093/molbev/mst024.

Naganathan, S.R. *et al.* (2014) 'Active torque generation by the actomyosin cell cortex drives left–right symmetry breaking', *eLife*. Edited by J. Ferrell, 3, p. e04165. Available at: https://doi.org/10.7554/eLife.04165.

Noda, T., Satoh, N. and Asami, T. (2019) 'Heterochirality results from reduction of maternal diaph expression in a terrestrial pulmonate snail', *Zoological Letters*, 5. Available at: https://doi.org/10.1186/s40851-018-0120-0.

Okumura, T. *et al.* (2008) 'The development and evolution of left-right asymmetry in invertebrates: Lessons from Drosophila and snails', *Developmental Dynamics*, 237(12), pp. 3497–3515. Available at: https://doi.org/10.1002/dvdy.21788.

Orr, H.A. (1991) 'Is Single-Gene Speciation Possible?', *evolution*, 45(3), pp. 764–769. Available at: https://doi.org/10.2307/2409927.

Ortiz, E.M. (2019) *vcf2phylip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis.* Zenodo. Available at: https://doi.org/10.5281/ZENODO.2540861.

Oxford Nanopore Technologies (2020) *Versatile sequencing library preparation methods for MinION, GridION and PromethION*, *Oxford Nanopore Technologies*. Available at: http://nanoporetech.com/resource-centre/versatile-sequencing-library-preparation-methods-minion-gridion-and-promethion (Accessed: 25 July 2021).

Ponder, W.F and Lindberg, David R. (2008), '*Phylogeny and evolution of the Mollusca*' *edited by Winston F. Ponder, David R. Lindberg* University of California Press Berkeley

Posada, D. and Crandall, K.A. (2001) 'Selecting Models of Nucleotide Substitution: An Application to Human Immunodeficiency Virus 1 (HIV-1)', *Molecular Biology and Evolution*, 18(6), pp. 897–906. Available at: https://doi.org/10.1093/oxfordjournals.molbev.a003890.

Rahman, S.R., Cnaani, J., Kinch, L.N. *et al.* A combined RAD-Seq and WGS approach reveals the genomic basis of yellow colour variation in bumble bee *Bombus terrestris*. *Sci Rep* **11**, 7996 (2021). https://doi.org/10.1038/s41598-021-87194-y

Richards, P.M. *et al.* (2017) 'Single-gene speciation: Mating and gene flow between mirror-image snails', *Evolution Letters*, 1(6), pp. 282–291. Available at: https://doi.org/10.1002/evl3.31.

Rochette, N.C. and Catchen, J.M. (2017) 'Deriving genotypes from RAD-seq short-read data using Stacks', *Nature Protocols*, 12(12), pp. 2640–2659. Available at: https://doi.org/10.1038/nprot.2017.123.

Rochette, N.C., Rivera-Colón, A.G. and Catchen, J.M. (2019) 'Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics', *Molecular Ecology*, 28(21), pp. 4737–4754. Available at: https://doi.org/10.1111/mec.15253.

Saadi, Ahmed J, Davison, A. and Wade, C.M. (2020) 'Molecular phylogeny of freshwater snails and limpets (Panpulmonata: Hygrophila)', *Zoological Journal of the Linnean Society*, 190(2), pp. 518–531. Available at: https://doi.org/10.1093/zoolinnean/zlz177.

Saenko, S.V. *et al.* (2021a) 'The draft genome sequence of the grove snail Cepaea nemoralis', *G3 Genes|Genomes|Genetics*, 11(jkaa071). Available at: https://doi.org/10.1093/g3journal/jkaa071.

Saenko, S.V. *et al.* (2021b) 'The draft genome sequence of the grove snail Cepaea nemoralis', *G3 Genes|Genomes|Genetics*, 11(2), p. jkaa071. Available at: https://doi.org/10.1093/g3journal/jkaa071.

Schell, T. *et al.* (2017) 'An Annotated Draft Genome for Radix auricularia (Gastropoda, Mollusca)', *Genome Biology and Evolution*, 9(3), p. 0. Available at: https://doi.org/10.1093/gbe/evx032.

Schier, A.F. (2003) 'Nodal Signaling in Vertebrate Development', *Annual Review of Cell and Developmental Biology*, 19(1), pp. 589–621. Available at: https://doi.org/10.1146/annurev.cellbio.19.041603.094522.

Schilthuizen, M. (2003) 'Sexual selection on land snail shell ornamentation: a hypothesis that may explain shell diversity', *BMC Evolutionary Biology*, 3(1), p. 13. Available at: https://doi.org/10.1186/1471-2148-3-13.

Schilthuizen, M. *et al.* (2007) 'Sexual selection maintains whole-body chiral dimorphism in snails', *Journal of Evolutionary Biology*, 20(5), pp. 1941–1949. Available at: https://doi.org/10.1111/j.1420-9101.2007.01370.x.

Schilthuizen, M. and Davison, A. (2005) 'The convoluted evolution of snail chirality', *Naturwissenschaften*, 92(11), pp. 504–515. Available at: https://doi.org/10.1007/s00114-05-0045-2.

Schultz, J.H., Adema, C.M. and Kamel, B. (2018) 'The Genome of Physella Acuta', *Unpublished* [Preprint]. Available at: https://www.ncbi.nlm.nih.gov/nuccore/1586278147.

Schultz, J.H., Bu, L. and Adema, C.M. (2018) 'Comparative immunological study of the snail Physella acuta (Hygrophila, Pulmonata) reveals shared and unique aspects of gastropod immunobiology', *Molecular Immunology*, 101, pp. 108–119. Available at: https://doi.org/10.1016/j.molimm.2018.05.029.

Seppälä, O. *et al.* (2021) 'Transcriptome profiling of Lymnaea stagnalis (Gastropoda) for ecoimmunological research', *BMC Genomics*, 22(1), p. 144. Available at: https://doi.org/10.1186/s12864-021-07428-1.

Shafer *et al.* (2017) ' Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference', Methods *in Ecology and Evolution,* 8(8), pp.907-917. Available at: https://doi.org/10.1111/2041-210X.12700

Shimizu, K. *et al.* (2019) 'Insights into the Evolution of Shells and Love Darts of Land Snails Revealed from Their Matrix Proteins', *Genome Biology and Evolution*, 11(2), pp. 380–397. Available at: https://doi.org/10.1093/gbe/evy242.

Spéder, P. and Noselli, S. (2007) 'Left–right asymmetry: class I myosins show the direction', *Current Opinion in Cell Biology*, 19(1), pp. 82–87. Available at: https://doi.org/10.1016/j.ceb.2006.12.006.

Sturtevant, A.H. (1923) 'Inheritance of Direction of Coiling in Limnaea', *Science*, 58(1501), pp. 269–270. Available at: https://doi.org/10.1126/science.58.1501.269.

Sutcharit *et al.* (2007) 'Evolution of whole-body enantiomorphy in the tree snail genus *Amphidromus*', Journal of Evolutionary Biology 20(2), pp. 661-672. Available at: https://doi.org/10.1111/j.1420-9101.2006.01246.x

Taylor, S.A., Anderson, D.J. and Friesen, V.L. (2013) 'Evidence for Asymmetrical Divergence-Gene Flow of Nuclear Loci, but Not Mitochondrial Loci, between Seabird Sister Species: Blue-Footed (Sula nebouxii) and Peruvian (S. variegata) Boobies', *PLOS ONE*, 8(4), p. e62256. Available at: https://doi.org/10.1371/journal.pone.0062256.

Teasdale, L.C. *et al.* (2016) 'Identification and qualification of 500 nuclear, single-copy, orthologous genes for the Eupulmonata (Gastropoda) using transcriptome sequencing and exon capture', *Molecular Ecology Resources*, 16(5), pp. 1107–1123. Available at: https://doi.org/10.1111/1755-0998.12552.

Tee, Y.H. *et al.* (2021) 'Actin polymerization and crosslinking drive left-right asymmetry in single cell and cell collectives'. bioRxiv, p. 2021.04.22.440942. Available at: https://doi.org/10.1101/2021.04.22.440942.

Tennessen, J.A. *et al.* (2020) 'Clusters of polymorphic transmembrane genes control resistance to schistosomes in snail vectors', *eLife*. Edited by D. Ebert, D. Soldati-Favre, and Y. Bourgeois, 9, p. e59395. Available at: https://doi.org/10.7554/eLife.59395.

Tills, O. *et al.* (2018) 'Transcriptomic responses to predator kairomones in embryos of the aquatic snail Radix balthica', *Ecology and Evolution*, 8(22), pp. 11071–11082. Available at: https://doi.org/10.1002/ece3.4574.

Tills, O., Truebano, M. and Rundle, S. (2015) 'An embryonic transcriptome of the pulmonate snail Radix balthica', *Marine Genomics*, 24, pp. 259–260. Available at: https://doi.org/10.1016/j.margen.2015.07.014.

Tingler, M. *et al.* (2018) 'A Conserved Role of the Unconventional Myosin 1d in Laterality Determination', *Current Biology*, 28(5), pp. 810-816.e3. Available at: https://doi.org/10.1016/j.cub.2018.01.075.

Ubrihien, R.P. *et al.* (2017) 'The response of Isidorella newcombi to copper exposure: Using an integrated biological framework to interpret transcriptomic responses from RNA-seq analysis', *Aquatic Toxicology*, 185, pp. 183–192. Available at: https://doi.org/10.1016/j.aquatox.2017.02.014.

Ueshima, R. and Asami, T. (2003) 'Single-gene speciation by left–right reversal', *Nature*, 425(6959), pp. 679–679. Available at: https://doi.org/10.1038/425679a.

Utsuno, H. and Asami, T. (2010) 'Maternal Inheritance of Racemism in the Terrestrial Snail Bradybaena similaris', *Journal of Heredity*, 101(1), pp. 11–19. Available at: https://doi.org/10.1093/jhered/esp058.

Vandenberg, L.N. and Levin, M. (2013) 'A unified model for left–right asymmetry? Comparison and synthesis of molecular models of embryonic laterality', *Developmental Biology*, 379(1), pp. 1–15. Available at: https://doi.org/10.1016/j.ydbio.2013.03.021.

Wade, C. and Davison, A. (no date) ‘Acroloxus sp. Transcriptome', *Unpublished* [Preprint].

Wood, W.B. (1991) 'Evidence from reversal of handedness in C. elegans embryos for early cell interactions determining cell fates', *Nature*, 349(6309), pp. 536–538. Available at: https://doi.org/10.1038/349536a0.

Wood, W.B. (2005) 'The Left–Right Polarity Puzzle: Determining Embryonic Handedness', *PLoS Biology*, 3(8), p. e292. Available at: https://doi.org/10.1371/journal.pbio.0030292.

Yu, G. *et al.* (2017) 'ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data', *Methods in Ecology and Evolution*, 8(1), pp. 28–36. Available at: https://doi.org/10.1111/2041-210X.12628.

Zhao, M. *et al.* (2018) 'eSnail: A transcriptome-based molecular resource of the central nervous system for terrestrial gastropods', *Molecular Ecology Resources*, 18(1), pp. 147–158. Available at: https://doi.org/10.1111/1755-0998.12722.