

**EVOLUTION OF THE TERRESTRIAL INVASION  
IN PANPULMONATA (MOLLUSCA,  
GASTROPODA): MOLECULAR ADAPTATIONS  
IN THE CONTEXT OF REALM TRANSITIONS**

Dissertation (verfasst in Englisch)  
zur Erlangung des Doktorgrades  
der Naturwissenschaften

vorgelegt beim Fachbereich Biowissenschaften  
der Johann Wolfgang Goethe - Universität  
in Frankfurt am Main

von

**Pedro Eduardo Romero Condori**

aus Lima, Perú

Frankfurt am Main, 2016

(D 30)

vom Fachbereich Biowissenschaften der  
Johann Wolfgang Goethe - Universität als Dissertation angenommen.

Dekanin: Prof. Dr. Meike Piepenbring.

Gutachter: Prof. Dr. Markus Pfenninger.

Gutachterin: Prof. Dr. Imke Schmitt.

Datum der Disputation:

# CONTENTS

<b>Summary</b>	<b>5</b>
<b>Zusammenfassung</b>	<b>7</b>
<b>1. Chapter 1: General Introduction</b>	<b>13</b>
1.1. The land invasion	13
1.2. Evolution and phylogeny of the Panpulmonata	16
1.3. Molecular adaptations in the context of realm transitions	21
1.4. Convergent and parallel evolution	23
1.5. Thesis questions and objectives	27
<b>2. Chapter 2: General Discussion</b>	<b>29</b>
2.1. Tempo and mode of the terrestrialization within Panpulmonata	29
2.2. Adaptive evolution in the context of terrestrialization	32
2.3. Convergent evolution	40
2.4. Conclusions and outlook	42
<b>Acknowledgments</b>	<b>45</b>
<b>References</b>	<b>46</b>

**Appendix: Publications and manuscripts**

53

- *Publication 1: Romero PE, Pfenninger M, Kano Y, Klussmann-Kolb A.*  
Molecular phylogeny of the Ellobiidae (Gastropoda: Panpulmonata) supports independent terrestrial invasions. *Molecular Phylogenetics and Evolution*. 2016;97:43-54. <http://dx.doi.org/10.1016/j.ympev.2015.12.014> 53
- *Publication 2: Romero PE, Weigand AM, Pfenninger M.*  
Positive selection on panpulmonate mitogenomes provide new clues on adaptations to terrestrial life. *BMC Evolutionary Biology*. 2016;16(1):164. <http://dx.doi.org/10.1186/s12862-016-0735-8> 69
- *Manuscript 1: Romero PE, Feldmeyer B, Pfenninger M.*  
Panpulmonate transcriptomes reveal candidate genes involved in the adaptation to terrestrial and freshwater habitats in gastropods. <http://dx.doi.org/10.1101/072389> 86

**Curriculum Vitae**

110



## SUMMARY

The transition from the marine to the terrestrial realm is one of the most fascinating issues in evolutionary biology for it required the appearance, in different organisms, of several novel adaptations to deal with the demands of the new realm. Adaptations include, for instance, modifications in different metabolic pathways, development of body structures to facilitate movement and respiration, or tolerance to new conditions of stress. The transition to the land also gives an extraordinary opportunity to study whether evolution used similar changes at the genomic level to produce parallel adaptations in different taxa.

Mollusks are among taxa that were successful in the conquest of the land. For instance, several lineages of the molluscan clade Panpulmonata (Gastropoda, Heterobranchia) invaded the intertidal, freshwater and land zones from the marine realm. In my dissertation, using tools from bioinformatics, phylogenetics, and molecular evolution, I used panpulmonates as a suitable model group to study the independent invasions into the terrestrial realm and the adaptive signatures in genes that may have favored the realm transitions.

My work includes two peer-reviewed published papers and one manuscript *under review*. In Publication 1 (Romero et al., 2016a), I used mitochondrial and nuclear molecular markers to resolve the phylogeny of the Ellobiidae, a family that possesses intertidal and terrestrial species. The phylogeny provided an improved resolution of the relationships within inner clades and a framework to study the tempo and mode of the land transitions. I showed that the terrestrialization events occurred independently, in different lineages (Carychiinae, Pythiinae) and in different geological periods. (Mesozoic, Cenozoic). In addition, the diversification in this group may not have been affected by past geological or climate changes as the Cretaceous-Paleogene (K-Pg) event or the sea-level decrease during the Oligocene.

In Publication 2 (Romero et al., 2016b), I generated new mitochondrial genomes from terrestrial species and compared them with other panpulmonates. I

used the branch-site test of positive selection and detected significant nonsynonymous changes in the terrestrial lineages from Ellobioidea and Stylommatophora. Two genes appeared under positive selection: *cob* (Cytochrome b) and *nad5* (NADH dehydrogenase 5). Surprisingly, I found that the same amino acid positions in the proteins encoded by these genes were also under positive selection in several vertebrate lineages that transitioned between different habitats (whales, bats and subterranean rodents). This result suggested an adaptation pattern that required parallel genetic modifications to cope with novel metabolic demands in the new realms.

In Manuscript 1 (Romero et al., *under review*), I *de novo* assembled transcriptomes from several panpulmonate specimens resulting in thousands of genes that were clustered in 702 orthologous groups. Again, I applied the branch-site test of positive selection in the terrestrial lineages from Ellobioidea and Stylommatophora and in the freshwater lineages from Hygrophila and Acochlidia. Different sets of genes appeared under positive selection in land and freshwater snails, supporting independent adaptation events. I identified adaptive signatures in genes involved in gas-exchange surface development and energy metabolism in land snails, and genes involved in the response to abiotic stress factors (radiation, desiccation, xenobiotics) in freshwater snails.

My work provided evidence that supported multiple land invasions within Panpulmonata and provided new insights towards understanding the genomic basis of the adaptation during sea-to-land transitions. The results of my work are the first reports on the adaptive signatures at the codon level in genes that may have facilitated metabolic and developmental changes during the terrestrialization in the phylum Mollusca. Moreover, they contribute to the current debate on the conquest of land from the marine habitat, a discussion that has been only based in vertebrate taxa. Future comparative genome-wide analyses would increase the number of genes that may have played a key role during the realm transitions.

# ZUSAMMENFASSUNG

## ENTWICKLUNG DER LANDINVASION IN PANPULMONATA (MOLLUSCA, GASTROPODA): MOLEKULARE ADAPATIONEN IM KONTEXT DES BEREICHS TRANSITIONEN

### Hintergrund

Der Übergang vom marinen zum terrestrischen Lebensraum ist eines der faszinierendsten Studienfelder der Evolutionsbiologie, denn um den Anforderungen dieser neuen Umgebung gerecht zu werden, sind verschiedene Anpassungen nötig. Der Lebensraumwechsel vom Meer zum Land vollzog sich unabhängig in verschiedenen Organismen, beispielsweise in Flechten, grünen Pflanzen, Arthropoden, Mollusken und Wirbeltieren. Dieser Vorgang setzte Modifikationen unterschiedlicher somatischer Systeme und Organe voraus, welche zunächst an aquatische Habitate angepasst waren. Zu den Beispielen dieser Anpassungen zählen die Entwicklung interner Atmungsorgane (Lungen) zur Sauerstoffaufnahme aus der Luft, verschiedene Hautmodifikationen wie die Entwicklung von Kutikula- und Keratinschichten zum Schutz vor Wasserverlust, die Produktion neuer Stoffe zur Ausscheidung von Stickstoff (z.B. als Harnsäure und Harnstoff) und die Präsenz eines Skeletts sowie einer starken Muskulatur, um den Körper zu stützen und zu bewegen.

Der Übergang vom Meer zum Land eröffnet ebenfalls außergewöhnliche Möglichkeiten, um zu untersuchen, ob die Evolution konvergenter Merkmale in unterschiedlichen Taxa auf den gleichen Veränderungen der genetischen Ebene beruht. Dies würde bedeuten, dass in unterschiedlichen Taxa dieselben Gene unter Selektion stünden und konvergente Merkmale auf verschiedenen organismischen Ebenen, wie Proteinstruktur und -funktion, Stoffwechselprozesse und Gewebeentwicklung letztlich auf den gleichen molekulargenetischen Mechanismen beruhen.

Mollusken sind eines der Taxa, welche erfolgreich das Land eroberten. So besiedelten mehrere Abstammungslinien der Panpulmonata (Gastropoda, Heterobranchia) erfolgreich die Gezeitenzone und Brackwasser, Binnengewässer (Süßwasser) und Land sowie marine Lebensräume (Salzwasser). Unter Zuhilfenahme von Methoden der Bioinformatik, Phylogenetik und Molekulargenetik, untersuchte ich im Zuge meiner Dissertation, unabhängige Landgänge verschiedener Gruppen innerhalb der Panpulmonata, und die damit verbundenen genetischen Veränderungen, welche durch die Wasser-Land-Transition favorisiert wurden.

### **Durchgeführte Studien**

Meine Arbeit umfasst zwei in Fachzeitschriften publizierte Artikel sowie ein Manuskript, welches sich noch im Begutachtungsprozess befindet. In der ersten Publikation (Romero et al., 2016a) untersuchte ich die Familie der Ellobiidae (Panpulmonata, Ellobioidea). Diese Familie umfasst Arten, welche in den Gezeitenzonen sowie in terrestrischen Habitaten vorkommen. In dieser Arbeit nutzte ich mitochondriale und nukleäre Marker (ribosomale Gene: 28S, 18S, 16S und 12S; sowie proteinkodierende Gene: *cox1* und *H3*), um phylogenetische Beziehungen innerhalb dieser Familie zu entschlüsseln. Anschließend rekonstruierte ich den phylogenetischen Stammbaum der Ellobiidae unter Verwendung der maximum likelihood-Methode und Bayes'scher Verfahren. Die von mir abgeleitete Phylogenie enthüllte eine verbesserte Auflösung der Beziehungen innerhalb der Familie. Dieses Ergebnis bestätigt frühere phylogenetische Hypothesen, welche auf morphologischen Daten basieren, beispielsweise die Monophylien der Subfamilien Carychiinae, Ellobiinae und Pythiinae, sowie die Einordnung von Carychiinae und Ellobiinae als Schwestertaxa.

Die resultierende Phylogenie diente ebenfalls als Bezugssystem zur Untersuchung der Dauer und Art des Lebensraumübergangs. Ich zeigte, dass sich

der Landgang innerhalb der Ellobiidae unabhängig in den Subfamilien Carychiinae und Pythiinae vollzog. Zusätzlich wies ich nach, dass diese Ereignisse während unterschiedlicher geologischer Ären geschahen (Mesozoikum bzw. Cenozoikum). Desweiteren konnte ein Einfluss geologischer oder klimatischer Ereignisse, wie das Massenaussterben der Kreide-Paläogen-Grenze (K-P-Grenze) oder der Abfall des Meeresspiegels während des Oligozäns, auf Vielfalt der Ellobiidae ausgeschlossen werden.

In meiner zweiten Publikation (Romero et al., 2016b) sequenzierte ich mitochondriale Genome der terrestrischen Panpulmonaten *Carychium tridentatum* (Ellobioidea), *Arion rufus* und *Helicella itala* (beide Stylommatophora). Diese Sequenzen wurden mit 47 weiteren mitochondriellen Genomen von bedeutenden Abstammungslinien der Panpulmonata verglichen. Anschließend führte ich einen branch site-Test durch, um positive Selektion nachzuweisen. Der Test ergab signifikante nicht-synonyme Mutationen in den terrestrischen Zweigen, die zu *Carychium* und den Stylommatophora führten. Diese Veränderungen waren in den Genen *cob* (Cytochrom b) und *nad5* (NADH- Dehydrogenase 5) lokalisiert, die beide Teil der oxidativen Phosphorylierung sind (OXPHOS Komplex III bzw. I). Die positive Selektion beider Gene legt nahe, dass sich die Anpassung an eine terrestrische Umgebung episodenhaft vollzog. So wies ich eine Beeinflussung der Massenwirkungskonstante durch Aminosäureänderungen (Ionisation der Carboxygruppe) in beiden Genen nach. Diese physiochemische Eigenschaft von Aminosäuren beeinflusst die Produktion reaktiver Sauerstoffspezies (ROS). Die Fähigkeit der Reduktion von ROS unter neuen Stresskonditionen (z.B. Extremtemperaturen oder Austrocknung), könnte die Anpassung einer Art an terrestrische Habitate begünstigen. Desweiteren verglich ich die Aminosäuresequenzen der Proteine terrestrischer Panpulmonata, welche von *cytb* und *nad5* kodiert werden, mit Wirbeltieren, die ebenfalls eine Anpassung an neue Habitate vollzogen: Wale, Fledermäuse und unterirdisch lebende Nager. Überraschenderweise fand ich in beiden Genen übereinstimmende positiv

selektierte Aminosäuren. Dieses Ergebnis weist auf ein Anpassungsmuster hin, welches parallele genetische Modifikationen voraussetzt, um mit neuen metabolischen Anforderungen in einer neuen Umwelt umgehen zu können. Daraus lässt sich ableiten, dass neue energetische Anforderungen des Landganges sowie die Notwendigkeit der Toleranz gegenüber neuen abiotischen Stressfaktoren, die physiologische Einschränkungen in terrestrischer Panpulmonata veränderten und Anpassungen mitochondrialer Gene auslösten.

In Manuskript 1 (Romero et al., *in review*) erweiterte ich anhand von Transkriptomdaten die Suche nach positiv selektierten Genen in terrestrischen Panpulmonaten. In diesen Datensätzen waren die terrestrische Abstammungslinien von *Carychium* sp., *Pythia pachyodon* (Ellobioidea) sowie *Arion vulgaris* (Stylommatophora) repräsentiert, während *Strubellia wawrai* (Acochlidia), *Biomphalaria glabrata*, *Planorbarius corneus* und *Radix balthica* (Hygrophila) als Vertreter limnischer Linien analysiert wurden. Ich erstellte de novo Transkriptome verschiedener Individuen resultierend in 50.000 bis 100.000 Transkripten (10% bis 15% funktional annotiert), welche sich auf 702 orthologe Cluster verteilen. Erneut führte ich einen branch site- Test auf positive Selektion in den terrestrischen Abstammungslinien von Ellobioidea und Stylommatophora sowie in den limnischen Abstammungslinien von Hygrophila und Acochlidia durch. In terrestrischen und limnischen Schnecken wurden unterschiedliche Gene positiv selektiert, was die Hypothese unabhängiger Anpassungsereignisse unterstützt. In Landschnecken identifizierte ich adaptive Signaturen in sieben Genen, welche am Aufbau von Aktin, dem Transport von Glukose und dem Tyrosin-Stoffwechsel beteiligt sind. In Süßwasserschnecken fand ich 28 positiv selektierte Gene, welche am DNA-Reparaturmechanismus, dem Abbau von Xenobiotika, dem mitochondrialen Elektronentransport oder der Biogenese der Ribosomen beteiligt sind. Positiv selektierte Gene waren Teil der Signalwege des Energiemetabolismus und der Entwicklung von Atmungsoberflächen in terrestrischen Schnecken; und

Teil der Antwort auf abiotische Stressfaktoren (Strahlung, Austrocknung, Xenobiotika) in Süßwasserschnecken.

### **Fazit**

Meine Arbeit erbrachte Beweise für multiple Landgänge innerhalb der Panpulmonata und vielversprechende Einsichten in das Verständnis der genetischen Grundlage der Anpassung während des Landgangs. Ich konnte nachweisen, dass sich die Anpassung an das Leben an Land mindestens zweimal unabhängig voneinander in den Ellobiidae vollzog (Publikation 1). Dieses Ergebnis erweitert die Anzahl unabhängiger Landgänge in den Panpulmonata (Acochlidia, Stylommatophora, Systellommatophora) zu mindestens fünf unabhängige Ereignisse. Obwohl sich der Landgang in unterschiedlichen Zweigen des Baums des Lebens vollzog, haben es nur wenige Abstammungslinien geschafft, erfolgreich das Land zu besiedeln und dort zu bleiben. So ist meine erste Publikation ein bedeutendes Beispiel der inhärenten Fähigkeit von Panpulmonaten, sich an die terrestrische Umwelt anzupassen.

Desweiteren fokussierte ich auf die Untersuchung der genetischen Basis von Anpassungen an die nicht-marine Umwelt in Schnecken (Publikation 2, Manuskript 1), ein Thema, welches in Invertebraten wenig untersucht ist. Tatsächlich wurde bis zum jetzigen Zeitpunkt lediglich eine Studie veröffentlicht, welche sich mit dem Landgang in Springschwänzen (Collembola) und Insekten (Hexapoda) beschäftigt. Meine Arbeit ist der erste Bericht über adaptive Signale auf dem Codonlevel von Genen, die Stoffwechsel- und Entwicklungsänderungen während des Landgangs im Phylum Mollusca ermöglichten, und leistet einen wichtigen Beitrag in der wissenschaftlichen Debatte über Meer-Land-Transitionen, die bisher von Erkenntnissen in Wirbeltieren dominiert wurde.

So untersuchte ich mitochondriale Genome (Publikation 2) und Transkriptome (Manuskript 1) von Panpulmonaten und identifizierte Anpassungen auf dem genetischen Level, speziell in den Signalwegen des

Energiehaushaltes und der Stresstoleranz. Diese Anpassungen gewährleisteten wahrscheinlich das Überleben und den Erfolg während des Übergangs zu nicht-marinen Lebensräumen. Ich fand Hinweise auf parallele Evolution in den Mitochondrien von terrestrischen Panpulmonaten, jedoch nicht in den Transkriptomen. Eine mögliche Erklärung dieses Phänomens mag sein, dass sowohl die geringe Größe des mitochondrialen Genoms, als auch die niedrige Zahl kodierender Gene und die grundlegende Rolle jener Gene in der Sauerstoffatmung die Möglichkeiten für Selektion begrenzt, was häufigere parallele Evolutionsereignisse begünstigt. Im Falle der Transkriptionsdaten, in denen die Anzahl von Genen wesentlich höher ist, kann Selektion unabhängig in mehreren verschiedenen Genen (und Stoffwechselwegen) agieren. Da ich lediglich Hunderte von Genen analysierte, bleibt die Möglichkeit auf parallele Evolution im nukleären Genom bestehen. Weitere vergleichende genomweite Studien, welche die Zahl von Genen mit Schlüsselrollen in der Transition zwischen Lebensräumen erweitern könnten, den Einblick in eine mögliche parallele Entwicklung verbessern.



## Chapter 1: GENERAL INTRODUCTION

The transition from the marine to the terrestrial realm is one of the most fascinating issues in evolutionary biology for it required the appearance of several novel adaptations to deal with novel demands in the new realm. The colonization of land from the marine habitat occurred many times in different organisms previously adapted to the marine habitat and involved numerous modifications. Adaptations include, for instance, modifications in different metabolic pathways and organs, development of body structures to facilitate movement and respiration, or tolerance to new conditions of stress. Independent transitions to the land give an extraordinary opportunity to study whether evolution used similar changes at the genomic level to produce parallel adaptations in different taxa.

Mollusks are among taxa that were successful in the conquest of the land. For instance, several lineages of the molluscan clade Panpulmonata (Gastropoda, Heterobranchia) invaded the intertidal, freshwater and land zones from the marine realm. For this reason, in my dissertation, I will use panpulmonates as a model group to study the independent invasions into the terrestrial realm and the adaptive signatures in genes that may have favored the realm transitions.

In the following, I will present the topic of the land invasion and describe the evolution of Panpulmonata. Then, I will review some molecular adaptations related to realm transitions. Furthermore, I will elaborate on the topic of convergent and parallel evolution and link my research in this broader evolutionary context. Finally, I will define my research questions and objectives.

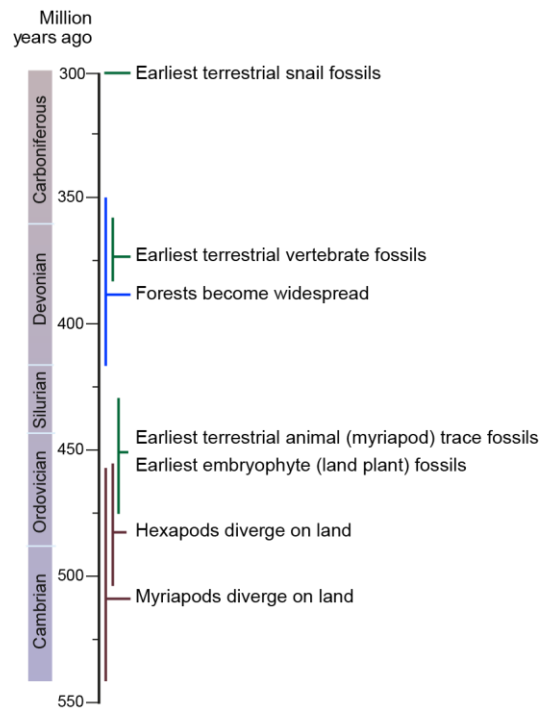
### 1.1. The land invasion

The conquest of land by organisms that evolved from marine ancestors is a fascinating topic in the evolution of life (Lillywhite, 2012). Transitions occurred multiple times across different taxa, *e.g.* bacteria, embryophytes, arthropods, mollusks, annelids and vertebrates (Laurin, 2010). The earliest fossils of land plants and animals

date back from the Paleozoic (~450 Ma) (Dunn, 2013); although, there is some evidence of other organisms living in the land before this period such as cyanobacteria or algae (~1200 Ma) (Strother et al., 2011). The transition in embryophytes (the clade including liverworts, hornworts, mosses, and vascular plants) occurred during the Ordovician and Silurian periods (Delaux et al., 2012) (Figure 1). This transition was benefited by the symbiotic association of the first land plants with arbuscular mycorrhizal fungi allowing them to absorb soil nutrients in the new realm (Delaux et al., 2015). The first direct evidence from animal fossils come from the myriapods (centipedes and millipedes) (Figure 1), followed by other arthropod groups and vertebrates (~370 Ma) (Laurin, 2010; Rota-Stabelli et al., 2013). Mollusks invaded the land later. For example, fossils from *Dawsonella meeki* (Gastropoda, Neritimorpha), one of the oldest land snail known, belong to the Late Carboniferous (~300 Ma) (Kano et al., 2002).

Previous major changes in the composition of the atmosphere during the late Proterozoic (~1000 Ma) especially, the increase of the ozone (O<sub>3</sub>) levels regulating the ultraviolet (UV) radiation flux and the decrease of the mean global temperature, likely, set the stage for the land invasion (Gordon and Olson, 1995). Still, the ancestors of land plants and animals needed to deal with many issues during the terrestrialization including desiccation and gas exchange.

Plants generated a waxy layer named cuticle in order to avoid desiccation. This layer is formed by a fatty acid polymer named cutin. Cuticle accumulation works as a permeability barrier to protect against external water, soil, and contamination by microorganisms, or confer rigidity in various tissues. Plants also produce various types of phenolic compounds (lignin, suberin, anthocyanin, flavonoids, etc.) in response to stress factors. In addition, they develop specialized epidermal structures (stomata) to facilitate carbon uptake and control the evaporation rate (Delaux et al., 2012).



**Figure 1.** Timeline of the colonization of land during the Paleozoic.

Fossil evidence in animals and embryophytes is indicated with green lines. Divergence times inferred from molecular data in myriapods and hexapods (Rota-Stabelli et al., 2013) is indicated with brown lines. The 95% credibility interval is indicated with vertical lines and the mean, with a horizontal line. Adapted from Dunn (2013).

Animals also generated a cuticle layer made from a combination of polysaccharides, lipids, and structural proteins. For example, chitin is present in the exoskeleton of arthropods, while keratins are present in vertebrate claws, hairs, nails and wool (Vandebergh and Bossuyt, 2012). Also, in order to obtain oxygen from the air, animals developed vascularized air-filled cavities (lungs). In addition, terrestrial animals have to deal with the nitrogenous waste products; fish (and amphibians) produce ammonia, a highly toxic compound but highly soluble in water, while reptiles and vertebrates excrete urea and uric acid, less toxic compounds that require less water to be excreted. Other adaptations in animals include the development of internal fertilization, olfactory, hearing and vision organs, structures for movement, and complex behavior (Little, 1990).

Although the transition to the land occurred multiple times, it has been argued that, actually, it only happened in few animal clades: annelids, arthropods, mollusks, onychophorans, vertebrates and worms (Vermeij and Dudley, 2000). According to these authors, organisms well adapted for marine life will be poorly suited for terrestrial life (competitive disadvantage) due to the drastic differences in physical properties of both realms. Thus, most of the invasions could have only happened during the Paleozoic when land animals and plants were absent or had a very modest competitive ability. However, there are some examples of invasions in crustaceans (*e.g.* isopods, amphipods, and brachyurans) and gastropod mollusks (*e.g.* cyclophoroideans, littorinoideans, and ellobiids) during the Mesozoic and Cenozoic, respectively (Vermeij and Dudley, 2000). It is likely that terrestrialization in gastropods occurred multiple times within the same lineages (Kameda and Kato, 2011). This is the reason why I considered to study the evolution of gastropod clades in the context of the land invasions.

## **1.2. Evolution and phylogeny of the Panpulmonata**

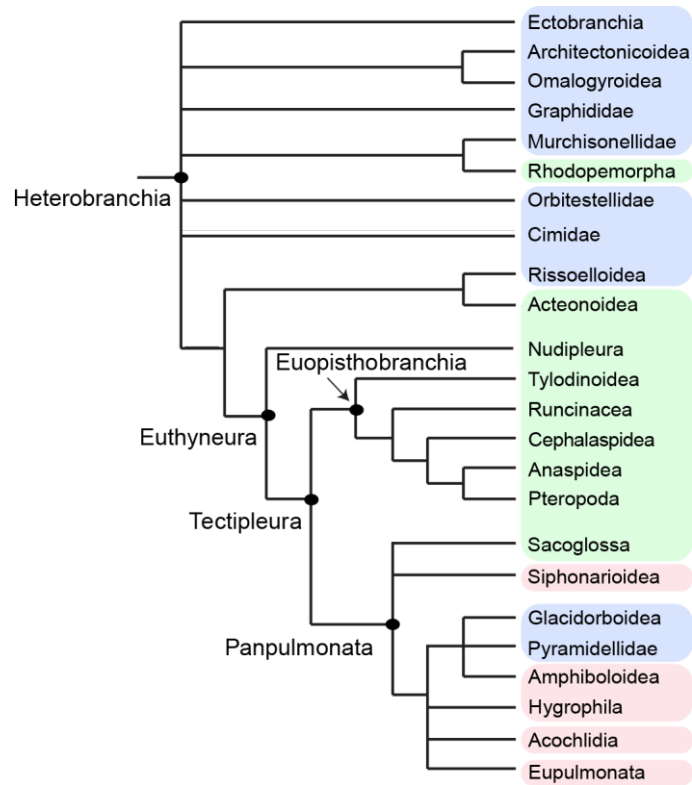
Mollusks are the second most diverse group after Arthropoda and one of the most morphologically variable metazoan phyla (Kocot, 2013). We can find within this group worm-like organisms (Caudofoveata, Solenogastres); chitons (Polyplacophora); clams, oysters and scallops (Bivalvia); squids and octopuses (Cephalopoda); snails and slugs (Gastropoda) (Ponder and Lindberg, 2008). Especially, the transition to non-marine habitats in gastropods resulted in one of the most significant adaptive radiations among animals (Kocot et al., 2013), leading to terrestrial species in several clades, *e.g.* Neritimorpha, Caenogastropoda, and Heterobranchia (Kameda and Kato, 2011).

The phylogeny of Heterobranchia has been in constant change. The clade was established uniting the former clades Allogastropoda and Pentaganglionata and was divided in three groups: Lower Heterobranchia, Opisthobranchia and Pulmonata (Hazprunar, 1985). Cladistic analyses based on morphological characters

showed unresolved relationships within Heterobranchia due to the high variability of euthyneuran anatomical characters (Dayrat and Tillier, 2002). Molecular phylogenies based on single, multi-locus mitochondrial or nuclear, and phylogenomic data also showed different trees (Schrödl, 2014; Wägele et al., 2014). Traditionally, Pulmonata was placed along Opisthobranchia in the clade Euthyneura. This clade possesses most of the species diversity (~30,000 spp.) within Heterobranchia (Kano et al., 2016). Mitochondrial protein-coding genes recovered a monophyletic Opisthobranchia and paraphyletic Pulmonata, whereas multi-locus datasets based on ribosomal genes recovered paraphyletic Opisthobranchia and monophyletic Pulmonata only if some clades previously assumed as Lower Heterobranchia were added to this clade (Kocot et al., 2013) (Figure 2). The clade Panpulmonata was recently established using a combination of ribosomal and mitochondrial markers (Jorger et al., 2010). This clade contains traditional pulmonates plus opisthobranch acochlidians and sacoglossans, and even lower heterobranch pyramidellids. Euopisthobranchia was also established in the same work, reuniting other traditional opisthobranchs such as Cephalaspidea and Anaspidea. Phylogenomic data supported the monophyly of Panpulmonata and Euopisthobranchia and their sister relationship within the monophyletic clade Tectipleura (Kocot et al., 2013; Zapata et al., 2014).

In particular, Panpulmonata are one of the most successful lineages that invaded the intertidal, freshwater and land zones (Figure 3). The transition to the land and freshwater realms occurred in several clades. For instance, Acochlidia possess marine and freshwater species; Ellobioidea, intertidal and terrestrial species; Hygrophila, freshwater species; Stylommatophora terrestrial species; and Systellommatophora, intertidal and terrestrial species (Dayrat et al., 2011; Jorger et al., 2010; Kano et al., 2015; Klussmann-Kolb et al., 2008). However, relationships within Panpulmonata are still unresolved (Figure 4). A recent analysis supported the association of Stylommatophora, Systellommatophora, and Ellobioidea, all of them containing terrestrial snails and slugs (Zapata et al., 2014). Sacoglossa slugs and

Siphonarioidea false-limpets appeared to be the first off shoots in the Panpulmonata. Both clades occur in the marine realm similar to Euopisthobranchia, the sister clade of Panpulmonata, alike to the Lower Heterobranchia, and likely, to the ancestor of all panpulmonates.

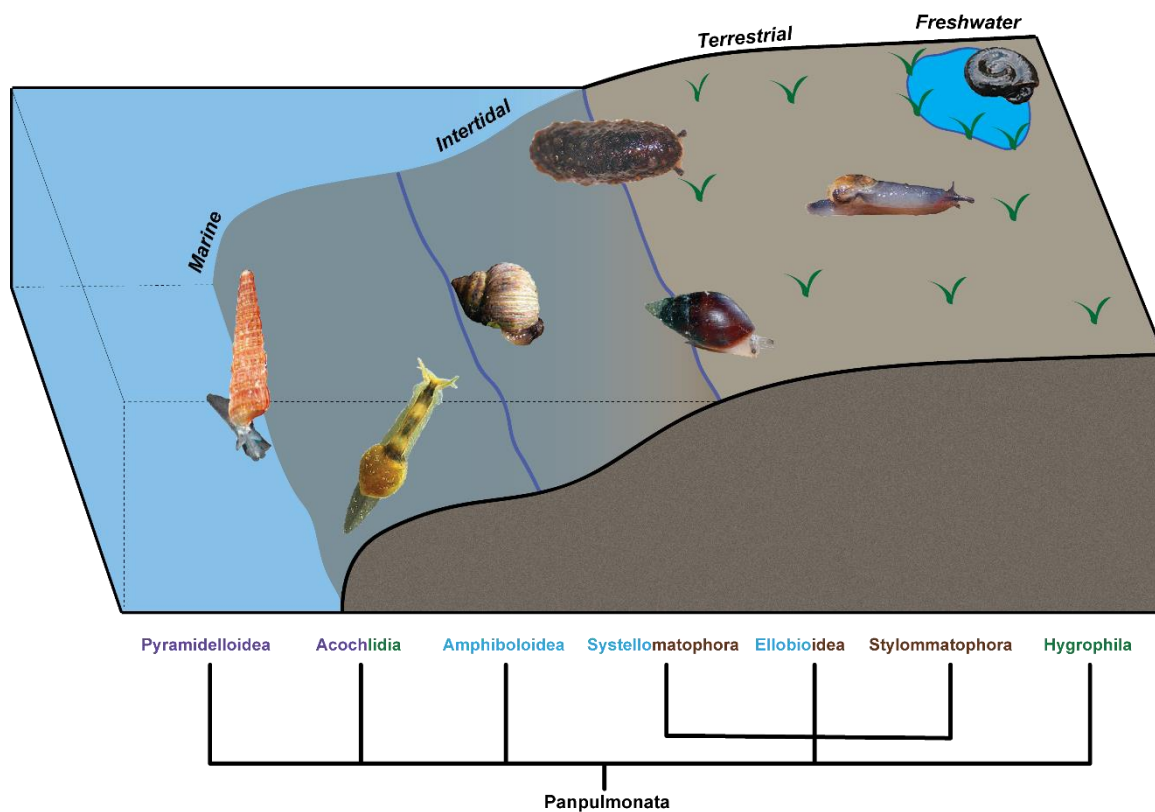


**Figure 2.** Phylogenetic relationships in Heterobranchia.

Consensus tree of the multiple phylogenetic hypotheses of the Heterobranchia based on multi-locus and phylogenomic datasets. Well-supported nodes are represented with dots. Colors represent classic classification of current polyphyletic groups, green: opisthobranchs, blue: lower heterobranchs, pink: pulmonates. Adapted from Wägele et al. (2014).

The Ellobioidea are one of the most intriguing clades within Panpulmonata. They possess a variety of morphologies, *e.g.* terrestrial and intertidal snails, plus intertidal slugs and false-limpets. The Ellobioidea mainly occur in the intertidal and supratidal zones (mangroves, muddy and rocky shores, and salt marshes) on tropical and temperate regions (Martins, 2007). In addition, four ellobiid genera include terrestrial species: *Pythia* Röding, 1798, occurring in tropical rainforests (Martins,

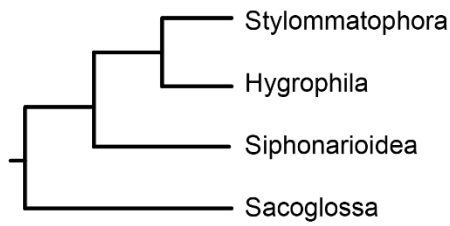
1995); *Carychium* Müller 1773, in Holarctic forests (Weigand et al., 2013); and, *Zospium* Bourguignat, 1856 and *Korozospeum*, Jochum & Prozorova, 2015, in Holarctic karst caves (Jochum et al., 2015). Previous phylogenies, based on morphological (Martins, 2007) and molecular data (Dayrat et al., 2011), were not able to resolve the relationships among ellobiid lineages. I decided to choose this group to investigate the tempo and mode of the land invasions in Panpulmonata (Publication 1, Romero et al., 2016a). Thus, I reconstructed a well-supported phylogeny in order to test if the land invasion occurred multiple times in the different lineages. Also, the new phylogeny allowed me to test if the land invasion had a correlation with past geological or climatic changes, specifically with the Cretaceous-Paleogene boundary (K-Pg) or the lowering of the sea level during the Oligocene (Romero et al., 2016a).



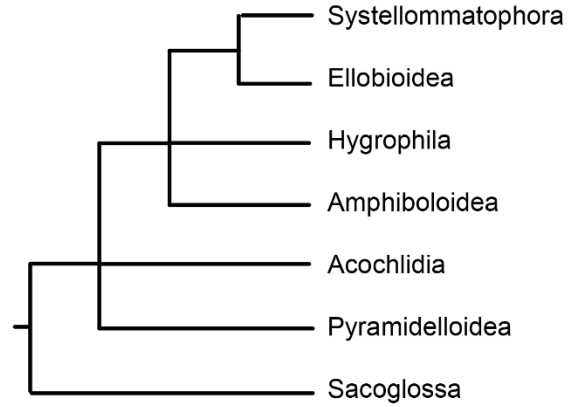
**Figure 3.** The realms of Panpulmonata.

Colors represent the different zones of panpulmonates occurrence. Purple: marine, light blue: intertidal, green: freshwater, brown: terrestrial. Consensus tree based on Romero et al. *under review* and Zapata et al., 2014. Snail images from Schrödl (2014).

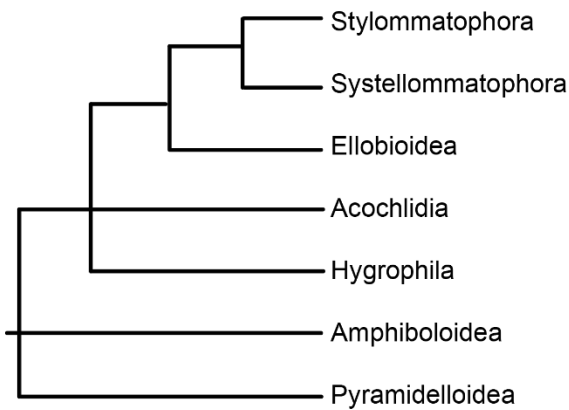
(a) Kocot et al., 2013



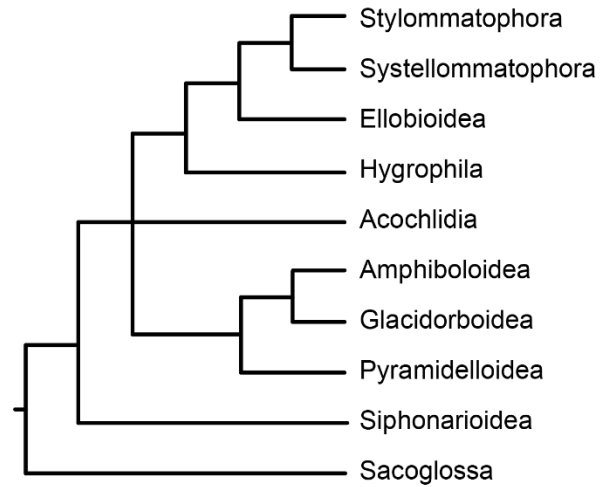
(b) Zapata et al., 2014



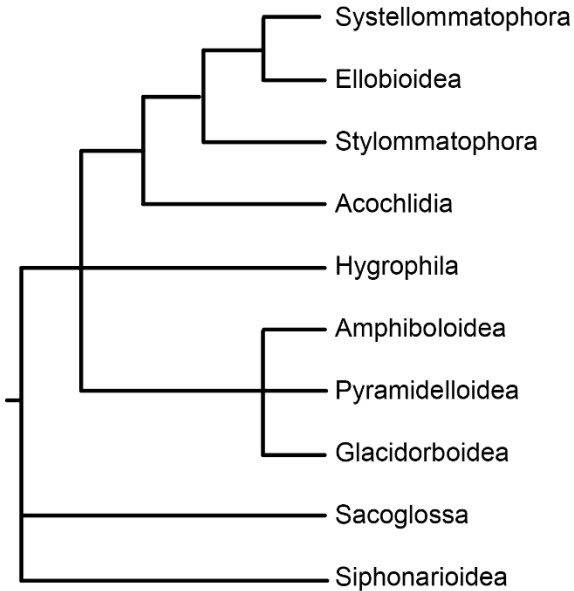
(c) Romero et al., *in review*



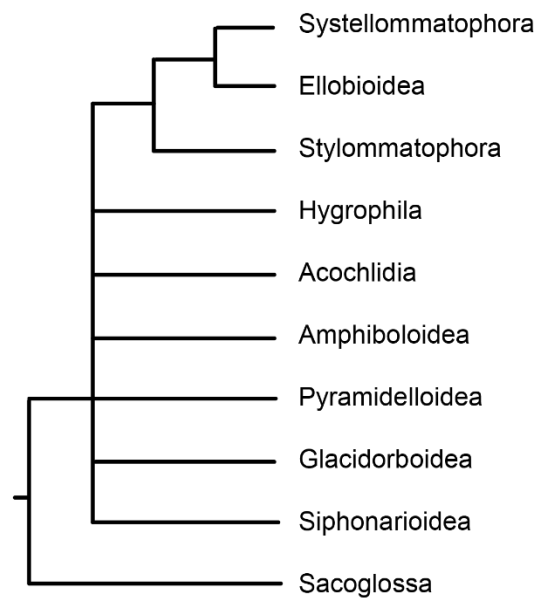
(d) Teasdale et al., *pers. comm.*



(e) Jörger et al., 2010



(f) Schrödl, 2014



**Figure 4.** Phylogenetic relationships in Panpulmonata.

Consensus trees based on phylogenomic studies (a, b, c, d) or multi-locus mitochondrial and nuclear data (e, f).



### 1.3. Molecular adaptations in the context of realm transitions

Novel adaptations may be caused by different mechanisms of gene evolution. For example, by changes in gene expression and in the coding sequence, expansion (gene duplication) or contraction of gene families, and *de novo* gene origination from non-coding sequences (Holloway et al., 2007; Innan and Kondrashov, 2010; Zhou et al., 2008). Several studies, specially focused in vertebrate evolution, employed a comparative approach to reveal adaptations related to evolutionary transitions. For example, mudskippers, amphibious fishes adapted to live on mudflats, showed an expansion of innate immune genes, possibly, to counteract novel pathogens on land; also, they lost a short wavelength-sensitive opsin gene (SWS1) in order to avoid damage produced by UV radiation. Moreover, they possess many V1R genes which produce receptors that bind to small air-borne chemicals and are common in typical terrestrial vertebrates (You et al., 2014). Coelacanths are one of the most primitive lineages from Sarcopterygii, a clade that includes lobe-finned fishes and tetrapods. The expansion of VR1 genes has also been demonstrated in the genome of the coelacanth *Latimeria* (Nikaido et al., 2013). Additionally, sequences that act as enhancers of genes for limb development were found in the coelacanth and in tetrapods, but not in ray-finned fishes (Actinopterygii) (Nikaido et al., 2013). The authors suggested that many genes already present in the primitive Sarcopterygii were later co-opted during the transition to terrestrial environments. A similar study found at least 50 genes present in *Latimeria* that were lost in the tetrapod lineage (Amemiya et al., 2013). These genes were implied, for instance, in fin, eye, ear and kidney development. In the latter study, adaptive evolution in tetrapods was suggested for the carbamoyl phosphate synthase I (CPS1) enzyme, a protein involved in the efficient production of hepatic urea, because the tetrapod and amniote branches showed signatures of positive selection (Amemiya et al., 2013).

Positive selection is a mode of natural selection that drives the increase in prevalence of advantageous traits (Sabeti et al., 2006). Selection can be identified at the population or species level. In the first case, population genetic approaches can

detect ongoing selection, whereas, in the second case, comparative approaches are more suitable to detect past selection (Nielsen, 2005). The principal tool to detect signatures of selection in comparative data is to calculate the non-synonymous/synonymous rate ratio ( $\omega = dN/dS$ ). If there is no selection, synonymous and non-synonymous substitutions should occur at the same rate and we would expect  $dN \approx dS$ . Moreover, positive selection ( $\omega > 1$ ) is expected if non-synonymous substitutions offer a fitness advantage and have a higher fixation probability than synonymous substitutions (Yang et al., 2000). It is possible, as in the case of the CPS1 enzyme, to detect  $\omega > 1$  in specific branches using the branch-site test of positive selection (Yang and dos Reis, 2011). This test compares two models, the null model assumes that neutral evolution occurred in a particular branch (called “foreground”), while the alternative model assumes positive selection in the foreground. In this way, different biological hypotheses about adaptive evolution can be tested statistically. Moreover, the information of amino acid positions under selection can be compared to protein structural or served as a guidance for further experimental work, *e.g.* site-directed mutagenesis (Zhai et al., 2012).

During my literature review about molecular signatures of adaptation related to the realm transition, I found only one publication that addresses this issue in invertebrates (Faddeeva et al., 2015). In this study, the authors compared transcriptomes of springtails, insects, and crustaceans, identifying signatures of adaptive evolution in the Collembola (springtails) and Hexapoda (springtails plus insects) lineages. Various genes related to ion transport, homeostasis, immune response and development appeared under positive selection suggesting their possible role during the evolution on land. The lack of studies focusing on adaptations involved in sea-to-land transitions at the molecular level, intensifies the importance of studying groups that exist in marine and non-marine realms, such as Panpulmonata. Therefore, in my dissertation, I looked for adaptive signatures related to the realm transition in Panpulmonata using two sets of data: mitogenomes (Publication

2, Romero et al., 2016b) and transcriptomes (Manuscript 3, Romero et al. *under review*). I started with mitochondrial genomes because they were better represented in the GenBank database. In addition, mitogenomes encode 13 proteins that are mainly conserved among the metazoan and are under high functional constraints for they intervene in the oxidative phosphorylation (OXPHOS) pathway (da Fonseca et al., 2008). Working with the multi-locus mitochondrial dataset was also my first approach to programming languages required to analyze next-generation sequencing (NGS) data from genomes or transcriptomes, faster and in an efficient way. Then, I expanded the analysis to a higher level exploring adaptive evolution in panpulmonate transcriptomes. Here, I analyzed hundreds of genes and learned how to curate raw high-throughput sequencing data, to *de novo* assemble transcript contigs, to annotate the transcripts and to cluster them in orthologous genes groups. In both Publication 2 and Manuscript 3, I reconstructed phylogenetics trees, and performed the branch-site test of positive selection in terrestrial lineages (Ellobioidea, Stylommatophora) for the mitochondrial data (Romero et al., 2016b), or terrestrial and freshwater lineages (Ellobioidea, Stylommatophora, Acochlidia, Hygrophila) for the transcriptomic data (Romero et al., *under review*)

My dissertation will present results of independent land invasions within Panpulmonata and will list genes and pathways that may have played a role during the adaptation to new non-marine realms. These results can be considered within a higher framework: the topic of convergent and parallel evolution.

#### **1.4. Convergent and parallel evolution**

Biologists have long distinguished between convergent and parallel evolution as labels for the independent origin of phenotypic similarity among populations or species (Arendt and Reznick, 2008). Parallel evolution is defined when the same phenotype evolves multiple times independently within a given species or among closely related species. In this case, it is assumed that similar phenotypic traits are based on the same genes and developmental pathways. Contrastingly, convergent

evolution is assumed when the same phenotype appears in unrelated species, supposedly, via different genes and developmental pathways (Arendt and Reznick, 2008). Cases where closely related taxa used different genetic mechanisms or distantly related taxa used the same pathways to obtain similar phenotypes can complicate the application of both concepts. Furthermore, many authors have used the terms only based on phylogenetic relationships, applying “parallelism” for closely related species and “convergence” for distant relatives, without considering whether similar phenotypes arose from similar or different genetic mechanisms. Others use the terms in reference to whether the genetic mechanism is the same (parallelism) or different (convergence) without taking into account the phylogenetic relatedness (Figure 5a) (Rosenblum et al., 2014).

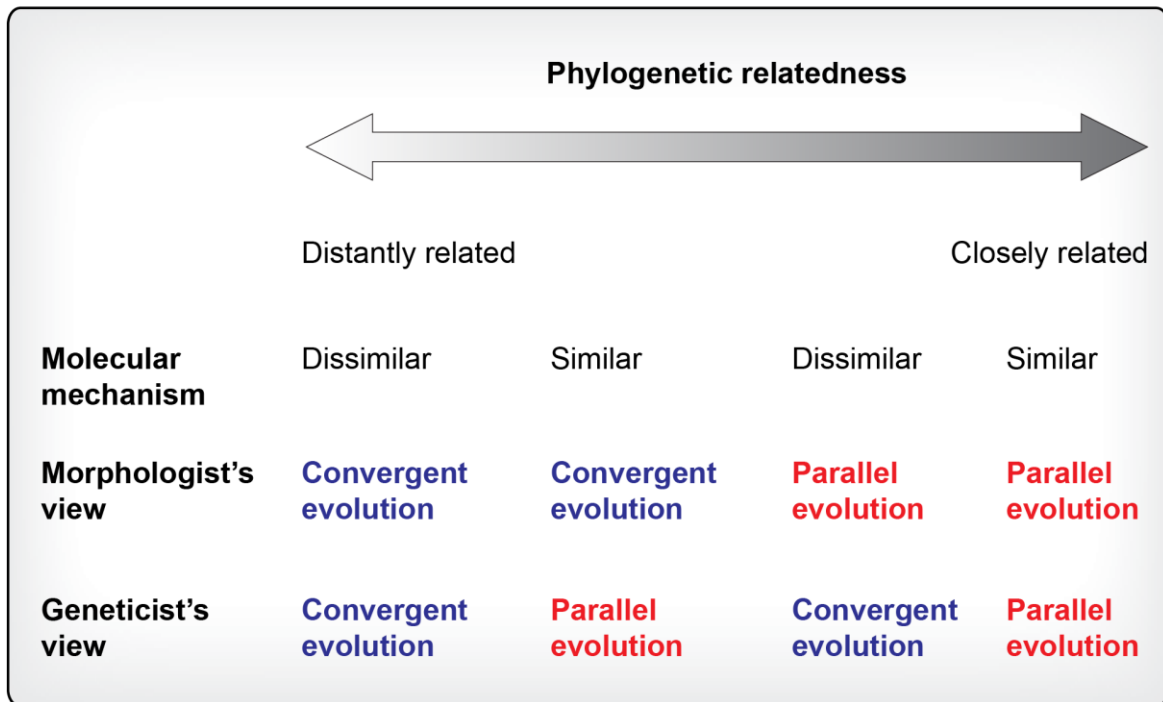
I decided to follow Rosenblum et al. (2014) definitions on convergent evolution (Figure 5b). The authors propose different levels of convergence. First, at the phenotypic level, convergent evolution describes the independent evolution of similar phenotypes. Phenotypes can be similar in close or distant species and can be a product of similar or different the genetic mechanisms. Second, parallel evolution is a special case of convergence at the molecular level, *i.e* the use of shared molecular mechanisms to produce similar phenotypes. They recommend that parallel evolution must be described in a hierarchical level of organization: allele, gene, pathway, or function. The authors also proposed that convergence should be described in a taxonomic context, for instance, convergence within a genus versus convergence within an order.

Perhaps, one of the most documented cases in convergent evolution triggered by a realm transition is the invasion of aquatic environments in mammals. Recent genome-wide screens found in the Cetacea (whales, dolphins and porpoises) several genes under positive selection related to osmoregulation, hypoxia, and DNA repair in dolphins and whales (Yim et al., 2014; Zhou et al., 2013). For example, myoglobin, the main oxygen carrier and storage protein in the muscle, has been associated to tolerance to hypoxia and enhanced oxygen storage during

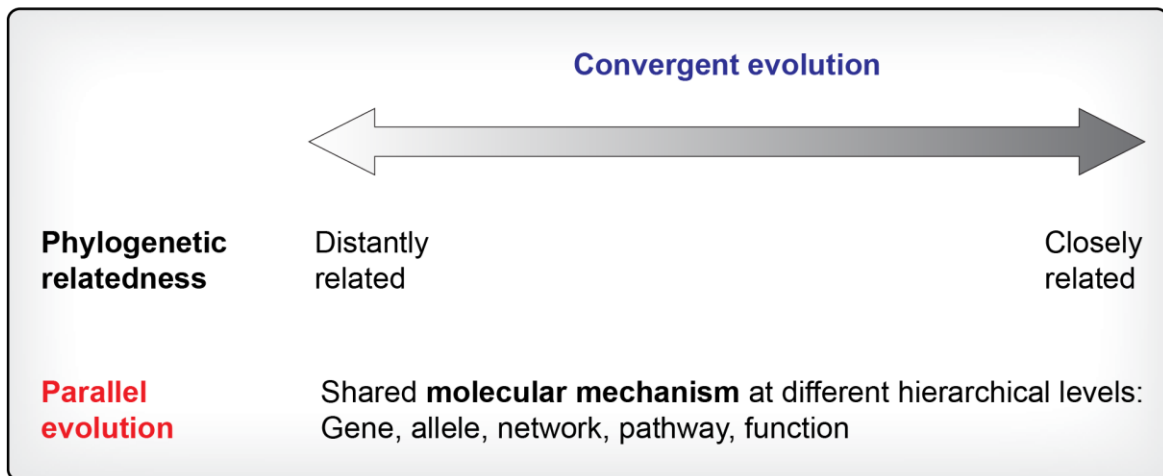
diving. Higher concentrations of myoglobin have been found in cetaceans and pinnipeds (sea lions and seals) and parallel amino acid replacements in the myoglobin from cetaceans, pinnipeds and beavers incremented the protein surface charge probably improving the folding stability and avoiding deleterious self-associations of these molecules (McGowen et al., 2014). Toothed whales (Odontoceti) have developed an echolocation system to communicate and navigate in the marine realm. Prestine, a protein involved in electromotility of the outer hair cells of the cochlea enabling sensitivity to sound, has been found under positive selection in the Odontoceti branch. In addition, parallel amino acid replacements, likely related to sensitivity to higher frequencies associated to echolocation, were shared between the Odontoceti and, surprisingly, Chiroptera bats (McGowen et al., 2014). These examples of parallel evolution show shared molecular mechanisms that are involved in similar phenotypic traits and adaptation to new realms.

My main hypothesis is that convergent terrestrial lineages in Panpulmonata shared molecular mechanisms in order to adapt to the land. Thus, I will explore if parallel evolution occurred in land snails from terrestrial Ellobioidea and Stylommatophora species using mitochondrial and nuclear protein-coding genes. Then, I will analyze adaptive evolution at the gene level. Finally, I will compare my results with recent literature to evaluate similar patterns of adaptive evolution in genes or pathways that have been related to realm transitions.

a) Conflicting definitions of convergence and parallelism.



b) Definitions used in this thesis according to Roseblum et al. (2014).



**Figure 5.** Previous and proposed definitions of convergence and parallelism.

Parallel evolution is a case of convergent evolution that occurs when shared molecular mechanisms produce a similar phenotype and it will be described only in a hierarchical level, from genes to functions. Adapted from Rosenblum et al. (2014).

## 1.5. Thesis questions and objectives

The main purpose of my thesis is to understand the evolution of the terrestrialization in Panpulmonata. Thus, I propose the following objectives in order to comprehend this process:

**1. To study the tempo and mode of the terrestrialization using the panpulmonate clade Ellobioidea (Publication 1).**

Here, I choose the family Ellobiidae which possess intertidal and terrestrial taxa to study the convergent land invasions and enquired these questions:

- a. Were the sea-to-land transitions in the Ellobiidae independent?
- b. Do they occur at the same geological era and were influenced by the similar historical geoclimatic changes?




**2. To analyze adaptive evolution in the context of the terrestrial invasion within Panpulmonata (Publication 2 and Manuscript 1).**

Here, I used mitogenomic and transcriptomic data from marine, intertidal, freshwater and terrestrial panpulmonates to test parallel signatures of adaptation at the molecular level during the land invasion and considered these questions:

- a. Are there signatures of positive selection in the mitogenomes or transcriptomes of terrestrial panpulmonates?
- b. What is the function of the genes under positive selection and in which pathways are they involved?

My objectives aim to answer different questions related to conquest of the land realm by panpulmonates using tools from bioinformatics, phylogenetics, and molecular evolution. Furthermore, I will discuss the results in the context of convergent evolution, and show parallel molecular mechanisms that may have involved during the transition to the new realms. A conceptual summary of the thesis is provided in the next page (Figure 6).

# Evolution of the terrestrial invasion in Panpulmonata: Molecular adaptations in the context of realm transitions

	Publication 1	Publication 2	Manuscript 1
<b>Title of the manuscript or publication</b>	Molecular phylogeny of the Ellobiidae (Gastropoda: Panpulmonata) supports independent terrestrial invasions 	Positive selection on panpulmonate mitogenomes provide new clues on adaptations to terrestrial life 	Panpulmonate transcriptomes reveal genes involved in the adaptation to terrestrial and freshwater habitats in gastropods 
<b>Genetic markers used in each work</b>	Multi-locus protein-coding and ribosomal mitochondrial and nuclear genes	Protein-coding mitochondrial genes	Mainly nuclear protein-coding gene transcripts
<b>Sequencing methodology</b>	Sanger sequencing of gene fragments	Sanger sequencing of complete mitochondrial genomes	Next-generation sequencing (NGS): RNAseq
<b>Main analyses</b>	Multi-locus alignments Phylogenetic reconstruction Divergence times	Multi-locus alignments Phylogenetic reconstruction Branch-site test of positive selection	Transcriptome assembly Clustering of orthologous genes Phylogenomic reconstruction Branch-site test of positive selection Functional annotation
<b>Main questions</b>	Were the sea-to-land transitions in the Ellobiidae independent?  Do they occur at the same geological time, and were influenced by the same historical geoclimatic changes?	Are there signatures of adaptive evolution in the mitogenomes of terrestrial panpulmonates?  Are the signatures present in other taxa that transitioned between different habitats?	Are there signatures of adaptive evolution in the transcriptomes of terrestrial or freshwater panpulmonates?  What is the function of the genes under positive selection and in which pathways are they involved?
<b>Patterns and processes</b>	<b>Tempo and mode of the terrestrialization</b>	<b>Adaptive evolution related to the terrestrialization</b>	<b>Adaptive evolution related to the terrestrialization</b>



**Figure 6.** Conceptual summary of the thesis.



## Chapter 2: GENERAL DISCUSSION

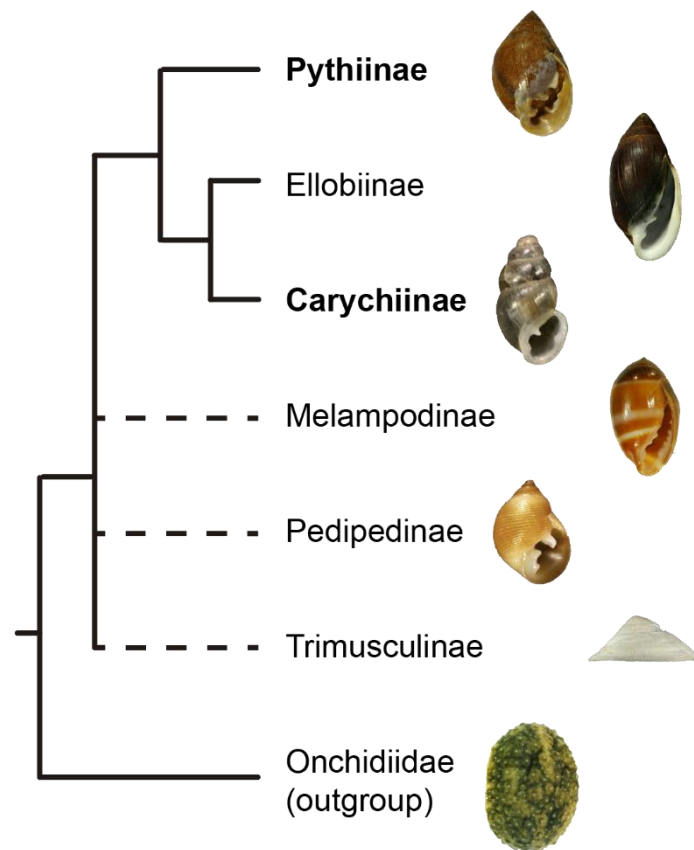
My work includes two peer-reviewed publications (Romero et al., 2016a; Romero et al., 2016b) and one manuscript (Romero et al., *under review*). In this chapter, I will synthesize the work presented in these three reports, try to answer the questions presented in the objectives section (Figure 6), and evaluate convergent evolution during the conquest of the land by panpulmonates.

### 2.1. Tempo and mode of the terrestrialization within Panpulmonata

In Publication 1 (Romero et al., 2016a), I studied the land invasion in the family Ellobiidae, the only family within the clade Ellobioidea. The latest classification divides this group in five subfamilies: Carychiinae, Ellobiinae, Melampodinae, Pedipedinae, and Pythiinae (Martins, 2007). Terrestrial species, whose development and reproduction are completely independent from aquatic environments, have been reported in the subfamily Carychiinae (Weigand et al., 2013) and in some species of the subfamily Pythiinae (Kano et al., 2015). However, phylogenetic analyses based on morphological or molecular markers were not able to resolve the evolutionary relationships in the Ellobiidae. Previous trees particularly failed to support the monophyly of the subfamilies (Dayrat et al., 2011).

My work expanded the number of genetic markers and samples in comparison with previous works. I proposed a new tree based on six molecular markers: mitochondrial 16S rRNA, 12S rRNA, and cytochrome oxidase I (*coi*); nuclear 18S rRNA, 28S rRNA, and histone 3 (H3). In addition, I expanded the taxon sampling using almost all of the Ellobiidae genera: Carychiinae (2 genera used/3 accepted genera), Ellobiinae (5/5), Melampodinae (3/5), Pedipedinae (2/4), and Pythiinae (8/8). The new phylogeny (Figure 7), reconstructed using maximum likelihood and Bayesian methods, supported the monophyly of the subfamilies Carychiinae, Ellobiinae, and Pythiinae. In addition, it supported the sister relationship within Ellobiinae and Carychiinae.

My results clearly supported that terrestrial ellobiids belong to different subfamilies, and that these subfamilies are reciprocally monophyletic. Thus, I propose that the land invasion occurred independently in different lineages (Carychiinae and Pythiinae), probably via the intertidal zone (rocky shores, estuaries and mangroves). The intertidal zone displays a wide range of variation in physical factors; therefore, organisms living in this kind of habitat must adapt to changing conditions and stress (Dayrat et al., 2011). In particular, ellobiids are able to breath oxygen from the air using their pulmonary cavity, and tolerate desiccation in their environment (Little, 1990). These adaptations are already present in intertidal lineages and should have helped the success during the invasion of the land.

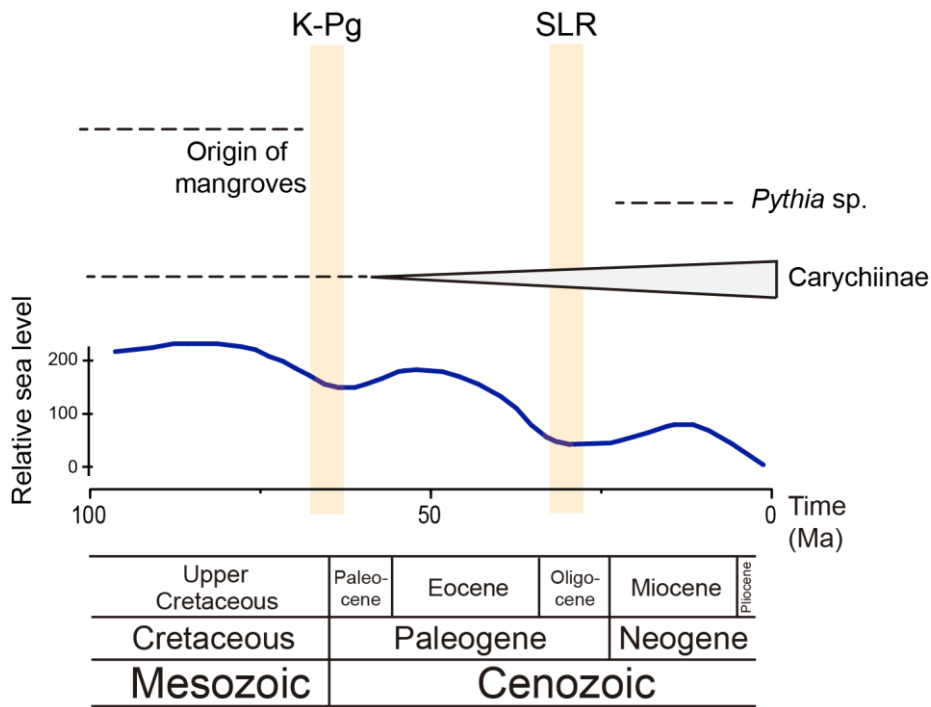


**Figure 7.** Phylogenetic relationships of the Ellobiidae (Heterobranchia, Ellobioidea) Consensus tree adapted from Romero et al. (2016a). Subfamilies highlighted in bold possess terrestrial species. Dashed lines represent subfamilies not supported in the phylogenetic reconstructions. Shell images (not to scale) from the Natural History Museum Rotterdam.

The terrestrialization in Ellobiidae occurred at least twice in different geological times (Figure 8). First, terrestrial Carychiinae likely originated during the Mesozoic (Upper Cretaceous) and diversified after the Cenozoic after the Cretaceous-Paleogene (K-Pg) boundary. The K-Pg event, caused by the impact of an asteroid, triggered a mass extinction that generated open niches for the diversification of different lineages of land snails, reptiles (including birds), and mammals (Breure and Romero, 2012; Proches et al., 2014). Mangroves also appeared during the Upper Cretaceous (Ellison et al., 1999) and could have provided transitional habitats in the intertidal zone facilitating the land invasion (Romero et al., 2016a). Provided that Carychiinae live in Holarctic forests (Weigand et al., 2013) and karst caves (Jochum et al., 2015), I suggest that they invaded the land first through the intertidal zone and then colonized their current habitat via active and passive dispersal. The origin of the Carychiinae seems to be long before the K-Pg, so there is no strong support that this event influenced the terrestrialization.

Second, the terrestrial *Pythia* species probably appeared and invaded the land during the Miocene. Provided that other species from the same genus are supratidal, I also suggest that the invasion occurred first through the intertidal zone and that they then actively dispersed to conquer tropical rainforests close to the sea achieving a life cycle totally independent from the marine environment. The invasion of the intertidal zone could have been facilitated by the regression of the sea level during the Oligocene. However, as the terrestrial *Pythia* species appeared later, there is not supported correlation between the land invasion and the marine regression.

Moreover, I analyzed the diversification in the Ellobiidae using lineage-through-time plots (LTT) and diversification models, and did not find a significant correlation among historical geoclimatic changes and the diversification rates within this family (Romero et al., 2016a). The addition of more ellobiid species and the use of better diversification models along with shell fossils would be useful to understand the origin and diversification in the Ellobiidae.



**Figure 8.** Time-scale of the terrestrialization in the Ellobiidae.

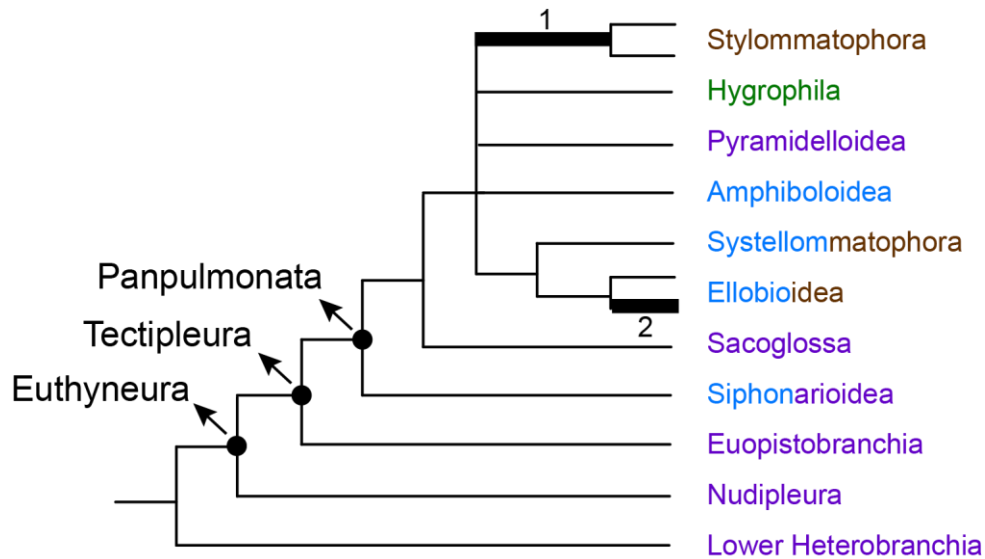
The possible origin of the terrestrial Ellobiidae (subfamily Carychiinae and *Pythia* sp.) is plotted along with different geoclimatic events such as the origin of the mangroves during the Upper Cretaceous, the Cretaceous-Paleogene boundary (K-Pg), and the sea level regression (SLR) during the Oligocene. Adapted from Romero et al. (2016a).

## 2.2. Adaptive evolution in the context of terrestrialization

### 2.2.1. Mitochondrial genomes

In Publication 2 (Romero et al 2016b), I generated new mitochondrial genomes from terrestrial panpulmonates. The mitochondria from *Carychium tridentatum* Risso, 1826 (Ellobioidea), *Arion rufus* Linnaeus, 1758 and *Helicella itala* Linnaeus, 1758 (both, Stylommatophora) were obtained using both Sanger sequencing and 454 pyrosequencing. These mitochondria were compared them with 47 other mitogenomes from all major panpulmonate lineages. Initially, I focused on mitochondrial evolution because mitochondria possess relatively few protein-coding genes (13) that are deeply involved in the oxidative phosphorylation (OXPHOS) pathway.

This means that amino acid changes in these genes are under high selective pressure for they can directly influence the metabolic performance of the cell (da Fonseca et al., 2008).



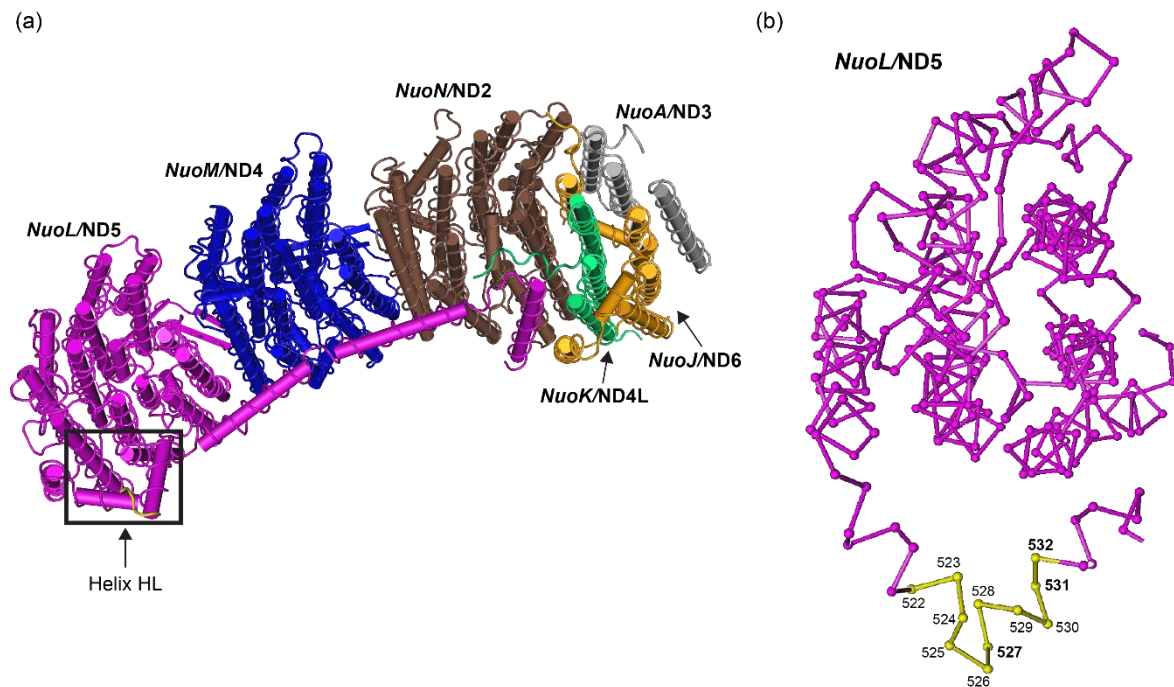
**Figure 9.** Consensus tree of the Euthyneura based on mitochondrial protein-coding genes.

Positive selection was found in the branch leading to stylommatophoran land snails (1) and in the terrestrial ellobiid *Carychium* (2). Colors represent the different realms where species occur: marine (purple), intertidal (light blue), freshwater (green), and terrestrial (brown). Adapted from Romero et al. (2016).

I applied the branch-site test of positive selection to the mitochondrial euthyneuran tree (Figure 10). I was particularly interested in the terrestrial lineages from Ellobioidea (*Carychium*) and Stylommatophora. Two genes appeared under positive selection in both branches: *nad5* (NADH dehydrogenase subunit 5) and *cob* (cytochrome *b*). These genes belong to the mitochondrial respiratory complex I and III, respectively (Publication 2, Romero et al., 2016b).

Positive selection has been consecutively reported in the mitochondrial complex I (Figure 10a) from different vertebrates such as fishes and rodents (Garvin et al., 2015). This complex participates in the oxidation of NADH (reduced nicotina-

mid adenine dinucleotide) and the transfer of electrons to the coenzyme ubiquinone, which passes these electrons to the respiratory complex III. Electron transfer is associated with proton movement across the membrane and it is driven by proton pumps. This process may be coordinated by the lateral helix HL which also provides structural stability needed for long range energy transmission (Efremov and Sazanov, 2011; Wirth et al., 2016).



**Figure 10.** Three-dimensional structure of the respiratory complex I from *Escherichia coli* (PDB ID: 3RKO).

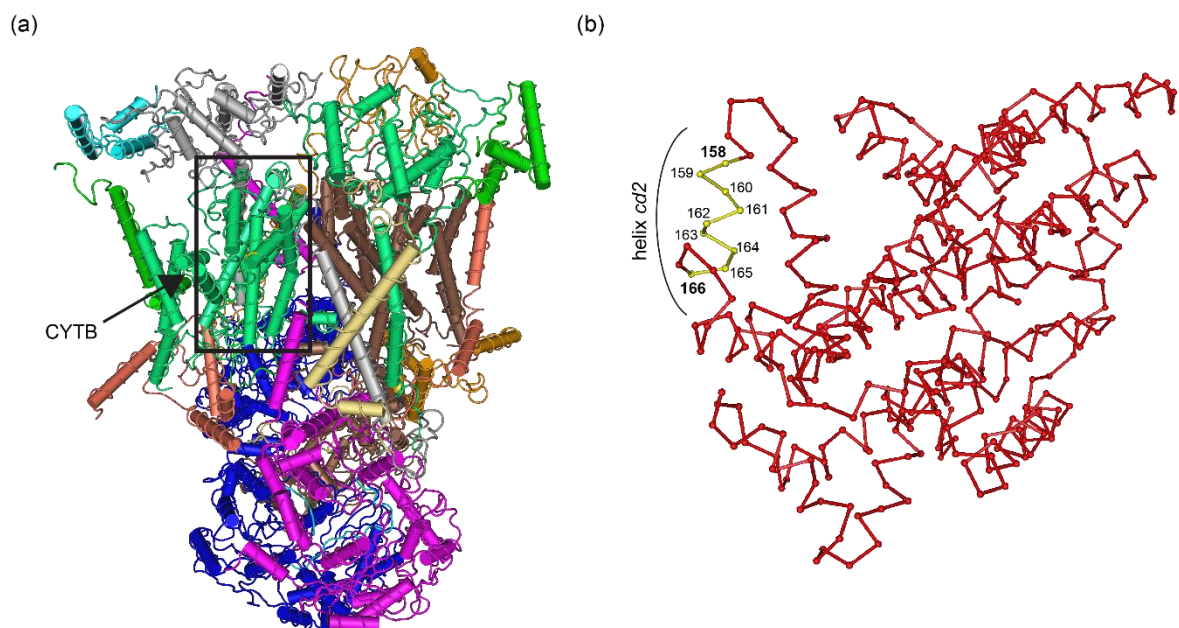
(a) Homologous mitochondrial proteins (NADH) are shown for every subunit. Alpha helices are drawn as tubes. (b) A section of the ND5 subunit showing the first portion of the helix HL. Amino acid positions that appeared under positive selection in different taxa are highlighted in yellow (Garvin et al., 2015; Romero et al., 2016b; Tomasco and Lessa, 2011). Positions under positive selection in terrestrial panpulmonates (*Carychium tridentatum* and the branch leading to Stylommatophora) are shown in bold. The 3D structure was obtained from GenBank and visualized in the software Cn3D.

I found signatures of positive selection on three sites from the ND5 helix HL in the terrestrial panpulmonate branches (Figure 9 and 10b). These sites are homologous to positions 527, 531 and 532 in *Escherichia coli* (Figure 10b) and are situated in a region that has been under positive selection in cetaceans, salmon, otariids and caprines (Garvin et al., 2015). In particular, position 527 appeared under positive selection in cetaceans, subterranean rodents and in terrestrial snails (Romero et al., 2016b; Tomasco and Lessa, 2011). It has been suggested that mutations in the proton pumping proteins may influence fitness because they are deeply involved in the respiratory control (translocation of H<sup>+</sup>) and can alter the production of reactive oxygen species (ROS) (Garvin et al., 2011).

In addition, I found positive selection in the mitochondrial cytochrome *b* (*cob*) from the terrestrial panpulmonate branches, a protein part of the respiratory complex III (Figure 11a). The protons released at this level also contribute to create the electrochemical gradient in the intermembrane space of the mitochondria (Garvin et al., 2015). The position under positive selection, 158 in the *Bos taurus* protein structure (Figure 11b), is located on the *cd2* helix and has also been implicated in the creation of the proton gradient necessary for the production of ATP (McClellan et al., 2005).

Cytochrome *b* has also been under selection in vertebrates such as bats and whales. In bats, amino acid changes in the OXPHOS proteins were associated with the adaptation to fly (Shen et al., 2010). The authors suggested that as flying requires an increment in the metabolic rate compared to moving on land, adaptive evolution favored mutations in the mitochondrial genes that are involved in energy metabolism. Additionally, whales presented signatures of positive selection in this position in comparison with artiodactyls, their sister group. In this case, amino acid changes in CYTB were probably a response to new metabolic demands during cetacean cladogenesis during the transition from land to water habitats (McClellan et al., 2005).

Given that the same positions were under selection, in both *nad5* and *cob*, in terrestrial panpulmonate branches and in other vertebrates that transitioned from different realms (whales or subterranean rodents), and that the respiratory complex genes were also under selection in bats that transitioned from moving on land to flying in the air, it is possible that parallel adaptive evolution also occurred in the terrestrial panpulmonate OXPHOS genes in order to cope with the new energetic demands on land, *e.g.* to move and sustain the body mass without the aid of the buoyancy force.



**Figure 11.** Three-dimensional structure of the respiratory complex III from the domestic cow *Bos taurus* (PDB ID: 1BGY).

(a) The complex consists of 11 dimeric subunits corresponding to the cytochromes b and c1. The section in the box corresponds to a subunit of the cytochrome b (CYTB). Alpha helices are drawn as tubes. (b) A section of the CYTB showing the *cd2* helix highlighted in yellow. Positions under positive selection are shown in bold: Cetaceans, position 158 and position 166 (McClellan et al., 2005); terrestrial panpulmonates (*Carychium tridentatum* and the branch leading to Stylommatophora), position 158 (Romero et al., 2016b). The 3D structure was obtained from GenBank and visualized in the software Cn3D.



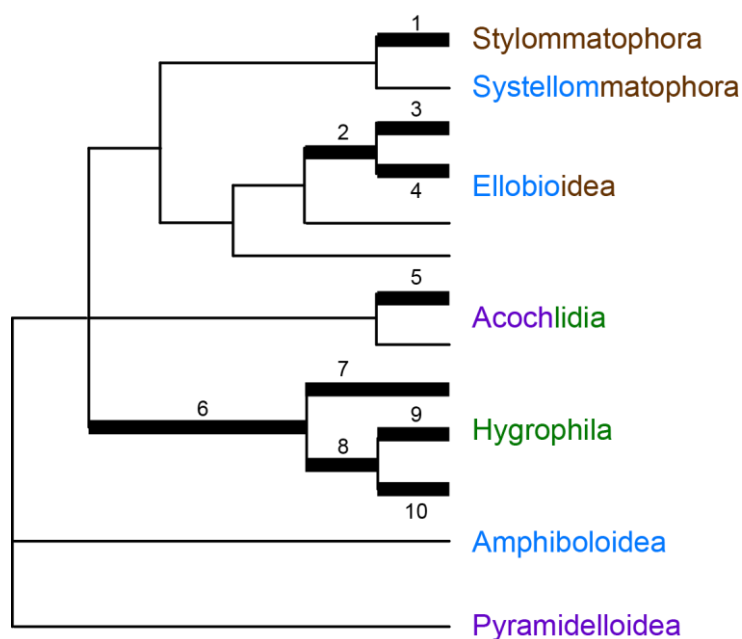
Finally, I found that the changes alter the equilibrium constant physicochemical property of the amino acids (ionization of COOH). This property may influence the overall metabolic efficiency reducing the production of reactive oxygen species (ROS). ROS reduction has been linked to a tolerance to desiccation in supratidal algae (Flores-Molina et al., 2014), and to the increase of individual longevity (Beckstead et al., 2009). Thus, I propose that adaptations in the mitochondrial genes were triggered by the necessity to deal with new energetic demands and to tolerate new abiotic stress factors in the terrestrial realm.

### 2.2.2. Transcriptomes

In Manuscript 1 (Romero et al., *under review*), I expanded the search of genes under positive selection in the terrestrial panpulmonates using transcriptome data. In this dataset, terrestrial lineages were represented by *Carychium* sp., *Pythia pachyodon* Pilsbry and Hirase, 1908 (Ellobioidea), and *Arion vulgaris* Moquin-Tandon, 1855 (Stylommatophora), while freshwater lineages were represented by *Strubellia wawrai* Brenzinger, Neusser, Jörger and Schrödl, 2011 (Acochlidia), *Biomphalaria glabrata* Say 1818, *Planorbarius corneus* Linnaeus, 1758 and *Radix balthica* Linnaeus, 1758 (Hygrophila). In this work, I *de novo* assembled transcriptomes from these and other panpulmonates resulting in 50,000 – 100,000 transcripts per species, of which 10 – 15% were functionally annotated. These transcripts were clustered in 702 orthologous groups, of which 382 were used to reconstruct the phylogenetic relationships within Panpulmonata (Figure 12). Again, I applied the branch-site test of positive selection to the terrestrial lineages from Ellobioidea and Stylommatophora. I also applied it to the freshwater lineages from Hygrophila and Acochlidia.

Different sets of genes appeared under positive selection in land and freshwater snails. In land snails (Fig. 12; 1, 3, 4), I identified a gene involved in the uptake of carbohydrates (sodium-glucose cotransporter). In this case, I suggest that improving glucose absorption could be an adaptive response to increased energetic

demands in the terrestrial realm as previously discussed in the mitochondrial section. In addition, one gene selected in *Pythia* (Fig 12; 3) is linked to the tyrosine metabolism. In invertebrates, tyrosine is ligated to the element iodine during the production of thyroid hormones (TH's) (Heyland et al., 2006). Remarkably, a previous study noted that TH's are necessary for the development of the lung in vertebrates. They influence the expression of fibroblast growth factors and bone morphogenic proteins required for differentiation of the thyroid gland and lung tissue from foregut mesoderm during embryo development (Crockford, 2009).



**Figure 12.** Unrooted tree showing the relationships within Panpulmonata based on 382 orthologous groups.

Positive selection was tested in terrestrial branches (1: *Arion*; 2: the branch leading to terrestrial ellobiids; 3: *Carychium*; 4: *Pythia*), and in freshwater branches (5: *Strubellia*; 6: the branch leading to freshwater Hygrophila species; 7: *Radix*; 8: the branch leading to *Planorbarius* and *Biomphalaria*; 9: *Planorbarius*; 10: *Biomphalaria*). Colors represent the different realms in which species occur: marine (purple), intertidal (light blue), freshwater (green), and terrestrial (brown). Adapted from Romero et al. (*under review*).

Land snails needed to breath air in the terrestrial realm. Thus, they modified their respiratory system by losing gills and altering the inner surface of the mantle into a lung (Dayrat et al., 2011). Therefore, I propose that, in mollusks, the tyrosine pathway may be linked to the development of novel gas-exchange surfaces (lungs) that facilitated the conquest of the land.

In freshwater snails, I identified genes related to stress tolerance. For example, in *Strubellia* (Fig. 12; 5), one gene under positive selection encodes structural ribosomal protein. Adaptive evolution during the realm transition has been suggested in ribosomal proteins in freshwater plants. Given that the ribosomal machinery is salt-sensitive, and that freshwater animals are hypertonic, whereas marine animals are hypotonic, mutations at this level could have triggered tolerance to new osmotic conditions (Wissler et al., 2011). Another gene under positive selection in this taxon encodes the cytochrome P450. This protein is involved in the metabolism of xenobiotics such as carcinogens or environmental pollutants. Adaptive evolution in this protein has been reported for insects in response to plant defense and insecticides. Also, cytochrome P450 appeared under selection in terrestrial hexapods in comparison to other marine arthropods (Faddeeva et al., 2015). Thus, it is likely that mutations in cytochrome P450 in mollusks may have been useful to metabolize or tolerate new organic pollutants present in the freshwater realm.

In *Radix* (Fig 12; 7), I found a DNA methyltransferase gene under positive selection. This protein belongs to the DNA repair pathway against damage produced by ultraviolet (UV) radiation (Lee et al., 2015). DNA repair genes have been shown under positive selection in terrestrial hexapods in comparison to marine ones (Faddeeva et al., 2015), in mudskipper fishes living in transitional habitats like mudflats (You et al., 2014), and in antelopes living in high altitude environments (Ge et al., 2013). Provided that different UV radiation could differentially affect marine and freshwater realms, mutation in DNA genes could have helped in the maintenance of genomic integrity and tolerance to new conditions of radiation in the freshwater realm.

In summary, genes in some terrestrial lineages were related to energy metabolisms and the differentiation of gas-exchange tissues; while in some freshwater lineages, they were related to tolerance to different types of abiotic stressors, such as osmotic pressure, UV radiation and xenobiotics. The occurrence of multiple terrestrial and freshwater lineages in Panpulmonata highlights this group as suitable model to study convergent evolution. It is likely that a wider screening of orthologous clusters in panpulmonates would result in many other pathways related to the conquest of the terrestrial realm.

### 2.3. Convergent evolution

Here, I present a hierarchical evaluation of possible parallel changes that occurred in terrestrial snails as a response to the land invasion. First, in the mitochondria, there were two different branches that appeared under positive selection: *Carychium* and the branch leading to Stylommatophora. At the gene level, *nad5* and *cob* showed significant nonsynonymous changes compared to other mitochondrial genes according to the branch-site test of positive selection (Publication 2; Romero et al., 2016b). At the amino acid level, specific positions under selection in the terrestrial branches do not present identical amino acids for all taxa nor similar amino acid properties, *i.e.* they can be hydrophobic or hydrophilic (Table 1). Therefore, parallel changes occurred in both genes but they did not change codon or amino acid information in an identical way. Examples of parallel changes that resulted in identical amino acids were reported for whales and bats in the prestin gene related to echolocation (McGowen et al., 2014). However, when other hearing genes such as *Tmc1* (transmembrane cochlear-expressed gene 1) and *Pjvk* (Pejvakin) were analyzed in dolphins and bats, results were mixed; some positions under selection shared identical amino acids, while others showed different amino acids (Davies et al., 2012). At the pathway level, both genes, *nad5* and *cob*, are part of the mitochondrial respiratory complexes involved in the oxidative phosphorylation and production of ATP.

Second, results from the transcriptomes showed positive selection in several different genes but most of them occurred on single branches (*Carychium*, *Radix*, *Strubellia*) suggesting that molecular adaptations to terrestrial or freshwater realms occurred independently. It must be noted that in comparison to mitochondrial genomes, which possesses few genes present that are conserved and easily identified; the nuclear genome has many thousands of genes that can be expressed and, in some cases, are more difficult to identify or annotate. As I was able to analyze approximately 700 orthologous genes out of approximately 20 000 genes genome wide, it is likely that other genes not included in my work could show signatures of positive selection in the land or freshwater lineages. Moreover, it is possible that adaptive evolution occurred in similar genes in the mitochondrial genomes because there are fewer possibilities where mutations can occur (13 protein-encoding genes), and selection can act on. In addition, molecular pathways in the mitochondria are under high functional constraints; therefore, it is expected that many non-synonymous changes result in deleterious mutations. Thus, if some mutations can improve the metabolic performance, they would probably tend to be used in different organisms adapting to high demanding habitats. In contrast, there are many more genes in the nuclear genome involved in even more metabolic pathways. Therefore, parallel molecular adaptations could occur in higher levels such as pathways or networks, and not necessarily at the gene, codon, or amino acid level.

One example of parallelism at a higher level of organization is the interactions between mitochondrial and nuclear genomes and proteins. I provided several examples of adaptive evolution in mitochondrial genes from the OXPHOS pathway. There are other mitochondrial metabolic pathways, such as the tricarboxylic acid cycle (Krebs cycle). This cycle comprises by proteins encoded in the nuclear genome that relocate in the mitochondrial matrix. One of these proteins (IDH2) has been shown under positive selection possibly as an adaptation to increased energy demands required by bigger brains in cetaceans and mammals or to flight in bats

(Ai et al., 2014). Thus, despite the fact that different genes are under positive selection in both genomes, their metabolic pathways converge in the same purpose: to produce more energy for the cell. Given that mitochondrial and nuclear evolution are tightly connected, it would not be surprising to find more examples of adaptive evolution in different genes and pathways that interconnect in higher levels of organization.

**Table 1.** Amino acid classification of each site under positive selection in the mitochondrial genome of terrestrial panpulmonates\*.

(a) *nad5*

Taxon	<i>Nuol</i>		<i>Nuol</i>		<i>Nuol</i>		<i>Nuol</i>	
	527	Classification	528	Classification	531	Classification	532	Classification
<i>Aegista</i>	I	Hydrophobic	N	Polar neutral	S	Polar neutral	N	Polar neutral
<i>Arion</i>	N	Polar neutral	Y	Hydrophobic	S	Polar neutral	S	Polar neutral
<i>Camaena</i>	S	Polar neutral	F	Polar neutral	M	Hydrophobic	N	Polar neutral
<i>Carychium</i>	H	Polar positive	M	Hydrophobic	E	Polar negative	H	Polar positive
<i>Dolicheulota</i>	F	Polar neutral	P	Hydrophobic	I	Hydrophobic	N	Polar neutral
<i>Gastrocopta</i>	S	Polar neutral	H	Polar positive	S	Polar neutral	S	Polar neutral
<i>Helicella</i>	V	Hydrophobic	P	Hydrophobic	S	Polar neutral	S	Polar neutral
<i>Mastigeulota</i>	V	Hydrophobic	L	Hydrophobic	V	Hydrophobic	N	Polar neutral
<i>Pupilla</i>	S	Polar neutral	H	Polar positive	G	Hydrophobic	S	Polar neutral
<i>Vertigo</i>	S	Polar neutral	M	Hydrophobic	S	Polar neutral	S	Polar neutral
Baleanopteridae	F	Polar neutral	S	Polar neutral	S	Polar neutral	T	Hydrophobic

(b) *cob*

Taxon	<i>Bos cob</i>	
	158	Classification
<i>Aegista</i>	T	Hydrophobic
<i>Arion</i>	G	Hydrophobic
<i>Camaena</i>	E	Polar negative
<i>Carychium</i>	A	Hydrophobic
<i>Dolicheulota</i>	E	Polar negative
<i>Gastrocopta</i>	G	Hydrophobic
<i>Helicella</i>	D	Polar negative
<i>Mastigeulota</i>	D	Polar negative
<i>Pupilla</i>	E	Polar negative
<i>Vertigo</i>	G	Hydrophobic
Baleanopteridae	N	Polar neutral

\*For comparison, the same amino acid position in cetaceans (Baleanopteridae) from Garvin et al. (2015) is shown. Position numbering based on the homologous sequences in *E. coli* (*nad5*) and *B. taurus* (*cob*).

## 2.4. Conclusions and outlook

In my dissertation, I tried to answer a number of questions related to the land invasion in Panpulmonata. First, I found that land transitions are not uncommon within this group. For example, in the superfamily Ellobioidea, the land invasion occurred at least twice in independent branches: the Carychiinae and a subclade of the genus *Pythia*. In addition, both realm transitions occurred in different geological periods, Cretaceous and Miocene, respectively. Second, I found different genes in the mitochondrial and nuclear genomes that could be involved in the adaptation to the realm transition. These genes belong to pathways that may have facilitated the success during the transition out of the marine realm improving the overall energy metabolism and tolerance to new biotic and abiotic stress factors. Especially, parallel changes in the mitochondrial genes *nad5* and *cob* in invertebrates and even in vertebrates suggest a general pattern of convergent evolution to deal with novel metabolic demands in a new realm.

My thesis contributed new evidence about the tempo and mode of the terrestrialization in mollusks, as well as new insights to understand the genomic basis of the adaptation during the sea-to-land transitions. Remarkably, it included mollusks into the current debate on the molecular adaptations related to the conquest of the land, a discussion that has focused only on vertebrates and insects. My results showed that panpulmonates are a suitable group to study convergent and parallel evolution during the conquest on the land, and will definitively be useful for future comparative genomic analyses and experiments to study the key molecular adaptations during the realm transitions.

Hence, positions that appeared under selection and seem promising to understand the molecular adaptations could be tested using experimental procedures, to test the influence of nonsynonymous changes in different metabolic pathways. Furthermore, other approaches using more complete genome information can be used along with data from positive selection analyses. An example from mudskippers showed that studying the expansion or loss of gene families and the up- or

down-regulation of transcription factors could also illuminate the evolution of adaptations to non-marine habitats (You et al., 2014). In addition, recent studies showed that freshwater eukaryotes have higher rates than marine ones (Mitterboeck et al., 2016a) and, in the case of insects, terrestrial species have higher rates than freshwater ones (Mitterboeck et al., 2016b). Moreover, several genes linked to lipid metabolism, muscle physiology and sensory systems accelerated their evolutionary rates from different marine mammal lineages that transitioned from land to an aquatic realm (Chikina et al., 2016).

Indeed, accelerated evolution of the mtDNA has been found in several terrestrial snail and slug species (Pinceel et al., 2005; Thomaz et al., 1996), and freshwater snails (Nolan et al., 2014; Pfenninger et al., 2006). Furthermore, I also found that nonsynonymous changes in the mitochondrial protein-coding genes are more frequent in freshwater Hygrophila and even more frequent in terrestrial Stylomatophora in comparison to other intertidal and marine panpulmonates (Romero et al., 2016b). Thus, it would be necessary to extend the analysis to nuclear data in order to evaluate if evolutionary rates are also increased in terrestrial or freshwater panpulmonate genomes.



## ACKNOWLEDGMENTS

First of all, I would like to thank Dr. Markus Pfenninger for giving me the possibility to conduct my thesis in his group and to encourage me to finish this project. I really appreciate his comments during the writing periods, and his willingness to discuss new ideas that improved our research. In addition, I would like to thank Dr. Barbara Feldmeyer; her assistance and vast expertise made possible to finish this work. I also appreciate your kindness and good humor during all these years.

My thesis received funding from the program “*LOEWE – Landes-Offensive zur Entwicklung Wissenschaftlichökonomischer Exzellenz*” of Hesse’s Ministry of Higher Education, Research, and the Arts. In addition, I received a scholarship from the Peruvian Council of Science, Technology and Innovation (CONCYTEC) through the initiative *CIENCIACTIVA: Programa de becas de doctorado en el extranjero del Gobierno del Perú (291-2014-FONDECYT)*. I also received travel grants from the European Molecular Biology Organization (EMBO), the Society for the Study of Evolution, the University of Texas at Austin, and Unitas Malacologica, in order to participate in workshops and conferences, to learn state-of-the-art methods in phylogenetics and bioinformatics and to present my results to the scientific community.

I thank all members from the Molecular Ecology group and the Ecological Genomics group at the Senckenberg Biodiversity and Climate Research Centre, especially Andreas, Friederike, Susanne, Tilman, Vikas, and Will. Moreover, I met really good friends during my stay in Frankfurt. Thanks to Antonio, Andrés, Eugenia, Henrik, Kelly, Marcia, Melissa, Miguel, Pamela, and Rodrigo. I definitely enjoyed your company and friendship. Also, many thanks to the Goethe Graduate Academy (GRADE), especially Dr. Heike Zimmermann-Timm and Eva Niederlechner for their support in providing opportunities to develop my academic and soft skills during these years. Finally, I would like to thank my mother Emma and my sisters Cynthia and Giovana, for their continuous long-distance support all these years. They were my inspiration to finish this work.

## REFERENCES

- AI, W. M., CHEN, S. B., CHEN, X., SHEN, X. J., SHEN, Y. Y. 2014. Parallel evolution of IDH2 gene in cetaceans, primates and bats. *FEBS letters*, 588, 450-4.
- AMEMIYA, C. T., ALFOLDI, J., LEE, A. P., FAN, S., PHILIPPE, H., MACCALLUM, I., BRAASCH, I., MANOUSAKI, T., SCHNEIDER, I., ROHNER, N., ORGAN, C., CHALOPIN, D., SMITH, J. J., ROBINSON, M., DORRINGTON, R. A., GERDOL, M., AKEN, B., BISCOTTI, M. A., BARUCCA, M., BAURAIN, D., BERLIN, A. M., BLATCH, G. L., BUONOCORE, F., BURMESTER, T., CAMPBELL, M. S., CANAPA, A., CANNON, J. P., CHRISTOFFELS, A., DE MORO, G., EDKINS, A. L., FAN, L., FAUSTO, A. M., FEINER, N., FORCONI, M., GAMIELDIEN, J., GNERRE, S., GNIRKE, A., GOLDSTONE, J. V., HAERTY, W., HAHN, M. E., HESSE, U., HOFFMANN, S., JOHNSON, J., KARCHNER, S. I., KURAKU, S., LARA, M., LEVIN, J. Z., LITMAN, G. W., MAUCELI, E., MIYAKE, T., MUELLER, M. G., NELSON, D. R., NITSCHKE, A., OLMO, E., OTA, T., PALLAVICINI, A., PANJI, S., PICONE, B., PONTING, C. P., PROHASKA, S. J., PRZYBYLSKI, D., SAHA, N. R., RAVI, V., RIBEIRO, F. J., SAUKA-SPENGLER, T., SCAPIGLIATI, G., SEARLE, S. M., SHARPE, T., SIMAKOV, O., STADLER, P. F., STEGEMAN, J. J., SUMIYAMA, K., TABBAA, D., TAFER, H., TURNER-MAIER, J., VAN HEUSDEN, P., WHITE, S., WILLIAMS, L., YANDELL, M., BRINKMANN, H., VOLFF, J. N., TABIN, C. J., SHUBIN, N., SCHARTL, M., JAFFE, D. B., POSTLETHWAIT, J. H., VENKATESH, B., DI PALMA, F., LANDER, E. S., MEYER, A., LINDBLAD-TOH, K. 2013. The African coelacanth genome provides insights into tetrapod evolution. *Nature*, 496, 311-6.
- ARENDETT, J., REZNICK, D. 2008. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends in Ecology & Evolution*, 23, 26-32.
- BECKSTEAD, W. A., EBBERT, M. T., ROWE, M. J., MCCLELLAN, D. A. 2009. Evolutionary pressure on mitochondrial cytochrome b is consistent with a role of CytbI7T affecting longevity during caloric restriction. *PLoS ONE*, 4, e5836.
- BREURE, A. S. H., ROMERO, P. E. 2012. Support and surprises: molecular phylogeny of the land snail superfamily Orthalicoidea using a three-locus gene analysis with a divergence time analysis and ancestral area reconstruction (Gastropoda: Stylommatophora). *Archiv für Molluskenkunde: International Journal of Malacology*, 141, 1-20.
- CHIKINA, M., ROBINSON, J. D., CLARK, N. L. 2016. Hundreds of genes experienced convergent shifts in selective pressure in marine mammals. *Molecular Biology and Evolution*, 33, 2182-92.

- CROCKFORD, S. J. 2009. Evolutionary roots of iodine and thyroid hormones in cell-cell signaling. *Integrative and Comparative Biology*, 49, 155-66.
- DA FONSECA, R. R., JOHNSON, W. E., O'BRIEN, S. J., RAMOS, M. J., ANTUNES, A. 2008. The adaptive evolution of the mammalian mitochondrial genome. *BMC Genomics*, 9, 119.
- DAVIES, K. T., COTTON, J. A., KIRWAN, J. D., TEELING, E. C., ROSSITER, S. J. 2012. Parallel signatures of sequence evolution among hearing genes in echolocating mammals: an emerging model of genetic convergence. *Heredity*, 108, 480-9.
- DAYRAT, B., CONRAD, M., BALAYAN, S., WHITE, T. R., ALBRECHT, C., GOLDING, R., GOMES, S. R., HARASEWYCH, M. G., MARTINS, A. M. 2011. Phylogenetic relationships and evolution of pulmonate gastropods (Mollusca): new insights from increased taxon sampling. *Molecular Phylogenetics and Evolution*, 59, 425-37.
- DAYRAT, B., TILLIER, S. 2002. Evolutionary relationships of euthyneuran gastropods (Mollusca): a cladistic re-evaluation of morphological characters. *Zoological Journal of the Linnean Society*, 403-470.
- DELAUX, P.-M., NANDA, A. K., MATHÉ, C., SEJALON-DELMAS, N., DUNAND, C. 2012. Molecular and biochemical aspects of plant terrestrialization. *Perspectives in Plant Ecology, Evolution and Systematics*, 14, 49-59.
- DELAUX, P. M., RADHAKRISHNAN, G. V., JAYARAMAN, D., CHEEMA, J., MALBREIL, M., VOLKENING, J. D., SEKIMOTO, H., NISHIYAMA, T., MELKONIAN, M., POKORNY, L., ROTHFELS, C. J., SEDEROFF, H. W., STEVENSON, D. W., SUREK, B., ZHANG, Y., SUSSMAN, M. R., DUNAND, C., MORRIS, R. J., ROUX, C., WONG, G. K., OLDROYD, G. E., ANE, J. M. 2015. Algal ancestor of land plants was preadapted for symbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 13390-5.
- DUNN, C. W. 2013. Evolution: out of the ocean. *Current Biology*, 23, R241-3.
- EFREMOV, R. G., SAZANOV, L. A. 2011. Structure of the membrane domain of respiratory complex I. *Nature*, 476, 414-20.
- ELLISON, A. M., FARNSWORTH, E. J., MERKT, R. E. 1999. Origins of mangrove ecosystems and the mangrove biodiversity anomaly. *Global Ecology and Biogeography*, 8, 95-115.
- FADDEEVA, A., STUDER, R. A., KRAAIJEVELD, K., SIE, D., YLSTRA, B., MARIEN, J., OP DEN CAMP, H. J., DATEMA, E., DEN DUNNEN, J. T., VAN STRAALLEN, N. M., ROELOFS, D. 2015. Collembolan transcriptomes highlight molecular evolution of hexapods and provide clues on the adaptation to terrestrial life. *PLoS ONE*, 10, e0130600.
- FLORES-MOLINA, M. R., THOMAS, D., LOVAZZANO, C., NÚÑEZ, A., ZAPATA, J., KUMAR, M., CORREA, J. A., CONTRERAS-PORCIA, L. 2014. Desiccation stress in intertidal seaweeds:

- Effects on morphology, antioxidant responses and photosynthetic performance. *Aquatic Botany*, 113, 90-99.
- GARVIN, M. R., BIELAWSKI, J. P., GHARRETT, A. J. 2011. Positive Darwinian selection in the piston that powers proton pumps in complex I of the mitochondria of Pacific salmon. *PLoS ONE*, 6, e24127.
- GARVIN, M. R., BIELAWSKI, J. P., SAZANOV, L. A., GHARRETT, A. J. 2015. Review and meta-analysis of natural selection in mitochondrial complex I in metazoans. *Journal of Zoological Systematics and Evolutionary Research*, 53, 1-17.
- GE, R. L., CAI, Q., SHEN, Y. Y., SAN, A., MA, L., ZHANG, Y., YI, X., CHEN, Y., YANG, L., HUANG, Y., HE, R., HUI, Y., HAO, M., LI, Y., WANG, B., OU, X., XU, J., ZHANG, Y., WU, K., GENG, C., ZHOU, W., ZHOU, T., IRWIN, D. M., YANG, Y., YING, L., BAO, H., KIM, J., LARKIN, D. M., MA, J., LEWIN, H. A., XING, J., PLATT, R. N., 2ND, RAY, D. A., AUVIL, L., CAPITANU, B., ZHANG, X., ZHANG, G., MURPHY, R. W., WANG, J., ZHANG, Y. P., WANG, J. 2013. Draft genome sequence of the Tibetan antelope. *Nature communications*, 4, 1858.
- GORDON, M. S., OLSON, E. C. 1995. *Invasions of the land: The transitions of organisms from aquatic to terrestrial life*, New York, Columbia University Press.
- HAZPRUNAR, G. 1985. The Heterobranchia—a new concept of the phylogeny of the higher Gastropoda. *Zeitschrift für Zoologische Systematik und Evolutionsforschung*, 23, 15-37.
- HEYLAND, A., PRICE, D. A., BODNAROVA-BUGANOVA, M., MOROZ, L. L. 2006. Thyroid hormone metabolism and peroxidase function in two non-chordate animals. *Journal of experimental zoology. Part B, Molecular and developmental evolution*, 306, 551-66.
- HOLLOWAY, A. K., LAWNICZAK, M. K., MEZEY, J. G., BEGUN, D. J., JONES, C. D. 2007. Adaptive gene expression divergence inferred from population genomics. *PLoS Genetics*, 3, 2007-13.
- INNAN, H., KONDRASHOV, F. 2010. The evolution of gene duplications: classifying and distinguishing between models. *Nature Reviews Genetics*, 11, 97-108.
- JOCHUM, A., PROZOROVA, L., SHARYI-OOL, M., PALL-GERGELY, B. 2015. A new member of troglobitic Carychiidae, *Koreozospeum nodongense* gen. et sp. n. (Gastropoda, Eupulmonata, Ellobioidea) is described from Korea. *Zookeys*, 39-57.
- JÖRGER, K. M., STÖGER, I., KANO, Y., FUKUDA, H., KNEBELSBERGER, T., SCHRODL, M. 2010. On the origin of Acochlidia and other enigmatic euthyneuran gastropods, with implications for the systematics of Heterobranchia. *BMC Evolutionary Biology*, 10, 323.
- KAMEDA, Y., KATO, M. 2011. Terrestrial invasion of pomatiopsid gastropods in the heavy-snow region of the Japanese Archipelago. *BMC Evolutionary Biology*, 11, 118.

- KANO, Y., BRENZINGER, B., NUTZEL, A., WILSON, N. G., SCHRÖDL, M. 2016. Ringicolid bubble snails recovered as the sister group to sea slugs (Nudipleura). *Scientific Reports*, 6, 30908.
- KANO, Y., CHIBA, S., KASE, T. 2002. Major adaptive radiation in neritopsine gastropods estimated from 28S rRNA sequences and fossil records. *Proceedings of the Royal Society B: Biological Sciences*, 269, 2457-65.
- KANO, Y., NEUSSER, T. P., FUKUMORI, H., JÖRGER, K. M., SCHRÖDL, M. 2015. Sea-slug invasion of the land. *Biological Journal of the Linnean Society*, 116, 253-259.
- KLUSSMANN-KOLB, A., DINAPOLI, A., KUHN, K., STREIT, B., ALBRECHT, C. 2008. From sea to land and beyond--new insights into the evolution of euthyneuran Gastropoda (Mollusca). *BMC Evolutionary Biology*, 8, 57.
- KOCOT, K. M. 2013. Recent advances and unanswered questions in deep molluscan phylogenetics. *American Malacological Bulletin*, 31, 195-208.
- KOCOT, K. M., HALANYCH, K. M., KRUG, P. J. 2013. Phylogenomics supports Panpulmonata: opisthobranch paraphyly and key evolutionary steps in a major radiation of gastropod molluscs. *Molecular Phylogenetics and Evolution*, 69, 764-71.
- LAURIN, M. 2010. *How vertebrates left water*, Berkeley, University of California Press.
- LEE, J., RHEE, K.-J., KIM, S. H., KANG, Y. W., CHO, Y., YANG, S.-J., PAN, C.-H., KIM, Y. S. 2015. O6-Methylguanine-DNA methyltransferase (MGMT) gene expression is associated with ultraviolet B (UVB)-induced cell growth inhibition and recovery. *Genes & Genomics*, 37, 789-796.
- LILLYWHITE, H. B. 2012. Evolutionary history and the conquest of land. *Bioscience*, 62, 517-518.
- LITTLE, C. 1990. *The terrestrial invasion: An ecophysiological approach to the origins of land animals*, Cambridge, Cambridge University Press.
- MARTINS, A. M. 1995. A new species of *Pythia* Röding, 1798 (Pulmonata, Ellobiidae), from New Ireland, Papua New Guinea. *Molluscan Research*, 16.
- MARTINS, A. M. 2007. Morphological and anatomical diversity within the Ellobiidae (Gastropoda, Pulmonata, Archaeopulmonata). *Vita Malacologica*, 4, 1-28.
- MCCLELLAN, D. A., PALFREYMAN, E. J., SMITH, M. J., MOSS, J. L., CHRISTENSEN, R. G., SAILSBERY, J. K. 2005. Physicochemical evolution and molecular adaptation of the cetacean and artiodactyl cytochrome *b* proteins. *Molecular Biology and Evolution*, 22, 437-55.
- MCGOWEN, M. R., GATESY, J., WILDMAN, D. E. 2014. Molecular evolution tracks macroevolutionary transitions in Cetacea. *Trends in Ecology & Evolution*, 29, 336-46.
- MITTERBOECK, T. F., CHEN, A. Y., ZAHEER, O. A., MA, E. Y., ADAMOWICZ, S. J. 2016a. Do saline taxa evolve faster? Comparing relative rates of molecular evolution between freshwater and marine eukaryotes. *Evolution*, 70, 1960-78.

- MITTERBOECK, T. F., FU, J., ADAMOWICZ, S. J. 2016b. Rates and patterns of molecular evolution in freshwater versus terrestrial insects. *Genome*, 1-13.
- NIELSEN, R. 2005. Molecular signatures of natural selection. *Annual Review of Genetics*, 39, 197-218.
- NIKAIDO, M., NOGUCHI, H., NISHIHARA, H., TOYODA, A., SUZUKI, Y., KAJITANI, R., SUZUKI, H., OKUNO, M., AIBARA, M., NGATUNGA, B. P., MZIGHANI, S. I., KALOMBO, H. W., MASENGI, K. W., TUDA, J., NOGAMI, S., MAEDA, R., IWATA, M., ABE, Y., FUJIMURA, K., OKABE, M., AMANO, T., MAENO, A., SHIROISHI, T., ITOH, T., SUGANO, S., KOHARA, Y., FUJIYAMA, A., OKADA, N. 2013. Coelacanth genomes reveal signatures for evolutionary transition from water to land. *Genome Research*, 23, 1740-8.
- NOLAN, J. R., BERGTHORSSON, U., ADEMA, C. M. 2014. *Physella acuta*: atypical mitochondrial gene order among panpulmonates (Gastropoda). *Journal of Molluscan Studies*, 80, 388-399.
- PFENNINGER, M., CORDELLIER, M., STREIT, B. 2006. Comparing the efficacy of morphologic and DNA-based taxonomy in the freshwater gastropod genus *Radix* (Basommatophora, Pulmonata). *BMC Evolutionary Biology*, 6, 100.
- PINCEEL, J., JORDAENS, K., BACKELJAU, T. 2005. Extreme mtDNA divergences in a terrestrial slug (Gastropoda, Pulmonata, Arionidae): accelerated evolution, allopatric divergence and secondary contact. *Journal of Evolutionary Biology*, 18, 1264-80.
- PONDER, W. F., LINDBERG, D. R. 2008. *Phylogeny and Evolution of the Mollusca*, Berkeley and Los Angeles, California, University of California Press.
- PROCHES, S., POLGAR, G., MARSHALL, D. J. 2014. K-Pg events facilitated lineage transitions between terrestrial and aquatic ecosystems. *Biological Letters*, 10, 20140010.
- ROMERO, P. E., PFENNINGER, M., KANO, Y., KLUSSMANN-KOLB, A. 2016a. Molecular phylogeny of the Ellobiidae (Gastropoda: Panpulmonata) supports independent terrestrial invasions. *Molecular Phylogenetics and Evolution*, 97, 43-54.
- ROMERO, P. E., WEIGAND, A. M., PFENNINGER, M. 2016b. Positive selection on panpulmonate mitogenomes provide new clues on adaptations to terrestrial life. *BMC Evolutionary Biology*, 16, 164.
- ROMERO, P. E., FELDMEYER, B., PFENNINGER, M. Panpulmonate transcriptomes reveal candidate genes involved in the transition to terrestrial and freshwater habitats in gastropods. *Under review*.
- ROSENBLUM, E. B., PARENT, C. E., BRANDT, E. E. 2014. The molecular basis of phenotypic convergence. *Annual Review of Ecology, Evolution, and Systematics*, 45, 203-226.
- ROTA-STABELLI, O., DALEY, A. C., PISANI, D. 2013. Molecular timetrees reveal a Cambrian colonization of land and a new scenario for ecdysozoan evolution. *Current Biology*, 23, 392-8.

- SABETI, P. C., SCHAFFNER, S. F., FRY, B., LOHMUELLER, J., VARILLY, P., SHAMOVSKY, O., PALMA, A., MIKKELSEN, T. S., ALTSHULER, D., LANDER, E. S. 2006. Positive natural selection in the human lineage. *Science*, 312, 1614-20.
- SCHRÖDL, M. 2014. Time to say “Bye-bye Pulmonata”? *Spixiana*, 37, 161-164.
- SHEN, Y. Y., LIANG, L., ZHU, Z. H., ZHOU, W. P., IRWIN, D. M., ZHANG, Y. P. 2010. Adaptive evolution of energy metabolism genes and the origin of flight in bats. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 8666-71.
- STROTHER, P. K., BATTISON, L., BRASIER, M. D., WELLMAN, C. H. 2011. Earth's earliest non-marine eukaryotes. *Nature*, 473, 505-9.
- TOMASCO, I. H., LESSA, E. P. 2011. The evolution of mitochondrial genomes in subterranean caviomorph rodents: adaptation against a background of purifying selection. *Molecular Phylogenetics and Evolution*, 61, 64-70.
- THOMAZ, D., GUILLER, A., CLARKE, B. 1996. Extreme divergence of mitochondrial DNA within species of pulmonate land snails. *Proceedings of the Royal Society B: Biological Sciences*, 263, 363-368.
- VANDEBERGH, W., BOSSUYT, F. 2012. Radiation and functional diversification of alpha keratins during early vertebrate evolution. *Molecular Biology and Evolution*, 29, 995-1004.
- VERMEIJ, G. J., DUDLEY, R. 2000. Why are there so few evolutionary transitions between aquatic and terrestrial ecosystems? *Biological Journal of the Linnean Society*, 70, 541-554.
- WÄGELE, H., KLUSSMANN-KOLB, A., VERBEEK, E., SCHRÖDL, M. 2014. Flashback and foreshadowing—a review of the taxon Opisthobranchia. *Organisms Diversity & Evolution*, 14, 133-149.
- WEIGAND, A. M., JOCHUM, A., SLAPNIK, R., SCHNITZLER, J., ZARZA, E., KLUSSMANN-KOLB, A. 2013. Evolution of microgastropods (Ellobioidea, Carychiidae): integrating taxonomic, phylogenetic and evolutionary hypotheses. *BMC Evolutionary Biology*, 13, 18.
- WIRTH, C., BRANDT, U., HUNTE, C., ZICKERMANN, V. 2016. Structure and function of mitochondrial complex I. *Biochimica et Biophysica Acta*, 1857, 902-14.
- WISSELER, L., CODONER, F. M., GU, J., REUSCH, T. B., OLSEN, J. L., PROCACCINI, G., BORNBERG-BAUER, E. 2011. Back to the sea twice: identifying candidate plant genes for molecular evolution to marine life. *BMC Evolutionary Biology*, 11, 8.
- YANG, Z., DOS REIS, M. 2011. Statistical properties of the branch-site test of positive selection. *Molecular Biology and Evolution*, 28, 1217-28.
- YANG, Z., NIELSEN, R., GOLDMAN, N., PEDERSEN, A. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics*, 155, 431-449.

- YIM, H. S., CHO, Y. S., GUANG, X., KANG, S. G., JEONG, J. Y., CHA, S. S., OH, H. M., LEE, J. H., YANG, E. C., KWON, K. K., KIM, Y. J., KIM, T. W., KIM, W., JEON, J. H., KIM, S. J., CHOI, D. H., JHO, S., KIM, H. M., KO, J., KIM, H., SHIN, Y. A., JUNG, H. J., ZHENG, Y., WANG, Z., CHEN, Y., CHEN, M., JIANG, A., LI, E., ZHANG, S., HOU, H., KIM, T. H., YU, L., LIU, S., AHN, K., COOPER, J., PARK, S. G., HONG, C. P., JIN, W., KIM, H. S., PARK, C., LEE, K., CHUN, S., MORIN, P. A., O'BRIEN, S. J., LEE, H., KIMURA, J., MOON, D. Y., MANICA, A., EDWARDS, J., KIM, B. C., KIM, S., WANG, J., BHAK, J., LEE, H. S., LEE, J. H. 2014. Minke whale genome and aquatic adaptation in cetaceans. *Nature Genetics*, 46, 88-92.
- YOU, X., BIAN, C., ZAN, Q., XU, X., LIU, X., CHEN, J., WANG, J., QIU, Y., LI, W., ZHANG, X., SUN, Y., CHEN, S., HONG, W., LI, Y., CHENG, S., FAN, G., SHI, C., LIANG, J., TOM TANG, Y., YANG, C., RUAN, Z., BAI, J., PENG, C., MU, Q., LU, J., FAN, M., YANG, S., HUANG, Z., JIANG, X., FANG, X., ZHANG, G., ZHANG, Y., POLGAR, G., YU, H., LI, J., LIU, Z., ZHANG, G., RAVI, V., COON, S. L., WANG, J., YANG, H., VENKATESH, B., WANG, J., SHI, Q. 2014. Mudskipper genomes provide insights into the terrestrial adaptation of amphibious fishes. *Nature Communications*, 5, 5594.
- ZAPATA, F., WILSON, N. G., HOWISON, M., ANDRADE, S. C., JÖRGER, K. M., SCHRÖDL, M., GOETZ, F. E., GIRIBET, G., DUNN, C. W. 2014. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. *Proceedings of the Royal Society B: Biological Sciences*, 281, 20141739.
- ZHAI, W., NIELSEN, R., GOLDMAN, N., YANG, Z. 2012. Looking for Darwin in genomic sequences--validity and success of statistical methods. *Molecular Biology and Evolution*, 29, 2889-93.
- ZHOU, Q., ZHANG, G., ZHANG, Y., XU, S., ZHAO, R., ZHAN, Z., LI, X., DING, Y., YANG, S., WANG, W. 2008. On the origin of new genes in *Drosophila*. *Genome Research*, 18, 1446-55.
- ZHOU, X., SUN, F., XU, S., FAN, G., ZHU, K., LIU, X., CHEN, Y., SHI, C., YANG, Y., HUANG, Z., CHEN, J., HOU, H., GUO, X., CHEN, W., CHEN, Y., WANG, X., LV, T., YANG, D., ZHOU, J., HUANG, B., WANG, Z., ZHAO, W., TIAN, R., XIONG, Z., XU, J., LIANG, X., CHEN, B., LIU, W., WANG, J., PAN, S., FANG, X., LI, M., WEI, F., XU, X., ZHOU, K., WANG, J., YANG, G. 2013. Baiji genomes reveal low genetic variability and new insights into secondary aquatic adaptations. *Nature Communications*, 4, 2708.



## APPENDIX: PUBLICATIONS AND MANUSCRIPTS

*Publication 1: Romero PE, Pfenninger M, Kano Y, Klussmann-Kolb A. Molecular phylogeny of the Ellobiidae (Gastropoda: Panpulmonata) supports independent terrestrial invasions. *Molecular Phylogenetics and Evolution*. 2016;97:43-54.*

## Anlage 1

### Erklärung zu den Autorenanteilen

an der Publikation / des Manuskripts (Titel): Molecular phylogeny of the Ellobiidae (Gastropoda: Panpulmonata) supports independent terrestrial invasions.

Status (*submitted, under review, in revision, accepted*): Accepted, 21.12.2015. Published online, 24.12.2015.

Ab *accepted* Name der Zeitschrift: Molecular Phylogenetics and Evolution. 2016;97:43-54.

Beteiligte Autoren (Namen mit eindeutigen Kürzeln):

- PER: Pedro Eduardo Romero
- MP: Markus Pfenninger
- YK: Yasunori Kano
- AKK: Annette Klussmann-Kolb

### Was hat der Promovierende bzw. was haben die Koautoren beigetragen?

#### (1) zu Entwicklung und Planung

Promovierender PER: 35%

Co-Autor MP: 5%

Co-Autor YK: 20%

Co-Autor AKK: 40%

#### (2) zur Durchführung der einzelnen Untersuchungen und Experimente

Promovierender PER: 90%. DNA extraction, amplification (PCR) and sequencing. Gel electrophoresis.

Co-Autor AKK: 10%. Laboratory protocols and PCR initial setting.

#### (3) zur Erstellung der Datensammlung und Abbildungen

Promovierender PER: 30%. Submission of the DNA sequences to Genbank.

Co-Autor YK: 40%. Logistics and sampling in the Ryuku islands.

Co-Autor AKK: 30%. Gathering of DNA sequences from previous theses and collections.

#### (4) zur Analyse und Interpretation der Daten

Promovierender PER: 50%. DNA sequence analysis, multiple sequence alignments. Phylogenetic reconstruction, divergence times analyses.

Co-Autor MP: 5%. LTT plots.

Co-Autor YK: 25%. Interpretation of the phylogeny in a systematics framework, and analysis of divergence times.

Co-Autor AKK: 20%. Interpretation of the phylogeny in a systematics framework.

#### (5) zum Verfassen des Manuskripts

Promovierender PER: 50%

Co-Autor MP: 5%

Co-Autor YK: 30%

Co-Autor AKK: 15%

Datum/Ort: 01.12.2016, Frankfurt am Main.

---

Unterschrift Promovend: \_\_\_\_\_

#### Zustimmende Bestätigungen der oben genannten Angaben

Unterschrift Betreuer: \_\_\_\_\_ Datum/Ort: 01.12.2016, Frankfurt am Main.

Ggfs. Unterschrift *corresponding author* PER: \_\_\_\_\_ Datum/Ort: 01.12.2016, Frankfurt am Main.



Contents lists available at ScienceDirect

## Molecular Phylogenetics and Evolution

journal homepage: [www.elsevier.com/locate/ympev](http://www.elsevier.com/locate/ympev)Molecular phylogeny of the Ellobiidae (Gastropoda: Panpulmonata) supports independent terrestrial invasions<sup>☆</sup>Pedro E. Romero<sup>a,b,\*</sup>, Markus Pfenninger<sup>a,c</sup>, Yasunori Kano<sup>d</sup>, Annette Klussmann-Kolb<sup>e</sup><sup>a</sup> Senckenberg Biodiversity and Climate Research Centre (BiK-F), Senckenberg Gesellschaft für Naturforschung (SGN), Senckenberganlage 25, Frankfurt am Main 60325, Germany<sup>b</sup> Institute for Ecology, Evolution & Diversity, Goethe University, Max-von-Laue-Straße 13, Frankfurt am Main 60438, Germany<sup>c</sup> Institute for Ecology, Evolution & Diversity, Goethe University, Max-von-Laue-Straße 13, Frankfurt am Main 60439, Germany<sup>d</sup> Atmosphere and Ocean Research Institute, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8564, Japan<sup>e</sup> Zoologisches Forschungsmuseum König, Adenauerallee 160, 53113 Bonn, Germany

## ARTICLE INFO

## Article history:

Received 22 April 2015

Revised 20 November 2015

Accepted 21 December 2015

Available online 24 December 2015

## Keywords:

Carychiinae

Pythiinae

Terrestrial invasion

Realm transitions

Molecular clock

Topology tests

## ABSTRACT

Gastropods of the family Ellobiidae are an interesting group in which to study transitions from intertidal to terrestrial realms. However, the phylogenetic relationships within this family still lack resolution. We present a phylogenetic hypothesis of the Ellobiidae based on Bayesian and maximum likelihood phylogenies. We used nuclear (18S, 28S, H3) and mitochondrial (16S, 12S, COI) data, increasing the numbers of markers and data, and making this the most comprehensive phylogenetic study of the family to date. Our results support phylogenetic hypotheses derived from morphological data, and provide a supported framework to evaluate the internal relationships within Ellobiidae. The resulting phylogenetic trees support the previous hypothesis that the Ellobiidae are monophyletic only if the Trimusculinae (*Otina*, *Smeagol* and *Trimusculus*) are considered part of this family. In addition, we found that the Carychiinae, Ellobiinae and Pythiinae are reciprocally monophyletic and closely related, with the Carychiinae as sister group to Ellobiinae. Relationships within Melampodinae and Pedipedinae and their phylogenetic positions remain unresolved. Land invasion by the Ellobiidae occurred independently in Carychiinae and *Pythia* during different geological times (Mesozoic and Cenozoic, respectively). Diversification in the family does not appear to be related to past climate and biotic changes, neither the Cretaceous–Paleogene boundary nor the lowering of the sea level in the Oligocene.

© 2015 Elsevier Inc. All rights reserved.

## 1. Introduction

The conquest of land by organisms that evolved from aquatic ancestors represents one of the most astonishing events in Earth's history of life (Lillywhite, 2012). This step was achieved multiple times in different phyla by means of specific adaptations in osmoregulation and water balance, air breathing, nitrogen excretion, reproduction, locomotion and behavior (Little, 1990). Gastropods are one of the most successful taxa to have developed terrestrial lineages; land invasion has been achieved several times, for example in Neritimorpha, Caenogastropoda, and Heterobranchia (Kameda and Kato, 2011).

Within Heterobranchia, the clade Panpulmonata includes several lineages that invaded the intertidal zone and non-marine ecosystems, for instance, Acochlidia, Ellobiidae, Hygrophila, Stylommatophora and Systellommatophora (Klussmann-Kolb et al., 2008; Jörger et al., 2010; Kano et al., 2015). Previous phylogenetic studies suggested that the evolution of panpulmonate terrestriality has occurred multiple times (Barker, 2001), although there is still controversy about the direction of the transitions between realms. Only solid phylogenetic hypotheses can serve as a framework to clarify the origins of this ecological diversity and the habitat transitions. Panpulmonate phylogeny therefore has important implications for understanding how gastropods transitioned out of marine habitats, resulting in one of the most significant adaptive radiations among animals (Kocot et al., 2013).

The panpulmonate family Ellobiidae (hollow-shelled snails) are a characteristic component of the intertidal and supratidal zones of mangrove forests and muddy shores in tropical regions worldwide, as well as of salt marshes and upper littoral rocky areas in temperate regions (Martins, 1996). In addition, three ellobiid genera include truly terrestrial species, which are independent from any

<sup>☆</sup> This paper was edited by the Associate Editor Jan Strugnell.

\* Corresponding author at: Senckenberg Biodiversity and Climate Research Centre (BiK-F), Senckenberg Gesellschaft für Naturforschung (SGN), Senckenberganlage 25, Frankfurt am Main 60325, Germany.

E-mail addresses: [pedro.romero@senckenberg.de](mailto:pedro.romero@senckenberg.de), [quipu.romero@gmail.com](mailto:quipu.romero@gmail.com) (P.E. Romero), [pfenninger@bio.uni-frankfurt.de](mailto:pfenninger@bio.uni-frankfurt.de) (M. Pfenninger), [kano@aori.u-tokyo.ac.jp](mailto:kano@aori.u-tokyo.ac.jp) (Y. Kano), [a.klussmann-kolb@zfmk.de](mailto:a.klussmann-kolb@zfmk.de) (A. Klussmann-Kolb).

aquatic habitats: *Pythia* Röding, 1798 in the Indo-western Pacific (for instance *Pythia colmani* Martins, 1995 living in the rainforest of Papua New Guinea; Martins, 1995), *Carychium* Müller, 1773 in the litterfall of forests and Holarctic riparian zones, and *Zospeum* Bourguignat, 1856 in European karst caves (Barker, 2001). Approximately 800 species names have been proposed for ellobiids, of which 100 (Mordan and Wade, 2008) to 250 (Weigand et al., 2013) are likely to be valid. The latest classification proposes five subfamilies within the Ellobiidae: Ellobiinae, Carychiinae, Melampodinae, Pedipedinae, and Pythiinae (Martins, 2007), although some authors have recognized a family rank for Carychiinae (Jörger et al., 2010; Weigand et al., 2013).

Morphological data have not been sufficient in determining the phylogenetic relationships within Ellobiidae, nor the systematic position of the family among panpulmonates (Morton, 1955; Dayrat and Tillier, 2002; Martins, 2007). However, the subfamilies can be distinguished by anatomical characters in the reproductive and nervous systems. The Carychiinae, Pedipedinae and Pythiinae are monoaulic (i.e. they possess a single hermaphroditic duct), whereas the Ellobiinae are diaulic (male and female ducts run separately), and the Melampodinae are semidiaulic (male and female gonoducts distally separated). In addition, the Pedipedinae and Melampodinae can be distinguished from species in the other subfamilies by their short visceral nerve ring (Martins, 2007).

Previous molecular phylogenetic studies were also not able to resolve neither the systematic position of the Ellobiidae nor the relations of the clades within. Complete mitochondrial genomes of ellobiids have been used in previous studies aiming to reconstruct the phylogeny of Gastropoda (Grande et al., 2008) or Heterobranchia (Medina et al., 2011; White et al., 2011). However, these phylogenetic analyses based only on mitochondrial data received criticism (Stöger and Schrödl, 2013; Wägele et al., 2014) and contradicted the results from multi-locus (nuclear and mitochondrial) analyses (Dinapoli and Klussmann-Kolb, 2010; Jörger et al., 2010) and phylogenomic studies (Kocot et al., 2013; Zapata et al., 2014). Long branch attraction artifacts were assumed to infer an erroneous rooting in the mitochondrial trees (Schrödl et al., 2011; Schrödl, 2014). Multi-locus studies have used a core set of molecular markers, including the nuclear 18S rRNA and 28S rRNA, and mitochondrial 16S rRNA and cytochrome oxidase I (COI) (Klussmann-Kolb et al., 2008; Jörger et al., 2010; Dayrat et al., 2011). In addition, it has been hypothesized that the family includes the genera *Otina* Gray, 1847 and *Smeagol* Climo, 1980 (formerly *Otinidae*), and *Trimusculus* Schmidt, 1818 (*Trimusculidae*), also occurring in the intertidal zone (Dayrat et al., 2011). However, relationships among the subfamilies have remained unclear, making it difficult to interpret the time and occurrence of habitat transitions.

In this work, we reconstructed a phylogenetic hypothesis of the Ellobiidae with increased numbers of taxa (comprising all subfamilies) and loci (nuclear 18S rRNA, 28S rRNA and Histone H3, and mitochondrial 16S rRNA, 12S rRNA and COI). Our topology showed improved phylogenetic resolution within Ellobiidae, allowing us to evaluate the monophyly of the subfamilies and to test the current hypotheses of the tempo and mode of habitat transitions. We found that ellobiids colonized the land independently in Carychiinae and Pythiinae during the Mesozoic and Cenozoic periods, respectively.

## 2. Material and methods

### 2.1. Taxon sampling

A total of 63 ellobiid species comprising 23 genera were included in this study. We followed the suggestions by Martins (2007) for the nomenclature of the subfamilies. Recent studies

using a multilocus (Jörger et al., 2010; Dayrat et al., 2011) or phylogenomic approach (Zapata et al., 2014) have shown that the Systellommatophora are the sister group of the Ellobiidae. Thus, the outgroup included 11 species from Onchidiidae (Systellommatophora). Specimens of all five ellobiid subfamilies were included in our data matrix: Carychiinae (2 genera used/2 genera accepted in Martins (2007)), Ellobiinae (5/5) Melampodinae (3/5), Pedipedinae (2/4), and Pythiinae (8/8). Data sets from previous studies (Klussmann-Kolb et al., 2008; Dinapoli and Klussmann-Kolb, 2010; Dayrat et al., 2011; Vonnemann et al., 2005; Weigand et al., 2011, 2013) served as a starting point for our analyses, in particular sequences of 18S rRNA, 16S rRNA and COI genes. The matrix was expanded with three additional markers: nuclear 28S rRNA and Histone H3, and mitochondrial 12S rRNA. In addition, taxon sampling was extended, e.g. 20 of the 63 ellobiid species now included were previously absent in GenBank. The mean ratio of sequences per marker was 81%, ranging from 100% for 18S rRNA to 64% for H3. Samples are preserved in the collection of the Senckenberg Naturmuseum Frankfurt; detailed information is shown in Table 1. We also considered the genera *Otina*, *Smeagol* and *Trimusculus* as members of the Ellobiidae. The close relationships of these genera to the above five subfamilies were previously proposed by Klussmann-Kolb et al. (2008) and Dinapoli and Klussmann-Kolb (2010), and then confirmed by Dayrat et al. (2011) using a wider taxonomic sampling.

Ellobiids exhibit modest diversity on land (Barker, 2001). In our analysis, terrestrial genera were represented in Carychiinae (*Carychium* and *Zospeum*) and Pythiinae (*Pythia* sp.). The latter species of *Pythia* was collected from a tropical rainforest in the Republic of Palau, where associated mollusks were all typical land snails of Cyclophoroidea and Stylommatophora, and was verified as being terrestrial throughout its ontogeny (Kano et al., 2015). Other species of the genus living on mangrove trees, such as *Pythia scarabaeus* (Linnaeus, 1758), were considered as intertidal. Members of Carychiinae were also considered fully terrestrial because of their geographic and ecological distributions (forests or caves that are far from the sea).

### 2.2. DNA extraction, amplification and sequencing

Genomic DNA was isolated using the DNeasy Blood & Tissue Kit (Qiagen, Germany) from specimens stored in 100% ethanol. A small piece ( $\leq 2 \text{ mm}^3$ ) of the foot was used, and the shell was destructed in case of very small samples. Polymerase chain reactions (PCR) were performed to amplify fragments from the six markers mentioned above; Supplementary Table S1 shows primers used in the PCR reactions. The reactions followed this protocol: Each 25  $\mu\text{L}$  PCR mixture included 1  $\mu\text{L}$  (10 pmol) of each primer, 2.5  $\mu\text{L}$  10 $\times$  PCR buffer, 2  $\mu\text{L}$  (100 mM)  $\text{MgCl}_2$ , 0.2  $\mu\text{L}$  (20 mM) dNTPs, 0.3  $\mu\text{L}$  Taq-polymerase (Fermentas), 1.5  $\mu\text{L}$  (10 mg/mL) bovine serum albumin, 12.5  $\mu\text{L}$  ddH<sub>2</sub>O and 4  $\mu\text{L}$  template DNA. PCR cycles were done with these conditions: 1 min at 95 °C, followed by 30 cycles of 30 s at 95 °C, 30 s at 52 °C and 30 s at 72 °C, and finally, 3 min at 72 °C. Visualization of PCR products was performed on a 1.4% agarose gel. Amplicons were cleaned using the QIAquick PCR Purification Kit or the QIAquick Gel Extraction kit (Qiagen) whenever multiple bands were detected. PCR products were bidirectionally sequenced using the PCR primer pair (5 pmol) and the BigDye<sup>®</sup> Terminator v.3.1 Cycle Sequencing Kit (Life Technologies, Inc.) on an ABI 3730 capillary sequencer following the manufacturer's instructions. The sequencing reactions were performed in the BiK-F Laboratory Centre, Frankfurt am Main. Sequences were deposited in GenBank with the following accession numbers: 18S: KM280960–KM281008, 28S: KM281048–KM281088, 16S: KM281010–KM281044, 12S: KJ920288–KJ920333, COI: KM281099–KM281117, H3: KM281118–KM281155.

**Table 1**  
Species used in this work with GenBank accession numbers and collection sites and habitats (H) of specimens. Accession numbers in bold denote sequences from previous studies. Suprageneric classification reflects topology of new trees. Abbreviations: I, intertidal; T, terrestrial.

Taxon	H	Locality	18S	28S	16S	12S	COI	H3	Voucher #	References
<b>ELLOBIDAE</b>										
<b>Carychiinae</b>										
<i>Carychium exiguum</i> (Say, 1822)	T	USA, New York, Ontario, Naples	KM280960	KM281045	<b>KC206155</b>	KJ920288	<b>HQ171504</b>	<b>KC206211</b>	AMW1322	Weigand et al. (2013)
<i>Carychium floridanum</i> Clapp, 1918	T	USA, Florida, Levy, Gulf Hammock	KM280961	KM281046	KC206144	KJ920289	KM281089	<b>KC206222</b>	AMW1446	Weigand et al. (2013)
<i>Carychium hachijoensis</i> Pilsbry, 1802	T	Japan, Hachijo-jima, Mitsuue	KM280962	KM281047	KM281036	KJ920290	KM281090	KM281118	AMW2184	Weigand et al. (2013)
<i>Carychium minimum</i> Müller, 1774	T	Italy, Sudtiroi, Auer	KM280963	<b>EF489361</b>	<b>KC206173</b>	KJ920291	<b>HQ171538</b>	<b>KC206193</b>	AMW1213	Klussmann-Kolb et al. (2008) and Weigand et al. (2013)
<i>Carychium namodes</i> Clapp, 1905	T	USA, Tennessee, Overton Co, Silt Cave	KM280964	KM281048	<b>KC206176</b>	KJ920292	<b>HQ171543</b>	<b>KC206190</b>	AMW1174	Weigand et al. (2011)
<i>Zospeum alpestre</i> (Freyer, 1855)	T	Slovenia, Olseva, Lozkekarjeva zijalka	KM280965	KM281049	<b>KC206150</b>	KJ920293	<b>HQ171580</b>	<b>KC206216</b>	AMW1373	Weigand et al. (2011) and Weigand et al. (2013)
<i>Zospeum frauinfeldii</i> (Freyer, 1855)	T	Slovenia, Laze Pri Planini, Mackovcice Jama	KM280966	KM281050	<b>KC206205</b>	KJ920294	<b>HQ171589</b>	<b>KC206205</b>	AMW1308	Weigand et al. (2011) and Weigand et al. (2013)
<i>Zospeum starezi</i> Gittenberger, 1980	T	Spain, Asturias, Inguanzo, Cueva del Bosque	KM280967	KM281051	<b>KC206121</b>	KJ920295	KM281091	<b>KC206245</b>	AMW1842	Weigand et al. (2013) and Weigand et al. (2013)
<i>Zospeum vasconicum</i> Prieto, De Winter, Weigand, Gómez & Jochum, 2015	T	Spain, Bizkaia, Guipuzcoa, Ermita de Sandali	KM280968	KM281052	<b>KC206117</b>	KJ920296	KM281092	<b>KC206247</b>	AMW1847	Weigand et al. (2013) and Jochum et al. (2015)
<b>Ellobiinae</b>										
<i>Auriculastrea duplicata</i> (Pfeiffer, 1854)	I	Japan, Kyushu Island, Oita Pref.	KM280969	KM281053	KM281015	KJ920297	KM281093	KM281119	SMF 346409	This work
<i>Auriculastrea subula</i> (Quoy & Gaimard, 1832)	I	Thailand, Phuket, Phuket Town	KM280970	KM281054	KM281012	KJ920298	KM281094	KM281120	ZSM Mol 20032295	This work
<i>Auriculastrea bidentata</i> (Montagu, 1808)	I	Portugal, Azores	KM280971	KM281055	<b>NC_016168</b>	<b>NC_016168</b>	<b>NC_016168</b>	KM281121	ZSM Mol 20000918/CASIZ 184730	White et al. (2011)
<i>Blaumenia heteroclitia</i> (Montagu, 1808)	I	Bermuda, St. George's Parish, Bartram's Pond	KM280972	KM281056	KM281013	KJ920299	KM281095	KM281122	SMF 346410	This work
<i>Blaumenia quadrasii</i> Möllendorff, 1895	I	Japan, Okinawa, Kunigami, Nakijin-son	KM280973	KM281057	KM281014	KJ920300	KM281096	KM281123	SMF 346411	This work
<i>Cylindronis quadrasii</i> Möllendorff, 1895	I	Japan, Okinawa, Kunigami, Nakijin-son	KM280974	KM281058	KM281016	KJ920301	KM281097	KM281124	No tissue left	This work
<i>Ellobium chinense</i> (Pfeiffer, 1864)	I	Japan, Kyushu Island, Oita Pref.	KM280975	KM281059	KM281018	KJ920302	KM281098	KM281125	SMF 346412	This work
<i>Ellobium aurisjudae</i> (Linnaeus, 1758)	I	Thailand, Phuket, Dan-Yit	KM280976	KM281060	KM281017	KJ920303	KM281099	KM281126	SMF 346413	This work
<i>Ellobium scheppneri</i> (Pfeiffer, 1850) Saussaye, 1850	I	Malaysia, Sarawak, Mukah	KM280977	KM281061	KM281019	KJ920304	KM281100	KM281127	SMF 346414	This work
<b>Melampodinae</b>										
<i>Melampus bidentatus</i> Say, 1822	I	Jamaica	KM280979	KM281063	KM281042	KM281009	KM281102	KM281129	SMF 346415	This work
<i>Melampus bulboides</i> (Montagu, 1808)	I	Bermuda, Southampton Parish, Evans Pond	KM280980	KM281064	KM281043	KJ920306	KM281103	KM281130	SMF 346416	This work
<i>Melampus fasciatus</i> (Deshayes, 1830)	I	French Polynesia, Caroline Island	KM280981	-	<b>HQ659875</b>	KJ920307	KM281104	KM281131	UF 294608	Dayrat et al. (2011)
<i>Melampus pulchellus</i> (Petit De La Saussaye, 1843)	I	Thailand, Phuket, Phuket Town	KM280982	KM281065	-	KJ920308	-	KM281132	SMF 346417	This work
<i>Melampus singaporensis</i> Pfeiffer, 1855	I	Malaysia, Sarawak, Mukah	KM280983	KM281066	KM281044	KJ920309	-	KM281133	ZSM Mol 20040289	This work
<i>Microtrilla alba</i> (Gassies, 1865)	I	Australia, New South Wales	<b>HQ659942</b>	-	<b>HQ659876</b>	-	<b>HQ660008</b>	-	AMS 398688	Dayrat et al. (2011)
<i>Microtrilla occidentalis</i> (Pfeiffer, 1854)	I	Bermuda, St. George's Parish, Lover's Lake	KM280984	KM281067	KM281040	KJ920310	<b>KC206183</b>	<b>KC206239</b>	AW1760	Weigand et al. (2013)
<i>Pseudomelampus exiguus</i> (Lowe, 1832)	I	Portugal, Azores	KM280985	KM281068	KM281039	KJ920311	<b>HQ660009</b>	KM281134	CASIZ 180473	Dayrat et al. (2011)

(continued on next page)

Table 1 (continued)

Taxon	H	Locality	18S	28S	16S	12S	COI	HB	Voucher #	References
<b>Pedipodinae</b>										
<i>Marrulina jilohi</i> Hutton, 1878	I	New Zealand, Catham Island	HQ659944	–	HQ659878	–	HQ660010	–	CASIZ 180474	Dayrat et al. (2011)
<i>Pedipes jouani</i> Montrouzier, 1862	I	Japan, Okinawa, Miyakojima, Karimata Hirara	KM280986	KM281069	KM281011	KJ920312	KM281105	KM281135	SMF 346418	This work
<i>Pedipes ovatis</i> Adams, 1849	I	Bermuda, Hamilton Parish, Walsingham Pond	KM280987	KM281070	KM281041	KJ920313	–	KM281136	SMF 346419	This work
<i>Pedipes mirabilis</i> (Mühlhfeld, 1816)	I	Jamaica	HQ659945	A1465074	HQ659879	–	–	–	–	Dayrat et al. (2011)
<i>Pedipes pedipes</i> (Bruguière, 1789)	I	Portugal, Azores	KM280988	–	KM281010	KJ920314	HQ660012	KM281137	CASIZ 180476	Dayrat et al. (2011)
<b>Pythiinae</b>										
<i>Allochroa layardi</i> (Adams & Adams, 1855)	I	UAE	HQ659947	–	HQ659881	–	HQ660013	–	NHM-BM 20080090	Dayrat et al. (2011)
<i>Allochroa pleifferi</i> (Dunker, 1860)	I	Japan, Okinawa, Miyakojima, Karimata Hirara	KM280989	KM281071	KM281033	KJ920315	–	KM281138	SMF 346420	This work
<i>Cassidula angulifera</i> (Pettit de la Saussaye, 1841)	I	Australia, Queensland	KM280990	KM281072	KM281025	KJ920316	KM281106	KM281139	AMS 448736	This work
<i>Cassidula aurifelis</i> (Bruguière, 1789)	I	Thailand, Phuket, Dan-Yit	KM280991	KM281073	KM281026	KJ920317	–	KM281140	SMF 346421	This work
<i>Cassidula crassiuscula</i> Mousson, 1869	I	Japan, Okinawa, Nago, Arisu	KM280992	KM281074	KM281037	KJ920318	KM281107	KM281141	No tissue left	This work
<i>Cassidula cf. labrella</i> (Deshayes, 1830)	I	UAE	HQ659950	–	HQ659884	–	HQ660016	–	NHM-BM 20080095	Dayrat et al. (2011)
<i>Cassidula nucleus</i> (Gmelin, 1791)	I	Malaysia, Sabah, Kota Kinabalu, Mengkabong	KM280993	KM281075	KM281027	KJ920319	KM281108	KM281142	SMF 346423	This work
<i>Cassidula schmaecheriana</i> Möllendorff, 1895	I	Japan, Ishigaki-jima	KM280994	KM281076	KM281024	KJ920320	KM281109	KM281143	SMF 346424	This work
<i>Cassidula vesperilionis</i> (Lesson, 1831)	I	Malaysia, Sabah, Kota Kinabalu, Mengkabong	KM280995	KM281077	KM281023	KJ920321	KM281110	KM281144	SMF 346425	This work
<i>Laemodonta bella</i> (Adams & Adams, 1855)	I	Genbank	KM280996	KM281078	KM281020	KJ920322	KM281111	KM281145	ZSM Mol 20061924	This work
<i>Laemodonta cubensis</i> (Pleiffer, 1854)	I	Bermuda, Hamilton Parish, Walsingham Pond	KM280997	KM281079	KC206128	KJ920323	KC206182	KC206238	AW1750	Weigand et al. (2013)
<i>Laemodonta monilifera</i> (Adams & Adams, 1854)	I	UAE, Oman Gulf	KM280998	–	KM281022	KJ920324	KM281112	KM281146	NHM-BM 20080099	This work
<i>Laemodonta oceanifracta</i> (Jonas, 1845)	I	Japan, Kyushu Island, Oita Pref.	KM280999	KM281080	KM281032	KJ920325	KM281113	KM281147	SMF 346427	This work
<i>Laemodonta punctatostriata</i> Adams & Adams, 1853	I	Hong Kong	HQ659952	–	HQ659886	–	HQ660018	–	–	Dayrat et al. (2011)
<i>Laemodonta typica</i> (Adams & Adams, 1854)	I	Japan, Kagoshima, Amami, Yamato-son	KM281000	KM281081	KM281021	KJ920326	KM281114	KM281148	SMF 346426	This work
<i>Myosotella myosotis</i> (Draparnaud, 1801)	I	Bermuda, Southampton Parish, Evans Pond	KM281001	EF489360	KM281038	KJ920327	–	–	MNB 104.546	Klussmann-Kolb et al. (2008)
<i>Ophicardelus costellaris</i> Adams & Adams, 1854	I	–	EF489342	EF489362	–	–	EF489387	–	MNB 104.553	Klussmann-Kolb et al. (2008)
<i>Ophicardelus ornatus</i> (Férussac, 1821)	I	–	DQ093442	DQ279994	DQ093486	–	DQ093530	DQ093512	–	Giribet et al. (2006)
<i>Ophicardelus sulcatus</i> Adams & Adams, 1854	I	Australia, New South Wales	HQ659955	–	HQ660021	–	HQ659889	–	AMS 405360	Dayrat et al. (2011)
<i>Ovatella firminii</i> Payraudreau, 1826	I	Greece, Crete	HQ659956	–	HQ659890	–	HQ660022	–	CASIZ 180479	Dayrat et al. (2011)
<i>Ovatella vulcani</i> (Morelet, 1860)	I	Portugal, Azores	KM281002	KM281082	KM281028	KJ920328	NC_016175	KM281149	CASIZ 180486	Dayrat et al. (2011)
<i>Pleuraloba quoyi</i> (Adams & Adams, 1854)	I	Australia, New South Wales	HQ659958	–	HQ659892	–	HQ660024	–	AMS 397375	Dayrat et al. (2011)
<i>Pythia borneensis</i> Adams, 1851	I	Malaysia, Sarawak, Mukah	KM281003	KM281083	KM281031	KJ920329	KM281116	KM281150	SMF 346428	This work
<i>Pythia cecilia</i> (Philippi, 1847)	I	Papua New Guinea	HQ659959	–	HQ659893	–	HQ660025	–	UF 339082	Dayrat et al. (2011)
<i>Pythia firminii</i> Möllendorff, 1885	I	Papua New Guinea	HQ659960	–	HQ659894	–	–	–	UF 339086	Dayrat et al. (2011)
<i>Pythia scarabaeus</i> (Linnaeus, 1758)	T	Papua New Guinea	KM281004	KM281084	KM281030	KJ920330	–	KM281151	ZSM Mol 20052266	This work
<i>Pythia sp.</i>	T	Palau, Ulung Island	KM281005	KM281085	KM281029	KJ920331	–	KM281152	–	Kano et al. (2015)

Table 1 (continued)

Taxon	H	Locality	18S	28S	16S	12S	COI	H3	Voucher #	References
<b>Trimusculinae</b>										
<i>Otira ovata</i> (Brown, 1827)	1	France, Roscoff	EF489344	EF489363	EF489310	–	EF489389	–	ZSM IMI 20034194	Klussmann-Kolb et al. (2008)
<i>Smegola philippensis</i> Thilier & Ponder, 1992	1	Australia, Victoria, Phillip Is., Sunderland Bay	FJ917210	FJ917229	FJ917263	KJ920332	FJ917283	KM281155	EED-Phy-878	Dhapoli and Klussmann-Kolb (2010)
<i>Trimusculatus afer</i> (Gmelin, 1791)	1	Genbank (as T. afra)	EF489343	–	EF489309	–	EF489388	–	JLU-SZBDF 2007.001	Klussmann-Kolb et al. (2008)
<i>Trimusculatus reticulatus</i> (Sowerby I, 1835)	1	USA, California	HQ659935	–	NC_016193	NC_016193	NC_016193	–	CASIZ 177988	White et al. (2011)
<i>Trimusculatus</i> sp.	1	USA, California, Santa Barbara	KM281008	KM281088	KM281035	KJ920333	KM281117	–	No tissue left	This work
<b>ONCHIDIIDAE</b>										
<i>Onchidella borealis</i> Dall, 1872	1	Chile, Coquimbo, Los Vilos, Pichidangui	KM281006	KM281086	DQ991936	DQ991936	DQ991936	KM281153	No tissue left	Medina et al. (2011)
<i>Onchidella celtica</i> (Cuvier, 1816)	1	France, Roscoff	KM281007	KM281087	NC_012376	NC_012376	NC_012376	KM281154	No tissue left	Grande et al. (2008)
<i>Onchidella floridana</i> (Dall, 1885)	1	Genbank	AY427521	AY427486	EF489317	–	EF489392	–	EED-Phy-462	Klussmann-Kolb et al. (2008) and Vonnemann et al. (2005)
<i>Onchidella hildae</i> (Hoffmann, 1928)	1	Panama	HQ659970	–	HQ659904	–	HQ660036	–	UF 372677	Dayrat et al. (2011)
<i>Onchidium tumidum</i> Semper, 1885	1	Australia, New South Wales	HQ659971	–	HQ659905	–	HQ660037	–	UF 395149	Dayrat et al. (2011)
<i>Onchidium vaigiense</i> Quoy & Gaimard, 1824	1	Papua New Guinea	HQ659974	–	HQ659908	–	HQ660040	–	UF 366435	Dayrat et al. (2011)
<i>Peronia peronii</i> (Cuvier, 1804)	1	Guam	HQ659975	–	HQ659909	NC_016181	HQ660041	–	CASIZ 180486	Dayrat et al. (2011) and White et al. (2011)
<i>Peronia verruculata</i> (Cuvier, 1830)	1	Genbank (as <i>O. verrucosum</i> )	AY427522	AY427487	EF489316	–	EF489391	–	EED-Phy-58	Klussmann-Kolb et al. (2008) and Vonnemann et al. (2005)
<i>Peronia</i> cf. <i>verruculata</i> (Cuvier, 1830)	1	Japan, Okinawa	HQ659977	–	HQ659911	–	HQ660043	–	UF 352288	Dayrat et al. (2011)
<i>Platvindex</i> cf. <i>coriaceum</i> (Semper, 1885)	1	Mozambique	HQ659983	–	HQ659917	–	HQ660049	–	NHM-BMI 20060274	Dayrat et al. (2011)
<i>Scaphis</i> sp.	1	Philippines	HQ659984	–	HQ659918	–	HQ660050	–	UF 368518	Dayrat et al. (2011)



### 2.3. Phylogenetic analyses

Multiple sequence alignments of each locus were obtained using MAFFT v.7 (Katoh and Standley, 2013). The alignment strategy used for 18S rRNA and 28S rRNA was *-linsi* as they both have one globally alignable domain. In the case of 16S rRNA and 12S rRNA the *-qinsi* algorithm was used to take the structural information into account. Regions with ambiguous positions that could not be rigorously aligned were removed using Gblocks server ([http://molevol.cmima.csic.es/castresana/Gblocks\\_server.html](http://molevol.cmima.csic.es/castresana/Gblocks_server.html)) (Talavera and Castresana, 2007), by allowing smaller final blocks, gaps within the final blocks and less strict flanking positions. Finally, COI and H3 sequences were aligned using *-ginsi*. These protein-coding alignments were similar to the ones produced by TranslatorX (<http://www.translatorx.co.uk/>). This software uses the translated amino acid alignment to guide the alignment of nucleotide sequences (Abascal et al., 2010). Alignment lengths after trimming were 1773, 1028, 342 and 261 positions for the ribosomal genes 18S rRNA, 28S rRNA, 16S rRNA and 12S rRNA, respectively, and 654 and 327 positions for the protein-coding genes COI and H3, respectively. Best-fit models for each DNA region were selected using jModelTest2 (Guindon and Gascuel, 2003; Durrin et al., 2012) based on the Akaike information criterion (AIC; Supplementary Table S2). Data were partitioned by gene and by codon position in case of the protein-coding sequences, using the partition scheme suggested in PartitionFinder (Lanfear et al., 2014).

Phylogenetic inference via maximum likelihood (ML) was performed using RAxML-HPC2 (8.0.9) (Stamatakis, 2006; Stamatakis et al., 2008) implemented on XSEDE (Miller et al., 2010) in the CIPRES Science Gateway web portal (<http://www.phylo.org/index.php/portal/>). We followed the “the hard and slow way” suggestions in the manual: the best-likelihood tree was selected after 1000 independent runs, branch support was evaluated using bootstrapping with 1000 replicates, and confidence values were drawn on the best-scoring tree. Bayesian inference (BI) was performed using MrBayes v3.2.2 (Ronquist et al., 2012) on XSEDE (CIPRES). Two simultaneous Monte Carlo Markov Chains (MCMC) were run, with the following parameters: eight chains of 100 million generations each, a sampling frequency of 1000 generations and a burn-in of 25%. Tracer 1.6 (Rambaut et al., 2014) was used to evaluate convergence, i.e. whether effective sample size (ESS) was sufficient for all estimated parameters (>200). In this study, we consider a bootstrap value of >70% (Douady et al., 2003) and a posterior probability of >0.95 (Leaché and Reeder, 2002) as a significant nodal support. Branch support is indicated in the following way in the text: bootstrap value/posterior probability. Finally, rogue taxa, i.e. taxa that deteriorate the resolution or branch support in a consensus tree were identified using RogueNaRok (<http://rnr.h-its.org/>) (Aberer et al., 2013). Ten taxa were removed from the concatenated alignment: *Auriculastra subula* (Quoy & Gaimard, 1832), *Cassidula* cf. *labrella* (Deshayes, 1830), *Cassidula nucleus* (Gmelin, 1791), *Carychium hachijoensis* Pilsbry, 1802, *Melampus bidentatus* Say, 1822, *Melampus fasciatus* (Deshayes, 1830), *Ophicardelus sulcatus* Adams & Adams, 1854, *Onchidella hildae* (Hoffman, 1928), *Platevindex* cf. *coriaceum* (Semper, 1885), *Smeagol philippensis* Tillier & Ponder, 1992, and *Zospeum alpestre* (Freyer, 1855). The final alignment resulted in 64 taxa (55 ellobiids and nine onchidiids).

### 2.4. Topology testing

Alternative topologies were compared using Bayes factors in BEAST 1.8.0 (Drummond et al., 2012) or likelihood-based tests using CONSEL (Shimodaira and Hasegawa, 2001). The unconstrained tree was compared with several topologies; in particular we evaluated the monophyly of the Ellobiidae, including or

excluding *Trimusculus* and *Otina*, the monophyly of subfamilies that were not well supported in the phylogeny, and relationships within subfamilies (see Section 3.2). Topological comparisons were done by examining the marginal likelihoods of the alternative topologies obtained via path sampling or stepping stone algorithms (Baele et al., 2012, 2013), and by comparing the Bayes factors of different hypotheses (Kass and Raftery, 1995; McVay and Carstens, 2013). For the likelihood-based test, alternative tree topologies were constrained using Arceopteryx (Han and Zmasek, 2009). Site-wise log-likelihood values for each topology were computed in RAxML. Approximately Unbiased (AU; Shimodaira, 2002), Shimodaira–Hasegawa (SH; Shimodaira and Hasegawa, 1999) and Kishino–Hasegawa (KH; Kishino and Hasegawa, 1989) tests were carried out in CONSEL.

### 2.5. Divergence times

The final alignment was used to calculate the approximate time of divergence among taxa with a relaxed molecular clock model (Drummond et al., 2006) implemented in BEAST. Different prior distributions were used for each calibration point following the suggestions proposed by Ho and Phillips (2009) and in the “Divergence Time Estimation” tutorial (<http://treethinkers.org/divergence-time-estimation-using-beast/>). We used previously suggested ages for three nodes in order to calibrate our topology. First, the Ellobiidae can probably be traced back to the Upper Jurassic (145–160 Ma; Bandel, 1994; Wägele et al., 2008). Moreover, Jörger et al. (2010) showed that the most recent common ancestor (MRCA) of the Ellobiidae probably appeared at ~140 Ma. In addition, Weigand et al. (2013) suggested that the Carychiinae could have descended from the extinct terrestrial genus *Carychiopsis* Sandberger, 1872, which originated no earlier than the beginning of the early Cenozoic (~65 Ma), coinciding with the beginning of the Alpine orogeny.

Thus, we used a log-normal distribution which places the highest probabilities on ages older than fossil records described above. A minimum age of 130 Ma was set for the first split within the Ellobiidae (including *Trimusculus* and *Otina*), using the following parameters; mean in real space: 5, standard deviation (SD): 1, 95% highest posterior density (HPD): 130–154, and offset: 130. In addition, a minimum age of 60 Ma was set for the first split within the Carychiinae; mean in real space: 10, SD: 1, 95% HPD: 60–67 and offset: 60. Finally, a normal distribution with a mean of 155 Ma, SD of 3 and 95% HPD of 148–160 was used for the root of the tree (Ellobiidae + Onchidiidae; Jörger et al., 2010). The normal distribution is used in cases when there is little justification to weight probabilities toward a minimum bound (Ho and Phillips, 2009).

Divergence time analyses were run with the relaxed uncorrelated lognormal clock model under the Yule process for each partition with an independent site model. ESS was evaluated in Tracer 1.6. The MCMC was run ten times independently for 20 million generations, sampling every 1000 generations, and discarding the first 25% as a burn-in; single runs were combined using LogCombiner to increase the ESS. Then, trees were combined using TreeAnnotator to produce a maximum clade credibility (MCC) tree. Posterior parameters, specifically divergence times, were calculated only for nodes with a posterior probability equal to or higher than 0.95.

### 2.6. Temporal patterns of diversification

Temporal dynamics of lineage diversification was evaluated under different models of diversification using a maximum likelihood approach. The gamma statistic ( $\gamma$ ) was calculated using the package *ape* (Paradis et al., 2004). Gamma ( $\gamma$ ) describes temporal



speciation shifts, in other words,  $\gamma < 0$  indicates a decrease in diversification rates and  $\gamma > 0$  corresponds to a rate increase. The MCCR test implemented in the R package *laser* (Rabosky, 2006a) evaluated the statistical significance of the  $\gamma$  statistic (Pybus and Harvey, 2000). Lineage-through-time (LTT) plots of 1000 random trees chosen from the BEAST analysis were compared graphically against a series of trees simulated under a pure-birth process. The Akaike Information Criterion (AIC) was used to test model-fit between constant-rate (pure-birth, birth–death) and variable-rate models (DDL, DDX, yule-n-rate). Constant-rate models can be assumed as the null hypothesis while variable-rate models serve as alternative hypotheses. Thus, we calculated the AIC score for each of the models, and compared the highest AIC score of the constant-rate models (AIC<sub>H0</sub>) against the highest AIC score of the variable-rate models (AIC<sub>H1</sub>) using the following formula:  $\Delta AIC_{rc} = AIC_{H0} - AIC_{H1}$ . A  $\Delta AIC_{rc}$  higher than 0 means that the variable-rate model is favored, while a  $\Delta AIC_{rc}$  lower than 0 means that the constant-rate model is favored (Yessoufou et al., 2014). Observed  $\Delta AIC_{rc}$  significance was tested using the function `fitdAICrc.batch` in *laser* by simulating 5000 trees of 55 tips under a pure-birth process.

### 3. Results

#### 3.1. Tree topologies and relationships among subfamilies

The monophyly of the Ellobiidae (without considering the former Otinidae and Trimusculidae (Ellobiidae *sensu stricto*; Martins, 2007) was not significantly supported in our final phylogenetic hypothesis, with bootstrap and posterior probability values below our threshold (35/0.80; Fig. 1, node B). On the other hand, the clade comprising the former Otinidae, Trimusculidae and Ellobiidae (i.e. Ellobiidae *sensu lato*; Dayrat et al., 2011) received the highest support values (100/1.00) (Fig. 1, node A). The association between *Trimusculus* and *Otina* was supported only in the Bayesian analysis (65/1.00).

Both Melampodinae and Pedipedinae appeared to be paraphyletic. Within Melampodinae, the monophyly of the type genus *Melampus* Montfort, 1810 received maximum support (100/1.00). The same occurred within Pedipedinae, where the studied species of the type genus *Pedipes* formed a robust clade (100/1.00). However, *Microtralia* Dall, 1894, *Pseudomelampus* Pallary, 1900 (Melampodinae) and *Marinula* King, 1832 (Pedipedinae) clustered in a third clade that was supported only in the Bayesian analysis (57/0.99).

The clade composed of Carychiinae, Ellobiinae and Pythiinae received significant support values (90/1.00). Of these, Carychiinae and Ellobiinae appeared as sister clades (74/0.98). The monophyly of Carychiinae and its component genera *Carychium* and *Zospeum* was strongly supported (100/1.00). The Ellobiinae also appeared to be monophyletic (100/1.00); within this subfamily, *Ellobium* Rödning, 1798 and *Blauneria* Shuttlesworth, 1854 were monophyletic (100/0.99).

The monophyly of Pythiinae was also strongly supported (99/1.00). This subfamily comprised two main subclades, one including the monophyletic genus *Pythia* and its sister taxon *Myosotella myosotis* (Draparnaud, 1801) (91/1.00), and the other comprising *Allochroa* Ancey, 1887, *Cassidula* Férussac, 1821, *Laemodonta* Philippi, 1846, *Ophicardelus* Beck, 1838, *Ovatella* Bivona-Bernardi, 1832 and *Pleuroloba* Hyman, Rouse & Ponder, 2005 (99/1.00). Each of the genera *Allochroa* (100/1.00), *Cassidula* (77/0.97), *Laemodonta* (74/1.00) and *Ovatella* (100/1.00) was revealed as monophyletic in the analyses.

The tree without pruning rogue branches (74 taxa) presented a similar topology and support values (Supplementary Fig. S1). Ellobiidae *sensu lato* (Dayrat et al., 2011) received significant support

(100/1.00). *Otina*, *Smeagol* and *Trimusculus* collectively formed a clade nested within the Ellobiidae, with significant support only in the Bayesian analysis (18/0.99). The same dataset presented significant support for the sister relationship of Carychiinae and Ellobiinae (88/0.99). On the other hand, *Cassidula* and *Ophicardelus* appeared polyphyletic.

Overall, our topologies, including 64 or 74 taxa, strongly support an independent and parallel colonization of terrestrial habitats within Ellobiidae in *Pythia* and Carychiinae.

#### 3.2. Topology tests

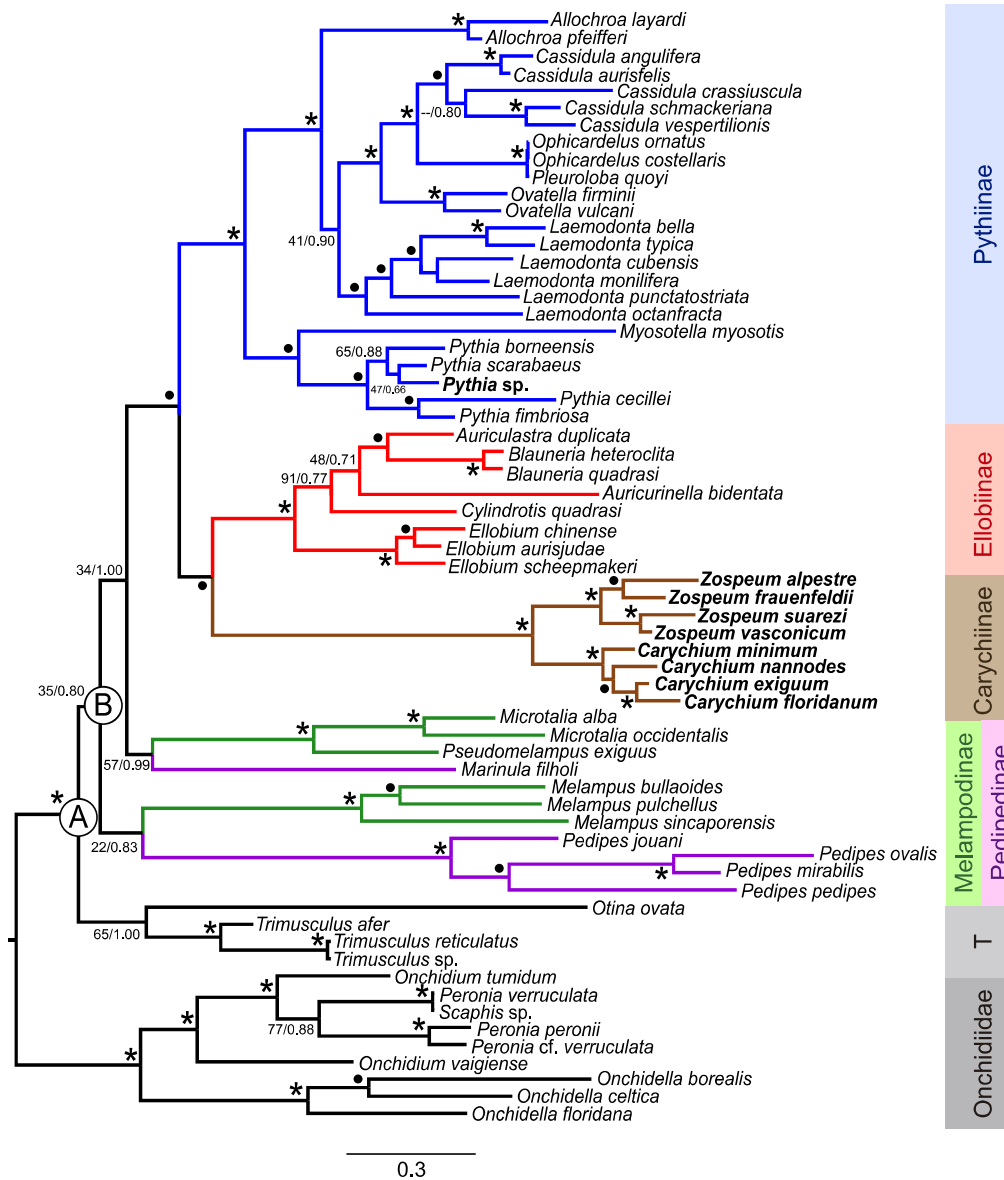
Our unconstrained phylogeny showed the following characteristics: Ellobiidae *sensu stricto* (Martins, 2007; node B in Fig. 1) did not receive significant support, whereas Ellobiidae *sensu lato* (Dayrat et al., 2011; node A in Fig. 1) constituted a robust clade. Carychiinae, Ellobiinae and Pythiinae collectively formed a sub-clade with significant statistical support. All alternative topologies performed better than the unconstrained topology. In particular, the constrained monophyly of the clade Carychiinae + Ellobiinae + Pythiinae (clade CEP) assuming Ellobiidae *sensu stricto* presented the best marginal likelihoods (Supplementary Table S3a). However, there is no decisive evidence to reject Ellobiidae *sensu lato* (Bayes Factors <2) using the stepping stone estimation (A4 vs. A3, and A6 vs. A5; Supplementary Table S4). In addition, our results showed that there is strong evidence for the clade CEP (A5 vs. A0, A6 vs. A0, and A8 vs. A0; Supplementary Table S4). Furthermore, we evaluated different hypotheses to resolve the sister group of Carychiinae (Supplementary Table S3b). There is positive evidence (BF > 2) in favor of Ellobiinae as its sister group in the unconstrained tree (B1 vs. B2), and in the topology assuming both monophyletic Ellobiidae *sensu lato* plus the monophyly of the clade CEP (B7 vs. B8: BF > 2). Also, there is very strong support for a sister group relationship of Carychiinae and Ellobiinae assuming Ellobiidae *sensu stricto* (B5 vs. B6: BF > 10; Supplementary Table S4). According to these results, our data support Ellobiinae as sister group to Carychiinae. Alternative topologies tested with CONSEL (Supplementary Table S5) in a maximum likelihood framework presented similar results: the monophyly of Pedipedinae or Melampodinae was clearly rejected ( $p < 0.05$ ), while Ellobiidae *sensu lato* cannot be rejected ( $p > 0.05$ ).

#### 3.3. Divergence times

The topology obtained in BEAST (Fig. 2) based on the concatenated six markers corroborates those obtained in RAxML and MrBayes analyses (Fig. 1). Previous chronograms showed that Ellobiidae originated in the Mesozoic (Jörger et al., 2010). Our calibration suggested that early divergences within the family Ellobiidae, including the splits into Pythiinae, Carychiinae and Ellobiinae, occurred in the Lower Cretaceous. Major diversification events occurred during the Upper Cretaceous (Mesozoic) and Paleogene (Cenozoic). Most recent common ancestors of Ellobiinae, *Pedipes*, and two major clades of Pythiinae can be traced back to the Upper Cretaceous. Finally, the invasion of the land habitats occurred independently and in different geological times in Ellobiidae, probably once in the Upper Cretaceous by the Carychiinae, and later in the Miocene by a subclade of the genus *Pythia*. The extant terrestrial genera *Carychium* and *Zospeum* diversified during the Oligocene and Miocene.

#### 3.4. Temporal patterns of diversification

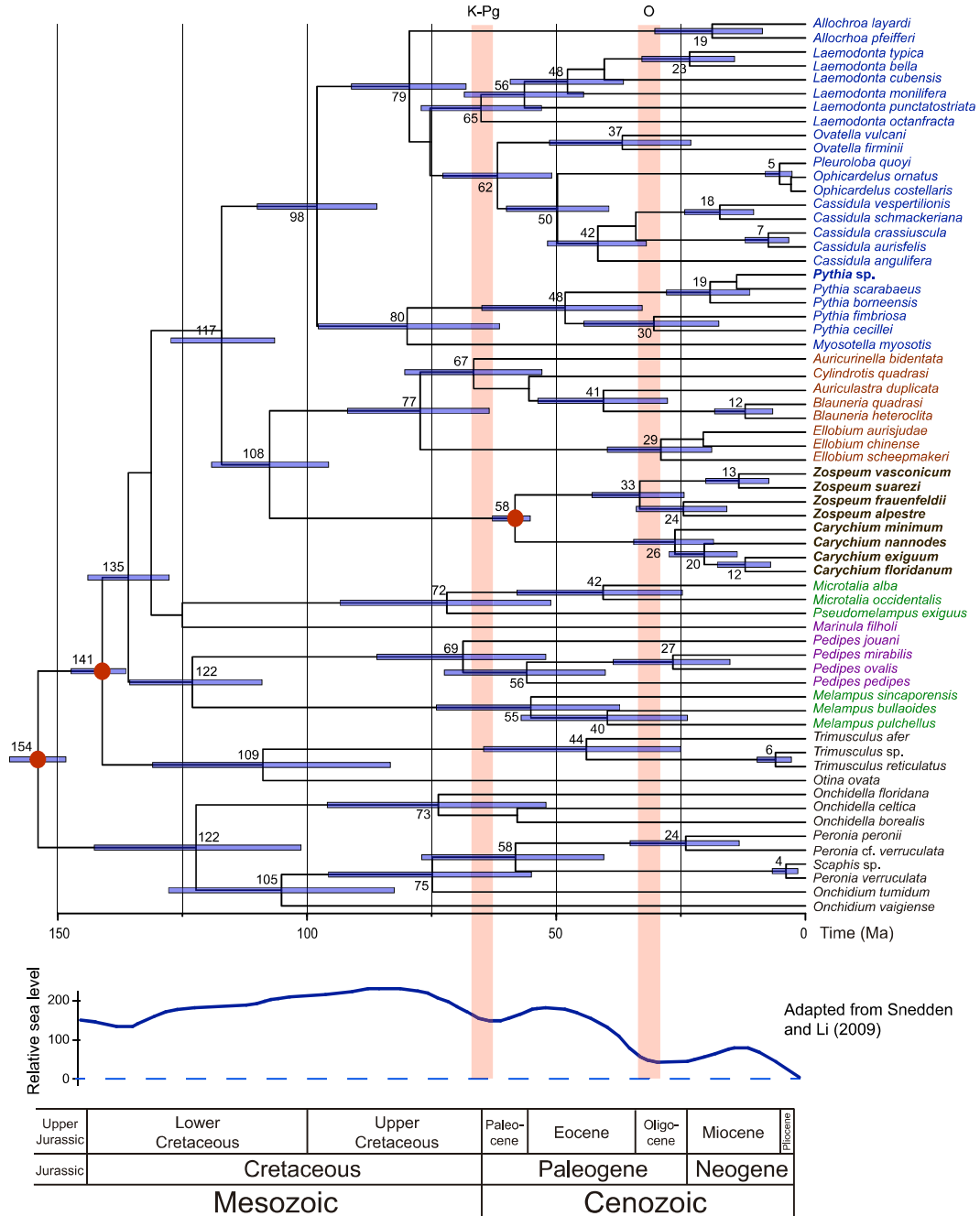
The lineage-through-time (LTT) plots fit well in a pattern of constant lineage aggregation through time (Supplementary



**Fig. 1.** Phylogram of the relationships within Ellobiidae inferred from Bayesian analysis of concatenated six-gene sequences. Posterior probabilities (PP, left) and bootstrap values (BS, right) are indicated on branches. Asterisks denote BS ≥ 90% and PP ≥ 0.99; dots are for clades with BS ≥ 70% and PP ≥ 0.95; see Section 3 for individual support values. Colors correspond to each of the subfamilies as delineated by Martins (2007); terrestrial taxa are highlighted in bold. Abbreviations: A, Ellobiidae *sensu lato* (Dayrat et al., 2011); B, Ellobiidae *sensu stricto* (Martins, 2007); T, Trimusculinae. For the interpretation of color references, the reader is referred to the web version of the article.

Fig. S2). We observed one minor stasis at around 75–95 Ma followed by a slight increase in the diversification rate between 45 Ma and 75 Ma, albeit within the distribution of the constant-rate model. In addition, the MCCR test found the value of the gamma statistic  $\gamma = -1.23$  as non-significant ( $p > 0.05$ ), assuming either 100 (Mordan and Wade, 2008) or 250 extant species in the family (Weigand et al., 2013). The AIC test in *laser* favored the

yule5rate model, i.e. a pure-birth model with five different diversification rates (Supplementary Table S6). However, three other models including DDL (logistic diversity-dependent diversification), yule3rate and yule4rate could not be rejected because their  $\Delta AIC$  were also below the critical value of 4 for small phylogenies (Rabosky, 2006b; Matos-Maravi et al., 2014). The yule5rate showed changes in diversification rates at 12 Ma and 40 Ma.



**Fig. 2.** BEAST chronogram of the Ellobiidae inferred from concatenated six-gene sequences and three calibration points (red circles). Bars express 95% highest posterior density for nodes with a PP of >0.5. Numbers indicate estimated nodal ages in millions of years. Terrestrial taxa are highlighted in bold. Light orange boxes indicate the Cretaceous–Paleogene boundary (K–Pg) and the regression of the sea level during the Oligocene (O). For the interpretation of color references, the reader is referred to the web version of the article.

## 4. Discussion

### 4.1. Phylogenetic implications

The most significant step in the evolution of the Heterobranchia was the invasion of freshwater and terrestrial realms by panpulmonates (Klussmann-Kolb et al., 2008; Jörger et al., 2010). These invasions triggered the adaptive radiation that led to about one third of the extant molluscan diversity (Kocot et al., 2013). We now have a solid hypothesis for the phylogeny of Ellobiidae to test the monophyly and sister relationships of the currently recognized subfamilies and to establish the temporal frame of the habitat transitions.

Our data strongly support a clade containing Carychiinae, Ellobiinae and Pythiinae (clade CEP), as well as its subclade Carychiinae+Ellobiinae and respective monophyly of the three subfamilies. Interestingly, Martins (1996) has described morphological characteristics that differentiate these subfamilies from Melampodinae and Pedipedinae; the visceral nerve ring is long in the former, while it is short in the latter. Moreover, the presence of an entirely glandular pallial gonoduct in the clade CEP contrasts with an only proximally glandular duct in Pedipedinae or a nonglandular condition in Melampodinae (Martins, 1996). We propose that these nervous and reproductive traits are shared synapomorphies for the clade of Carychiinae, Ellobiinae and Pythiinae. Within this clade, Pythiinae is morphologically differentiated from Carychiinae and Ellobiinae in having a longer right parietovisceral connective (Hyman et al., 2005). Pythiinae as the most ancestral taxon among the above five subfamilies (Martins, 2007) was rejected by our molecular phylogenetic reconstruction.

We can also compare our results to the previous molecular phylogenetic hypothesis by Dayrat et al. (2011). Notably, the relationships within Pythiinae are better resolved in our trees; *Myosotella* is closest to *Pythia* (but not to *Carychium*) and these constitute a robust clade that is sister to the remaining genera in the subfamily (Fig. 1). In addition, we found the monophyletic Ellobiinae comprises two subclades, one with *Ellobium* and the other with the rest of the subfamily including *Auriculinea* Tausch, 1886, which was recovered closer to some species of Melampodinae in Dayrat et al. (2011).

*Pseudomelampus* and *Microtralia* (currently classified in Melampodinae) constitute a strongly supported clade, with *Marinula* (Pedipedinae) as its possible sister group. Actually, the former two genera were proposed to belong to either Pedipedinae or Melampodinae based on morphological analyses (Martins, 2007), conforming to the non-monophyletic nature of these subfamilies in our reconstruction. *Pseudomelampus*, *Microtralia* and *Marinula* therefore may form an independent subfamily, however an increased taxonomic sampling for molecular analyses and re-evaluation of morphological characters are needed to formally introduce a new classification for the entire Ellobiidae.

We validated the inclusion of *Otina*, *Smeagol* and *Trimusculus* into Ellobiidae as the exclusive members of the subfamily Trimusculinae (Klussmann-Kolb et al., 2008; Dayrat et al., 2011). Although *Otina* and *Trimusculus* were found to potentially form the first extant offshoot of the family (Fig. 1), the inclusion of *Smeagol* rendered this Trimusculinae nested within Ellobiidae *sensu stricto* (Supplementary Fig. S1). Regardless, the strongly supported monophyly of Ellobiidae *sensu lato* in combination with such uncertainties in the early history of the family seem to justify the above taxonomic treatment. Within Trimusculinae, *Otina* and *Smeagol* share morphological characteristics such as the presence of the ocular ridge, division of the foot into the propodium and metapodium, and the position of the heart dorsal to the kidney (Tillier and Ponder, 1992). This clade was neither supported nor rejected in our analyses (Supplementary Fig. S1).

### 4.2. Multiple invasions of the land

The Ellobiidae represent an ecologically diverse family that has adapted to intertidal and terrestrial habitats. Our phylogenetic reconstruction indicates that the invasion of the land occurred at least twice in this family, first by the lineage leading to the extant Carychiinae and later within the genus *Pythia* (Fig. 2), presumably via the intertidal rocky shores, estuaries and mangroves, where all other ellobiids occur. The importance of these transitional environments has been stressed for terrestrialization in other subclades of Panpulmonata, including Acochlidia, Glacidorboidea, Hygrophila, Stylommatophora and Systellommatophora (Klussmann-Kolb et al., 2008; Holznagel et al., 2010; Jörger et al., 2010; Kano et al., 2015).

The intertidal zone usually exhibits a wide range of variation in physical factors, allowing species to develop new adaptations to changing conditions (Dayrat et al., 2011), although the shift from an intertidal to a terrestrial life requires further physiological and behavioral modification (Little, 1983). Experimental studies have demonstrated remarkable tolerance of ellobiids to desiccation and various salt concentrations. *Melampus bidentatus* survives a loss of almost 80% of its body water (Price, 1980), whereas *Myosotella myosotis* withstands changes in salt concentration from freshwater to nearly three times the salinity of seawater (Seelemann, 1968). Such tolerance to desiccation and freshwater, along with the ability to breath outside the water, should have facilitated the succeeding invasion of the land by the Ellobiidae.

### 4.3. Tempo and mode in diversification

It has been proposed that the Cretaceous–Paleogene (K–Pg) mass extinction event facilitated habitat transitions and succeeding species diversification of plants and animals (Procheş et al., 2014; but see Stadler, 2011). This onset of the realm transitions may be attributable not only to the filling of empty ecological niches, but also the creation of marginal habitats that facilitated the transitions. For instance, the diversification of angiosperms resulted in increased areas of transitional habitats between the sea and land, such as mangroves and seagrass beds (Procheş et al., 2014). In particular, the origin of mangroves during the Paleocene (Ellison et al., 1999) could have provided a new suitable environment for the diversification of intertidal species. Sea level fluctuations could also have opened novel niches that contributed to the diversification of marine and terrestrial species close to the marginal zones. The sea level was much higher during the Cretaceous than at present while significantly lower in the Oligocene (Haq et al., 1987; Miller et al., 2005; Snedden and Liu, 2010).

However, the past climatic and biotic changes do not seem to exhibit an obvious correlation with the habitat transitions from intertidal to terrestrial realms or historical diversification rates in Ellobiidae. Our molecular time calibration suggests that the entirely terrestrial Carychiinae diverged from its intertidal sister (Ellobiinae) in the warm Cretaceous period. The other terrestrialization event in the family, exemplified by an unidentified species of *Pythia* (Kano et al., 2015), seems to have occurred in the Miocene after the sea level regression in the Oligocene (Fig. 2). The LTT plots also did not reveal a significant deviation from a constant rate of diversification, suggesting that the ellobiid evolution was not strongly affected by the K–Pg event that wiped out nearly half of the terrestrial species of the time (Jiang et al., 2010; Breure and Romero, 2012). A previous analysis for the Carychiinae also suggests that major environmental changes did not affect their diversification rate, possibly because large-scale bioclimatic factors had less impact for the moist leaf-litter and subterranean habitats of *Carychium* and *Zospeum*, respectively (Weigand et al., 2013).



Similar protection in moist microhabitats and remarkable tolerance to harsh conditions might have contributed to the continuous diversification in the intertidal lineages of Ellobiidae.

## 5. Conclusions

The present molecular phylogeny of Ellobiidae supports and extends previous hypotheses, while offering a new framework for understanding their morphological, ecological and genetic diversity. Independent invasions of the land are strongly suggested, first by the Carychiinae in the Cretaceous and later by a subclade of the genus *Pythia* in the Miocene. These habitat transitions or historical diversification rates of Ellobiidae do not seem to be directly associated with the documented global change events including the Cretaceous–Paleogene mass extinction and the Oligocene regression, possibly due to their physiological tolerance to varying conditions. Future studies considering their morphological traits related to osmoregulation, excretion and breathing, as well as genomic approaches to identify candidate genes in such adaptation for the terrestrial environment, would further elucidate patterns, processes and mechanisms of habitat transitions in Gastropoda.

## Acknowledgments

This work was supported by the German funding program “LOEWE – Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz” of the Hesse’s Ministry of Higher Education, Research and the Arts, and by JSPS KAKENHI (No. 26291077). P.R. also received a PhD scholarship from CONCYTEC – Peru. We would like to thank Benoît Dayrat, Antonio de Frias Martins, Adrienne Jochum and Alexander Weigand for providing samples or suggestions for this work, and to Claudia Nesselhauf for her support in the laboratory. We also thank the editor and reviewers for their comments on the manuscript. This is Contribution #234, Bermuda Biodiversity Project (BBP), Bermuda Aquarium, Natural History Museum and Zoo, Department of Conservation Services.

## Appendix A. Supplementary material

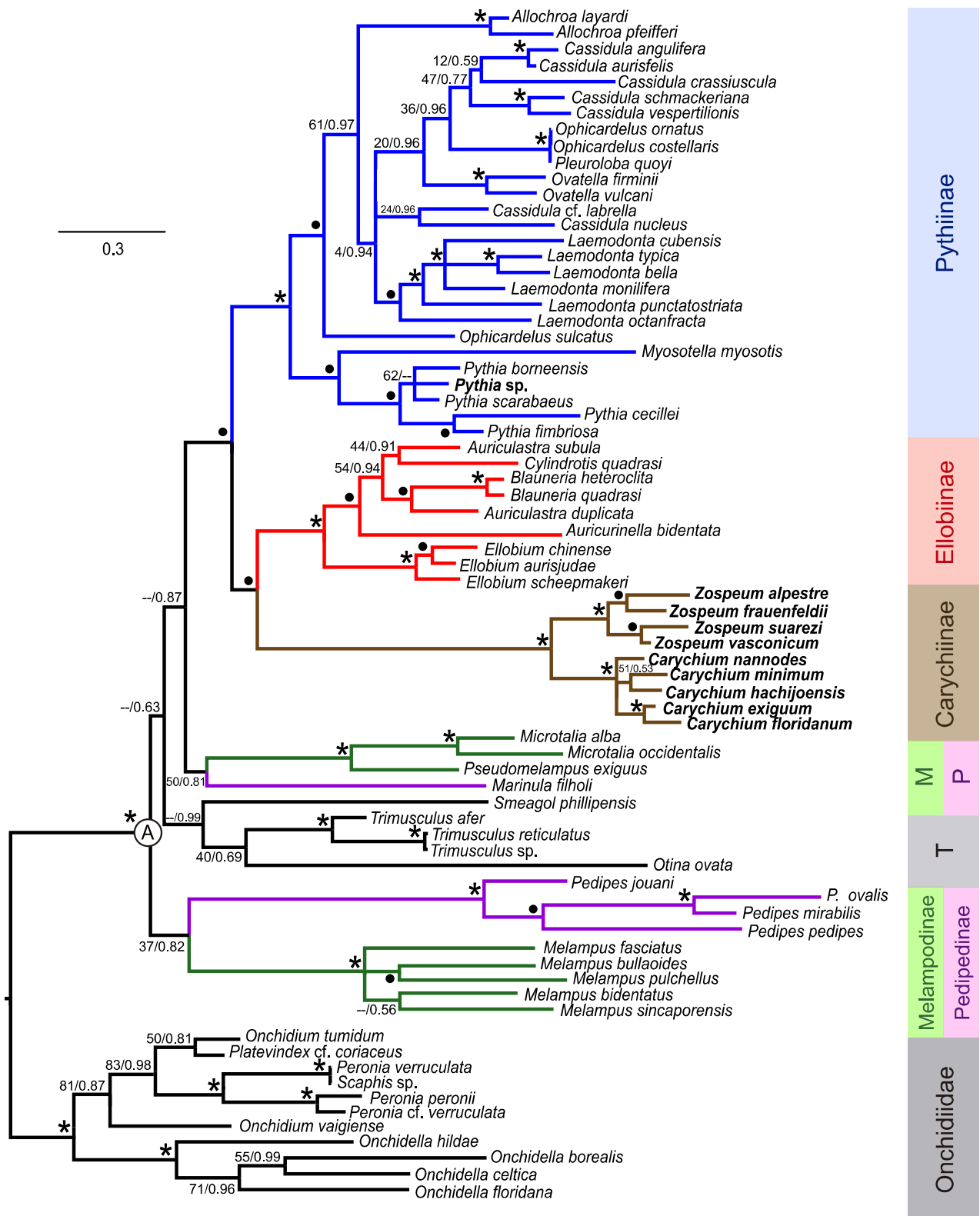
Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ympev.2015.12.014>.

## References

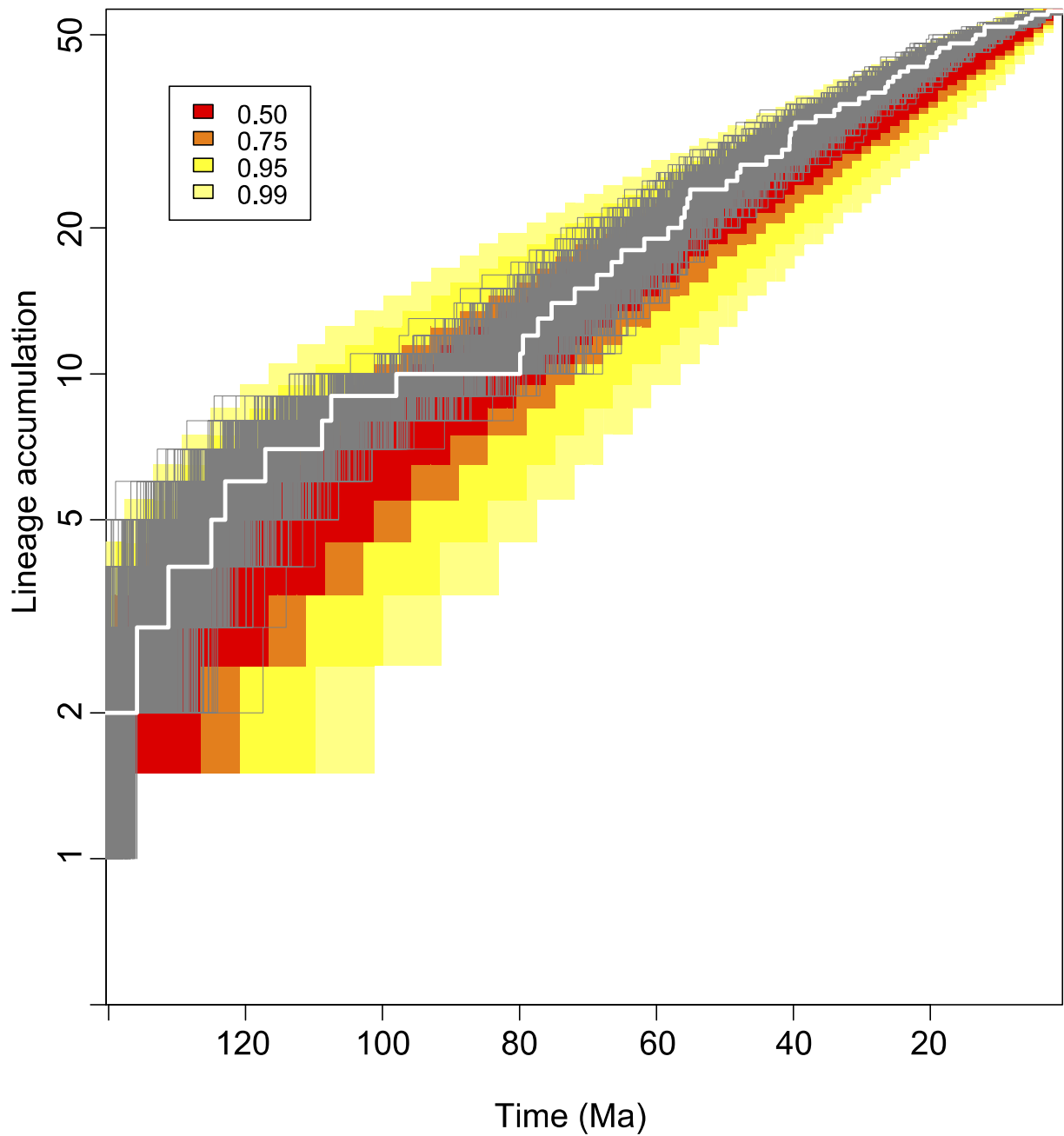
- Abascal, F., Zardoya, R., Telford, M.J., 2010. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucl. Acids Res.* 38, W7–W13.
- Aberer, A.J., Krompass, D., Stamatakis, A., 2013. Pruning rogue taxa improves phylogenetic accuracy: an efficient algorithm and webservice. *Syst. Biol.* 62, 162–166.
- Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M.A., Alekseyenko, A.V., 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol. Biol. Evol.* 29, 2157–2167.
- Baele, G., Li, W.L., Drummond, A.J., Suchard, M.A., Lemey, P., 2013. Accurate model selection of relaxed molecular clocks in bayesian phylogenetics. *Mol. Biol. Evol.* 30, 239–243.
- Bandel, K., 1994. Triassic Euthyneura (Gastropoda) from St. Cassian Formation (Italian Alps) with a discussion on the evolution of the Heterostropha. *Freib. Forschh. Reihe. C. 2*, 79–100.
- Barker, G.M., 2001. Gastropods on land: phylogeny, diversity and adaptive morphology. In: Barker, G.M. (Ed.), *The Biology of Terrestrial Molluscs*. CABI Publishing, Wallingford, pp. 1–146.
- Breure, A.S.H., Romero, P.E., 2012. Support and surprises: molecular phylogeny of the land snail superfamily Orthalicoidae using a three-locus gene analysis with a divergence time analysis and ancestral area reconstruction (Gastropoda: Stylommatophora). *Arch. Mollusk.* 141, 1–20.
- Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772.
- Dayrat, B., Conrad, M., Balayan, S., White, T.R., Albrecht, C., Golding, R., Gomes, S.R., Harasewych, M.G., Martins, A.M., 2011. Phylogenetic relationships and evolution of pulmonate gastropods (Mollusca): new insights from increased taxon sampling. *Mol. Phylogenet. Evol.* 59, 425–437.
- Dayrat, B., Tillier, S., 2002. Evolutionary relationships of euthyneuran gastropods (Mollusca): a cladistic re-evaluation of morphological characters. *Zool. J. Linn. Soc.* 135, 403–470.
- Dinapoli, A., Klussmann-Kolb, A., 2010. The long way to diversity—phylogeny and evolution of the Heterobranchia (Mollusca: Gastropoda). *Mol. Phylogenet. Evol.* 55, 60–76.
- Douady, C.J., Delsuc, F., Boucher, Y., Doolittle, W.F., Douzery, E.J., 2003. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Mol. Biol. Evol.* 20, 248–254.
- Drummond, A.J., Ho, S.Y., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4, e88.
- Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973.
- Ellison, A.M., Farnsworth, E.J., Merkt, R.E., 1999. Origins of mangrove ecosystems and the mangrove biodiversity anomaly. *Glob. Ecol. Biogeogr.* 8, 95–115.
- Giribet, G., Okusu, A., Lindgren, A.R., Huff, S.W., Schrod, M., Nishiguchi, M.K., 2006. Evidence for a clade composed of molluscs with serially repeated structures: monoplacophorans are related to chitons. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7723–7728.
- Grande, C., Templado, J., Zardoya, R., 2008. Evolution of gastropod mitochondrial genome arrangements. *BMC Evol. Biol.* 8, 61.
- Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704.
- Han, M.V., Zmasek, C.M., 2009. PhyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinform.* 10, 356.
- Haq, B.U., Hardenbol, J., Vail, P.R., 1987. Chronology of fluctuating sea levels since the Triassic. *Science* 235, 1156–1167.
- Ho, S.Y., Phillips, M.J., 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst. Biol.* 58, 367–380.
- Holzner, W.E., Colgan, D.J., Lydeard, C., 2010. Pulmonate phylogeny based on 28S rRNA gene sequences: a framework for discussing habitat transitions and character transformation. *Mol. Phylogenet. Evol.* 57, 1017–1025.
- Hyman, I.T., Rouse, G.W., Ponder, W.F., 2005. Systematics of *Ophicardelus* (Gastropoda: Heterobranchia: Ellobiidae). *Mollusc. Res.* 25, 14–26.
- Jiang, S.J., Bralower, T.J., Patzkowsky, M.E., Kump, L.R., Schueth, J.D., 2010. Geographic controls on nannoplankton extinction across the Cretaceous/Paleogene boundary. *Nat. Geosci.* 3, 280–285.
- Jochum, A., de Winter, A.J., Weigand, A.M., Gómez, B., Prieto, C., 2015. Two new species of *Zospeum* Bourguignat, 1856 from the Basque-Cantabrian Mountains, Northern Spain (Eupulmonata, Ellobioidea, Carychiidae). *ZooKeys* 483, 81–96.
- Jörger, K.M., Stöger, I., Kano, Y., Fukuda, H., Knebelberger, T., Schrödl, M., 2010. On the origin of Acochlidia and other enigmatic euthyneuran gastropods, with implications for the systematics of Heterobranchia. *BMC Evol. Biol.* 10, 323.
- Kameda, Y., Kato, M., 2011. Terrestrial invasion of pomatiopsid gastropods in the heavy-snow region of the Japanese Archipelago. *BMC Evol. Biol.* 11, 118.
- Kano, Y., Neusser, T.P., Fukumori, H., Jörger, K.M., Schrödl, M., 2015. Sea-slug invasion of the land. *Biol. J. Linn. Soc.* 116, 253–259.
- Kass, R.E., Raftery, A.E., 1995. Bayes factors. *J. Am. Stat. Assoc.* 90, 773–795.
- Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780.
- Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* 29, 170–179.
- Klussmann-Kolb, A., Dinapoli, A., Kuhn, K., Streit, B., Albrecht, C., 2008. From sea to land and beyond—new insights into the evolution of euthyneuran Gastropoda (Mollusca). *BMC Evol. Biol.* 8, 57.
- Kocot, K.M., Halanych, K.M., Krug, P.J., 2013. Phylogenomics supports Panpulmonata: opisthobranch paraphyly and key evolutionary steps in a major radiation of gastropod molluscs. *Mol. Phylogenet. Evol.* 69, 764–771.
- Lanfear, R., Calcott, B., Kainer, D., Mayer, C., Stamatakis, A., 2014. Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evol. Biol.* 14, 82.
- Leaché, A.D., Reeder, T.W., 2002. Molecular systematics of the Eastern Fence Lizard (*Sceloporus undulatus*): a comparison of Parsimony, Likelihood, and Bayesian approaches. *Syst. Biol.* 51, 44–68.
- Lillywhite, H.B., 2012. Evolutionary history and the conquest of land. *Bioscience* 62, 517–518.
- Little, C., 1983. *The Colonisation of Land: Origins and Adaptations of Terrestrial Animals*. Cambridge University Press, Cambridge.
- Little, C., 1990. *The Terrestrial Invasion: An Ecophysiological Approach to the Origins of Land Animals*. Cambridge University Press, Cambridge.
- Martins, A.M., 1995. A new species of *Pythia* Röding, 1798 (Pulmonata, Ellobiidae), from New Ireland, Papua New Guinea. *Mollusc. Res.* 16, 59–67.
- Martins, A.M., 1996. Relationships within Ellobiidae. In: Taylor, J.D. (Ed.), *Origin and Evolutionary Radiation of the Mollusca*. Oxford University Press, Oxford, pp. 285–294.
- Martins, A.M., 2007. Morphological and anatomical diversity within the Ellobiidae (Gastropoda, Pulmonata, Archaeopulmonata). *Vita Malacol.* 4, 1–28.
- Matos-Maravi, P., Aguila, R.N., Pena, C., Miller, J.Y., Sourakov, A., Wahlberg, N., 2014. Causes of endemic radiation in the Caribbean: evidence from the historical biogeography and diversification of the butterfly genus *Calisto* (Nymphalidae: Satyrinae: Satyrini). *BMC Evol. Biol.* 14, 199.
- McVay, J.D., Carstens, B., 2013. Testing monophyly without well-supported gene trees: evidence from multi-locus nuclear data conflicts with existing taxonomy in the snake tribe Thamnopini. *Mol. Phylogenet. Evol.* 68, 425–431.

- Medina, M., Lal, S., Valles, Y., Takaoka, T.L., Dayrat, B.A., Boore, J.L., Gosliner, T., 2011. Crawling through time: transition of snails to slugs dating back to the Paleozoic, based on mitochondrial phylogenomics. *Mar. Genomics* 4, 51–59.
- Miller, K.G., Kominz, M.A., Browning, J.V., Wright, J.D., Mountain, G.S., Katz, M.E., Sugarman, P.J., Cramer, B.S., Christie-Blick, N., Pekar, S.F., 2005. The Phanerozoic record of global sea-level change. *Science* 310, 1293–1298.
- Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE), IEEE, New Orleans, pp. 1–8.
- Mordan, P., Wade, C., 2008. Heterobranchia II: The Pulmonata. In: Ponder, W.F., Lindberg, D. (Eds.), *Phylogeny and Evolution of the Mollusca*. University of California Press, Berkeley, pp. 409–426.
- Morton, J.E., 1955. The evolution of the Ellobiidae with a discussion on the origin of the Pulmonata. *Proc. Zool. Soc. Lond.* 125, 127–168.
- Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of Phylogenetics and evolution in R language. *Bioinformatics* 20, 289–290.
- Price, C.H., 1980. Water relations and physiological ecology of the salt marsh snail, *Melampus bidentatus* Say. *J. Exp. Mar. Biol. Ecol.* 45, 51–67.
- Procheş, S., Polgar, G., Marshall, D.J., 2014. K–Pg events facilitated lineage transitions between terrestrial and aquatic ecosystems. *Biol. Lett.* 10, 20140010.
- Pybus, O.G., Harvey, P.H., 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc. R. Soc. B* 267, 2267–2272.
- Rabosky, D.L., 2006a. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evol. Bioinform.* 2, 247–250.
- Rabosky, D.L., 2006b. Likelihood methods for detecting temporal shifts in diversification rates. *Evolution* 60, 1152–1164.
- Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J., 2014. Tracer v1.6.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542.
- Schrödl, M., 2014. Time to say “Bye-bye Pulmonata”? *Spixiana* 37, 161–164.
- Schrödl, M., Jörger, K.M., Wilson, N.G., 2011. A reply to Medina et al. (2011): crawling through time: transition of snails to slugs dating back to the Paleozoic based on mitochondrial phylogenomics. *Mar. Genomics* 4, 301–303.
- Seelemann, U., 1968. Zur Überwindung der biologischen Grenze Meer–Land durch Mollusken. Untersuchungen an *Alderia modesta* (Opisth.) und *Ovatella myosotis* (Pulmonat.). *Oecologia* 1, 130–154.
- Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* 51, 492–508.
- Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16, 1114–1116.
- Shimodaira, H., Hasegawa, M., 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17, 1246–1247.
- Snedden, J., Liu, C., 2010. A compilation of Phanerozoic sea-level change, coastal onlaps and recommended sequence designations. *Search and Discovery* 40594.
- Stadler, T., 2011. Mammalian phylogeny reveals recent diversification rate shifts. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6187–6192.
- Stamatakis, A., 2006. RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690.
- Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAXML Web servers. *Syst. Biol.* 57, 758–771.
- Stöger, I., Schrödl, M., 2013. Mitogenomics does not resolve deep molluscan relationships (yet?). *Mol. Phylogenet. Evol.* 69, 376–392.
- Talavera, G., Castresana, J., 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56, 564–577.
- Tillier, S., Ponder, W.F., 1992. New species of *Smeagol* from Australia and New Zealand, with a discussion of the affinities of the genus (Gastropoda: Pulmonata). *J. Mollusc. Stud.* 58, 135–155.
- Vonnemann, V., Schrödl, M., Klussmann-Kolb, A., Wägele, H., 2005. Reconstruction of the phylogeny of the Opisthobranchia (Mollusca, Gastropoda) by means of 18S and 28S rRNA sequences. *J. Mollusc. Stud.* 71, 113–125.
- Wägele, H., Klussmann-Kolb, A., Verbeek, E., Schrödl, M., 2014. Flashback and foreshadowing—a review of the taxon Opisthobranchia. *Org. Divers. Evol.* 14, 133–149.
- Wägele, H., Klussmann-Kolb, A., Vonnemann, V., Medina, M., 2008. Heterobranchia I: the opisthobranchia. In: Ponder, W., Lindberg, D. (Eds.), *Phylogeny and Evolution of the Mollusca*. University of California Press, Berkeley, pp. 385–408.
- Weigand, A.M., Jochum, A., Pfenninger, M., Steinke, D., Klussmann-Kolb, A., 2011. A new approach to an old conundrum—DNA barcoding sheds new light on phenotypic plasticity and morphological stasis in microsnails (Gastropoda, Pulmonata, Carychiidae). *Mol. Ecol. Resour.* 11, 255–265.
- Weigand, A.M., Jochum, A., Slapnik, R., Schnitzler, J., Zarza, E., Klussmann-Kolb, A., 2013. Evolution of microgastropods (Ellobioidea, Carychiidae): integrating taxonomic, phylogenetic and evolutionary hypotheses. *BMC Evol. Biol.* 13, 18.
- White, T.R., Conrad, M.M., Tseng, R., Balayan, S., Golding, R., Martins, A.M., Dayrat, B. A., 2011. Ten new complete mitochondrial genomes of pulmonates (Mollusca: Gastropoda) and their impact on phylogenetic relationships. *BMC Evol. Biol.* 11, 295.
- Yessoufou, K., Bamigboye, S.O., Daru, B.H., van der Bank, M., 2014. Evidence of constant diversification punctuated by a mass extinction in the African cycads. *Ecol. Evol.* 4, 50–58.
- Zapata, F., Wilson, N.G., Howison, M., Andrade, S.C., Jörger, K.M., Schrödl, M., Goetz, F.E., Giriabet, G., Dunn, C.W., 2014. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. *Proc. R. Soc. B* 281, 20141739.

Additional File 1. Phylogram of the relationships within Ellobiidae inferred from concatenated six-gene sequences. Posterior probabilities (PP, left) and bootstrap values (BS, right) are indicated on branches. Asterisks denote BS=90% and PP=0.99; dots are for clades with BS=70% PP=0.95; see Results for individual support values. Colors correspond to each of the subfamilies as delineated by Martins (2007); terrestrial taxa are highlighted in bold. Abbreviations: A, Ellobiidae *sensu lato* (Dayrat et al., 2011); B, Ellobiidae *sensu stricto* (Martins, 2007); M, Melampodinae; P, Pedipedinae; T, Trimusculinae.



Additional File 2. Clade diversification within Ellobiidae. Lineage-through-time plots (LTT) of 1000 trees selected from the BEAST divergence time analysis were compared against a simulated pattern of constant lineage accumulation through time. Legend colors represent the confidence intervals of simulated trees under the null model (Yule: Pure-birth process). The white line is the consensus ultrametric maximum credibility tree obtained in BEAST. The number of lineages is presented in a logarithmic scale. Time scale is in million years before the present.





*Publication 2: Romero PE, Weigand AM, Pfenninger M. Positive selection on pan-pulmonate mitogenomes provide new clues on adaptations to terrestrial life. BMC Evolutionary Biology. 2016;16(1):164.*

## Anlage 2

### Erklärung zu den Autorenanteilen

an der Publikation / des Manuskripts (Titel): Positive selection on panpulmonate mitogenomes provide new clues on adaptations to terrestrial life.

Status (*submitted, under review, in revision, accepted*): Accepted, 08.08.2016. Published online, 22.08.2016.

Ab *accepted* Name der Zeitschrift: BMC Evolutionary Biology. 2016;16(1):164.

Beteiligte Autoren (Namen mit eindeutigen Kürzeln):

- PER: Pedro Eduardo Romero
- AMW: Alexander M. Weigand
- MP: Markus Pfenninger

### Was hat der Promovierende bzw. was haben die Koautoren beigetragen?

#### (1) zu Entwicklung und Planung

Promovierender PER: 50%

Co-Autor MP: 30%

Co-Autor AMW: 20%

#### (2) zur Durchführung der einzelnen Untersuchungen und Experimente

Promovierender PER: 50%. DNA extraction, amplification (PCR) and sequencing. Gel electrophoresis.

Co-Autor AMW: 50%. DNA extraction, amplification (PCR) and sequencing. Gel electrophoresis.

#### (3) zur Erstellung der Datensammlung und Abbildungen

Promovierender PER: 30%. Submission of the mitogenome sequences to Genbank.

Co-Autor AMW: 35%. Sampling of *Carychium*.

Co-Autor MP 35%. Sampling of *Candidula* and *Radix*.

#### (4) zur Analyse und Interpretation der Daten

Promovierender PER: 60%. DNA sequence analysis, multiple sequence alignments. Final mitogenome assembly and annotation. Phylogenetic reconstruction. Analysis of evolutionary rates. Positive selection analyses.

Co-Autor AMW: 20%. DNA sequence analysis, multiple sequence alignments. Initial mitogenome assembly

Co-Autor MP: 20%. Analysis of evolutionary rates. Interpretation of selection patterns. initial mitogenome assembly.

#### (5) zum Verfassen des Manuskripts

Promovierender PER: 70%

Co-Autor AMW: 10%

Co-Autor MP: 20%

Datum/Ort: 01.12.2016, Frankfurt am Main.

---

Unterschrift Promovend: \_\_\_\_\_

#### Zustimmende Bestätigungen der oben genannten Angaben

Unterschrift Betreuer: \_\_\_\_\_ Datum/Ort: 01.12.2016, Frankfurt am Main.

Ggfs. Unterschrift *corresponding author* PER: \_\_\_\_\_ Datum/Ort: 01.12.2016, Frankfurt am Main.

RESEARCH ARTICLE

Open Access



# Positive selection on panpulmonate mitogenomes provide new clues on adaptations to terrestrial life

Pedro E. Romero<sup>1,2,3\*</sup>, Alexander M. Weigand<sup>4,5</sup> and Markus Pfenninger<sup>1,2</sup>

## Abstract

**Background:** Transitions from marine to intertidal and terrestrial habitats resulted in a significant adaptive radiation within the Panpulmonata (Gastropoda: Heterobranchia). This clade comprises several groups that invaded the land realm independently and in different time periods, e.g., Ellobioidea, Systellomatophora, and Stylommatophora. Thus, mitochondrial genomes of panpulmonate gastropods are promising to screen for adaptive molecular signatures related to land invasions.

**Results:** We obtained three complete mitochondrial genomes of terrestrial panpulmonates, i.e., the ellobiid *Carychium tridentatum*, and the stylommatophorans *Arion rufus* and *Helicella itala*. Our dataset consisted of 50 mitogenomes comprising almost all major panpulmonate lineages. The phylogenetic tree based on mitochondrial genes supports the monophyly of the clade Panpulmonata. Terrestrial lineages were sampled from Ellobioidea (1 sp.) and Stylommatophora (9 spp.). The branch-site test of positive selection detected significant non-synonymous changes in the terrestrial branches leading to *Carychium* (Ellobioidea) and Stylommatophora. These convergent changes occurred in the *cob* and *nad5* genes (OXPHOS complex III and I, respectively).

**Conclusions:** The convergence of the non-synonymous changes in *cob* and *nad5* suggest possible ancient episodes of positive selection related to adaptations to non-marine habitats. The positively selected sites in our data are in agreement with previous results in vertebrates suggesting a general pattern of adaptation to the new metabolic requirements. The demand for energy due to the colonization of land (for example, to move and sustain the body mass in the new habitat) and the necessity to tolerate new conditions of abiotic stress may have changed the physiological constraints in the early terrestrial panpulmonates and triggered adaptations at the mitochondrial level.

**Keywords:** Codon models, Land invasion, Mitogenomics, Panpulmonata, Positive selection

## Background

The transition from water to land is a fascinating evolutionary issue. The realm change occurred several times and across different taxa, from microorganisms to lichens and green plants, and later, arthropods, mollusks, annelids and vertebrates [1]. The multiple transitions required modifications on several systems and organs previously

adapted to aquatic habitats. For example, the presence of internal gas exchangers (lungs) to uptake oxygen, or different skin modifications as cuticle and keratin layers to decrease evaporation rates. Other examples include the production of novel compounds (e.g., uric acid and urea) to excrete nitrogen, and the presence of a skeleton and thick body muscles to support the body mass [2].

The vast majority of mollusks are marine, but several lineages within the clade Gastropoda have conquered freshwater and terrestrial habitats [3], e.g., Neritimorpha, Cyclophoroidea, Littorinoidea, Rissooidea, and Panpulmonata [4]. Most terrestrial mollusks belong to the clade Panpulmonata (Gastropoda: Euthyneura) [5]. The invasion of the non-marine realms in Panpulmonata

\* Correspondence: pedro.romero@senckenberg.de

<sup>1</sup>Senckenberg Biodiversity and Climate Research Centre (BiK-F), Senckenberganlage 25, 60325 Frankfurt am Main, Germany

<sup>2</sup>Institute for Ecology, Evolution & Diversity, Faculty of Biological Sciences, Goethe University Frankfurt, Max-von-Laue-Straße 13, 60438 Frankfurt am Main, Germany

Full list of author information is available at the end of the article



© 2016 The Author(s). **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

occurred multiple times independently in the clades Ellobioidea, Hygrophila, Stylommatophora and Systellommatophora [6–8]. For this reason, panpulmonates are a suitable model to study parallel adaptations to new environments, in particular to land.

Examples from invertebrates like the Collembola (Arthropoda: Hexapoda) showed that adaptive changes during terrestrialization have probably occurred in genes related to ion transport, homeostasis, immune response and development [9]. On the other hand, several examples from vertebrates showed various molecular mechanisms of adaptation: duplication and functional diversification of keratin genes [10], expansion of genes encoding olfactory receptors to detect airborne ligands [11], and positive selection on either nuclear genes involved in the urea cycle [12], or mitochondrial genes responding to the increase of oxygen during the Devonian [13]. However, the genomic basis of the transition to the land in panpulmonates is still unknown.

The animal mitochondrial genome encodes 13 proteins that are involved in the production of almost all energy in the eukaryotic cells. These proteins belong to four of the five complexes of the oxidative phosphorylation (OXPHOS) pathway and are under high functional constraints [14]. For example, inefficiencies caused by amino acid changes in these proteins alter the OXPHOS performance, and produce reactive oxygen species (ROS), i.e., molecules that lead to cellular damage and metabolism disruption [15]. In addition, amino acid substitutions have been shown to improve aerobic capacity and adaptation to new environments [16].

Several studies have reported non-neutral changes in each of the 13 mitochondrial genes [17]. Cytochrome c oxidase genes *cox1* and *cox3* from the freshwater fish *Poecilia* spp. present substitutions involved in the adaptation to toxic (H<sub>2</sub>S-rich) environments. These substitutions trigger conformational changes that block the uptake of H<sub>2</sub>S [18]. In addition, repeated selection at the same structural amino acid location has been found in the mitochondrial complex I from other fish, rodents and snakes [19]. The changes likely impacted the biomechanical apparatus that affect the electrochemical gradient in the mitochondria [17]. Moreover, cytochrome b (*cob*) in whales demonstrates several signatures of positive selection, in comparison to other artiodactyls. The adaptive changes in *cob* have been related to changes in the demand of metabolic processes during cetacean cladogenesis and the transition from land to water habitats [20].

Mitochondrial genomes of euthyneuran gastropods represent a promising dataset to screen for adaptive signatures related to water-to-land transitions. As mentioned above, this clade contains terrestrial and intertidal panpulmonates (e.g., Stylommatophora, Systellommatophora,

and Ellobioidea), and also freshwater taxa (Hygrophila) as well as other marine clades (e.g., Euopisthobranchia, Nudipleura). Here, we sequenced and annotated three new panpulmonate mitogenomes from the terrestrial ellobiid *Carychium tridentatum* (Risso, 1826) and the stylommatophorans *Arion rufus* (Linnaeus, 1758) and *Helicella itala* (Linnaeus, 1758). We used these new mitogenomes in addition to 47 already published euthyneuran mitogenomes, to reconstruct the phylogenetic relationships of Euthyneura. Finally, we evaluate the magnitude of selective pressures that occurred on the branches leading to terrestrial taxa.

## Results and discussion

### Characteristics of the new panpulmonate mitogenomes

The length of the three new mitochondrial genomes from the terrestrial panpulmonates *Carychium tridentatum* (Ellobiidae, Ellobioidea), *Arion rufus* and *Helicella itala* (Arionidae and Hygromiidae, Stylommatophora) are 13908 bp, 14321 bp, and 13966 bp, respectively. The three mitogenomes all encode for 13 protein-coding genes (PCG), 22 tRNAs, and 2 rRNAs, as reported for most other animal mitogenomes [21]. A detailed overview of the gene annotations can be found in Table 1. Nine genes are encoded on the major strand: *cox1*, *cox2*, *cob*, *nad1*, *nad2*, *nad4*, *nad4L*, *nad5* and *nad6*; while four are encoded in the minor strand: *atp6*, *atp8*, *cox3* and *nad3*. The gene arrangement is similar to other panpulmonates [21, 22]. Basically, the coding-genes are organized as follows: *cox1–nad6–nad5–nad1–nad4L–cob–cox2–atp8–atp6–nad3–nad4–cox3–nad2*. The cluster *cox2–atp8–atp6* is conserved among other gastropods and cephalopods [20]. However, in *A. rufus*, we found the small rRNA subunit between *cox2* and *atp8* (*cox2–rrnS–atp8–atp6*). Furthermore, clusters *trnD–trnC–trnF* and *trnY–trnW–trnG–trnH–trnQ–trnL2* are typical in Ellobioidea and Systellommatophora [22]. We found these clusters in the ellobiid *C. tridentatum*, but also in the stylommatophorans *H. itala*, and *A. rufus*, the latter with a slight modification (*trnW–trnY*).

The total length of the PCG is 10923 bp in *C. tridentatum*, 10935 bp in *A. rufus*, and 11071 bp *H. itala*. The GC-content of the PCG is approximately 30 %, being slightly higher in *H. itala* (34 %). PCG start with five different initiation codons: ATA, ATG, ATT, GTG, TTG. Finally, an AT-rich intergenic spacer between *cox3* and *trnI* has been proposed as the potential origin of replication (POR) in other euthyneurans [22, 23]. We found the same intergenic region in each of the three new mitogenomes. The AT-mean value in the potential POR region was 83 %.

### Phylogenetic analyses

Our reconstructed tree is congruent with previous comprehensive phylogenetic analyses in Euthyneura, using a

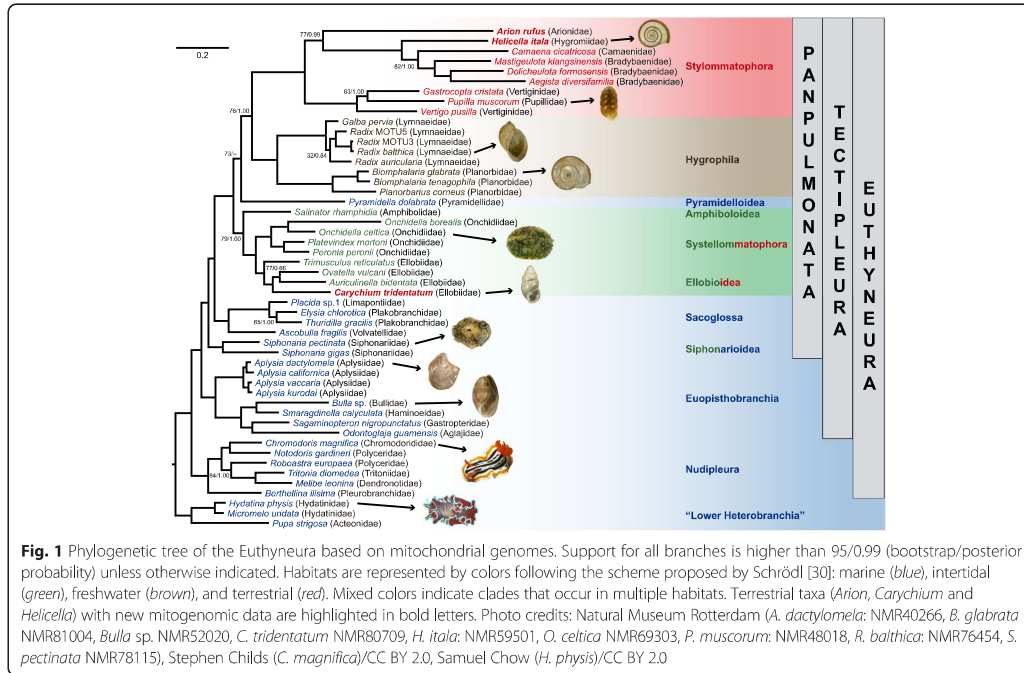
**Table 1** Gene features in the three new panpulmonate mitogenomes

Gene	Coding strand	Ellobioidea				Stylommatophora							
		<i>Carychium tridentatum</i>				<i>Arion rufus</i>				<i>Helicella itala</i>			
		From	To	Start codon	Stop codon	From	To	Start codon	Stop codon	From	To	Start codon	Stop codon
<i>cox1</i>	+	1	1536	TTG	TAA	1	1530	TTG	TAA	1	1530	ATG	TAA
<i>trnV</i>	+	1529	1591	-	-	1538	1600	-	-	1527	1587	-	-
<i>rnl</i>	+	1679	2645	-	-	1690	2578	-	-	1699	2583	-	-
<i>trnL1</i>	+	2588	2653	-	-	2592	2654	-	-	2595	2655	-	-
<i>trnP</i>	+	<b>2712</b>	<b>2773</b>	-	-	2657	2721	-	-	2656	2712	-	-
<i>trnA</i>	+	<b>2650</b>	<b>2712</b>	-	-	2739	2803	-	-	2714	2775	-	-
<i>nad6</i>	+	2778	3233	ATT	TAA	2838	3278	ATA	TAG	2776	3249	TTG	TAA
<i>nad5</i>	+	3258	4889	ATT	TAG	3238	4920	ATA	TAA	3239	4915	TTG	TAA
<i>nad1</i>	+	4867	5775	TTG	TAA	4911	5813	ATG	TAG	4867	5796	ATT	TAA
<i>nad4l</i>	+	5776	6114	GTG	TAA	5822	6124	ATA	TAG	5787	6098	ATT	TAG
<i>cob</i>	+	6056	7177	ATT	TAG	6103	7182	TTG	TAA	6113	7187	TTG	T
<i>trnD</i>	+	7175	7224	-	-	7184	7248	-	-	7188	7242	-	-
<i>trnC</i>	+	7228	7290	-	-	7284	7342	-	-	7243	7303	-	-
<i>trnF</i>	+	7291	7354	-	-	7346	7409	-	-	7307	7368	-	-
<i>cox2</i>	+	7340	8053	TTG	TAA	7412	8080	ATG	TAG	7369	8067	GTG	TAG
<i>trnY</i>	+	8028	8079	-	-	<b>8207</b>	<b>8270</b>	-	-	8042	8101	-	-
<i>trnW</i>	+	8080	8140	-	-	<b>8082</b>	<b>8149</b>	-	-	8097	8157	-	-
<i>trnG</i>	+	8150	8197	-	-	8465	8527	-	-	8158	8218	-	-
<i>trnH</i>	+	8199	8260	-	-	8522	8591	-	-	8215	8274	-	-
<i>trnQ</i>	-	8292	8354	-	-	<b>8672</b>	<b>8736</b>	-	-	8275	8332	-	-
<i>trnL2</i>	-	8355	8419	-	-	<b>9557</b>	<b>9624</b>	-	-	8333	8388	-	-
<i>atp8</i>	-	8394	8570	TTG	TAG	<b>9578</b>	<b>9775</b>	ATG	TAG	8360	8614	ATG	TAA
<i>trnN</i>	-	8571	8637	-	-	<b>9781</b>	<b>9843</b>	-	-	8615	8677	-	-
<i>atp6</i>	-	8638	9282	ATG	TAA	<b>9836</b>	<b>10489</b>	TTG	TAA	8679	9332	ATA	TAA
<i>trnR</i>	-	9283	9343	-	-	<b>10490</b>	<b>10555</b>	-	-	9330	9389	-	-
<i>trnE</i>	-	9344	9408	-	-	<b>8589</b>	<b>8655</b>	-	-	9390	9449	-	-
<i>rnlS</i>	-	9408	10086	-	-	<b>8822</b>	<b>9477</b>	-	-	9449	10126	-	-
<i>trnM</i>	+	10108	10176	-	-	<b>9491</b>	<b>9553</b>	-	-	10145	10210	-	-
<i>nad3</i>	-	10172	10528	ATA	TAA	10557	10904	ATA	TAA	10188	10556	ATG	TAA
<i>trnS2</i>	-	10540	10595	-	-	10914	10979	-	-	10554	10605	-	-
<i>trnS1</i>	+	10598	10653	-	-	10993	11053	-	-	10683	10734	-	-
<i>nad4</i>	+	10653	11960	TTG	TAA	11041	12444	ATT	TAA	10723	12036	ATC	TAG
<i>trnT</i>	-	11983	12047	-	-	12391	12455	-	-	12040	12101	-	-
<i>cox3</i>	-	12038	12826	ATG	TAG	12436	13233	ATT	TAA	12082	12882	ATA	TAA
<i>trnI</i>	+	12871	12933	-	-	13280	13342	-	-	12921	12980	-	-
<i>nad2</i>	+	12933	13871	ATT	TAG	13307	14230	TTG	TAA	12969	13913	ATT	TAG
<i>trnK</i>	+	13860	2	-	-	14262	8	-	-	13914	7	-	-

Annotations were performed in the MITOS server using default parameters, and then manually refined in Geneious R7. +/- signs indicate the sense of each annotation. Gene rearrangements with respect to the other two mitogenomes are indicated in bold

combination of mitochondrial and nuclear genes [5], and phylogenomics [3, 24] (Fig. 1). Tectipleura (Euopisthobranchia + Panpulmonata) [25] are highly supported in

both ML and BI analyses. The clade Nudipleura is monophyletic and within this clade, Pleurobranchidae (*Berthellina*) is the sister group of Nudibranchia, as



proposed by Göbbeler et al. [26]. In addition, there is high support for the clade Euopisthobranchia. This clade was defined by Jörger et al. [5] reuniting the clades Umbraculoidea, Anaspidea, Runcinacea, Pteropoda and Cephalaspidea. In our topology, Anaspidea (*Aplysia* spp.) and the cephalaspideans *Bulla*, *Odontoglossa*, *Sagaminopteron* and *Smaragdina* conform a monophyletic group.

Previous mitochondrial phylogenetic reconstructions recovered the monophyly of the former accepted clade “Opisthobranchia” and the paraphyly of “Pulmonata” [22, 23, 27]. However, the topologies derived from mitogenomics have received criticism, for long-branch attraction (LBA) artifacts affecting the topologies in Heterobranchia. In these cases, long-branched stylommatophorans were recovered closer to the root of the clade while they appeared as derived in the nuclear topologies [28]. On the other hand, recent genomic evidence rejected “Opisthobranchia” in favor of Euopisthobranchia as the sister group of Panpulmonata. Phylogenetic reconstructions based on concatenated nuclear and mitochondrial genes [5, 7, 29] as well new phylogenomic studies [3, 24] recovered the paraphyly of “Opisthobranchia”, and support for Panpulmonata.

We were aware of these rooting issues; thus, we choose members of the “Lower Heterobranchia” as outgroup taxa. Our topology recovered monophyletic Panpulmonata and

Euopisthobranchia as its sister group. The clade Panpulmonata, defined by Jörger et al. [5], comprises the clades Amphiboloidea, Ellobioidea, Glacidorboidea, Hygrophila, Siphonarioidea, Stylommatophora, and Systellomatophora plus Acochlidia and Sacoglossa, previously regarded as opisthobranchs [30]. Therefore, Panpulmonata possesses an extraordinary diversity in morphology (snails, slugs and intermediate forms), and habitats (marine, intertidal, freshwater and terrestrial).

Recently, the monophyly of “Pulmonata” has been challenged [30], i.e., evidence from phylogenomics did not recover “Pulmonata” as a monophyletic group [3]. In our tree, members from Amphibolidae (*Salinator*) and Pyramidelloidea (*Pyramidella*) appear between traditional “Pulmonata” clades, favoring Panpulmonata over “Pulmonata”. Our topology supports the Amphipulmonata clade (Ellobioidea + Systellomatophora) [29], and rejects the Geophila hypothesis (Stylommatophora + Systellomatophora) [30]. Finally, the association between Stylommatophora and Hygrophila has been also found using phylogenomic analyses [24], although with a small subset of euthyneuran taxa.

**Patterns of evolutionary rates**

The relative evolutionary rates (RER) for amino acids were not equally distributed over the alignment

(Additional file 1). The mean RER value in the *nad* genes was 2–3 times higher than in the *cox* genes. The RER for the *cox1* gene were below the mean rates in our dataset, indicating a higher number of conserved sites.

Stylommatophora ( $\gamma = 0.6514x$ ;  $R^2 = 0.9641$ ) along with Hygrophila ( $\gamma = 0.6708x$ ;  $R^2 = 0.9770$ ) presented the highest divergence slopes in our data (Additional file 2). This means that fewer nucleotide changes produced more amino acid changes in comparison to the other clades, i. e. non-synonymous changes are more frequent in both Stylommatophora and Hygrophila. Furthermore, Stylommatophora presented the highest absolute values for both nucleotide and amino acid divergence. This result explains the presence of long branches in this clade (Fig. 1).

Several hypotheses have been proposed to explain the extreme divergence found in the mitochondrial DNA of land snails (Stylommatophora) [31]: (1) exceptionally accelerated rate of evolution, (2) haplotype groups previously differentiated in isolated refuges getting into secondary contact, (3) natural (positive) selection preserving the variation, and (4) the particular population structure in pulmonates that allowed them to preserve ancient haplotypes. Accelerated evolution of the mtDNA has been found in several species of land snails and slugs (hypothesis 1). The evolutionary rate in the mtDNA (rRNA) of the snails *Euhadra* and *Mandarina* was 10 % per Ma [32, 33], and 5.2 % in the slug *Arion* [34]. However, this is not a general pattern for all pulmonates, for example the evolutionary rate of *Albinaria* and *Partula* was estimated to be 1–1.2 and 2.8 % per Ma, respectively [35]. The secondary contact after allopatric divergence of haplotypes (hypothesis 2) has been found in *Candidula* [36] and *Cepaea* [37]. Moreover, introgression of mitochondrial lineages as a result of hybridization has been observed in two species of the land snail genus *Trochulus* [38]. The effect of natural selection (hypothesis 3) shaping the genetic diversity has been proposed before for land

snails [39] although it was considered to be uncommon [40]. However, a recent study has shown that mitochondrial DNA undergoes substantial amounts of adaptive evolution, especially in mollusks [41]. The particular demographic pattern of land snails that produces highly structured populations (hypothesis 4), i.e., “islands” of isolated demes, affects the probability of reciprocal monophyly of two samples and the chance that a gene tree matches the species tree [42], and explained the persistence of ancestral polymorphisms and the extreme divergence in *Achatinella* [43], *Systrophia* [44], and *Xerocrassa* [45]. In addition, in the case of Hygrophila, some studies found high divergence rates in *Physella* [46], and *Radix* [47], although no clear hypothesis has been proposed to explain this pattern.

#### Analyses of selective pressures

Codon substitution models have been widely used to detect adaptive signatures affecting protein evolution [48]. First, we tested for the presence of positive selected codons across the alignments. All comparisons (M2a-M1a, M8-M7, M8-M8a) consistently favored positive selection models M2a and M8 in *cox1*, *cox2*, *cox3*, and *cob* ( $p < 0.05$ ) (Table 2). However, the proportion of sites with  $\omega > 1$  was extremely low (Additional files 3 and 4). High  $\omega$ -values in the positively selected genes can be explained by the presence of few synonymous sites affecting dN estimations. On the contrary, a context of strong negative selection could explain  $\omega < 1$  for most of the genes [49, 50].

Positive selection tests based on either sites or branches only, are conservative for many genes [51]. This is because the test is only significant if the average  $\omega > 1$  holds true for all sites or all branches. However, one might expect that positive selection affects only specific sites in specific branches or lineages [52]. For these reasons, we used the branch-site test of positive

**Table 2** Site test of positive selection

	M1a (np = 3)	M2a (np = 5)	M7 (np = 3)	M8 (np = 5)	M8a (np = 4)	M2a/M1a (df = 2)	M8/M7 (df = 2)	M8/M8a (df = 1)
ATP6	-23368.8643	-23368.7771	-22291.9994	-22292.0038	-22292.0038	0.1744	-0.008956	0.0000
COX1	-30899.0024	-30854.9357	-29758.9356	-29704.0913	-29746.9794	<b>88.1334</b>	<b>109.6887</b>	<b>85.7763</b>
COX2	-18042.2493	-18015.6742	-17256.1223	-17221.3147	-17245.3928	<b>53.1501</b>	<b>69.6151</b>	<b>48.1561</b>
COX3	-21196.1015	-21158.9525	-20023.0527	-20010.8195	-20023.0034	<b>74.2980</b>	<b>24.4664</b>	<b>24.3678</b>
CYTb	-30416.7613	-30401.6179	-28984.2418	-28973.9145	-28984.1611	<b>30.2867</b>	<b>20.6547</b>	<b>20.4933</b>
ND1	-28047.2931	-28045.7290	-26771.4013	-26771.4065	-26771.4065	3.1282	-0.010268	0.0000
ND2	-36553.7114	-36553.7117	-35518.2839	-35518.2961	-35518.2961	-0.0006	-0.024316	0.0000
ND4	-47287.5457	-47287.5457	-45097.7072	-45097.4722	-45097.4722	0.0000	0.46994	0.0000
ND5	-58590.6941	-58589.0944	-55650.7195	-55650.7278	-55650.7278	3.1994	-0.016656	0.0000

The comparisons within the site models were M8 vs. M8a/M7, and M2a vs. M1a. Values in bold represent highly significant differences ( $p < 0.01$ ) from the null model, np: number of parameters  
df degrees of freedom



selection [51] focusing on the branches leading to terrestrial taxa (foreground) within Panpulmonata. The alternative model (model A) fitted significantly better than the null model (model A1) in the genes *cox1*, *cox2*, *cob* and *nad5* (Additional file 5). From these four genes, only *cob* and *nad5* presented an  $\omega$ -ratio higher than one and positively selected codons in the foreground (two sites in *cob* and six in *nad5*) (Table 3).

The branch-site model has been shown to detect ancient episodes of positive selection [53]. A potential problem of the test can be the saturation over long evolutionary times; however, simulations have shown that extreme sequence divergence does not generate false positives although it can lead to a high rate of false negatives, especially in older nodes [54]. In the same study, significant levels of positive selection ( $\omega = 6$  or 12) were detected at the radiation of bony vertebrates (Euteleostomi), approximately 400–500 Ma [55]. In our data, divergence times are lower than in the vertebrate study: Euthyneura and Panpulmonata probably diverged from their sister groups 250–350 Ma and 150–250 Ma ago, respectively; while panpulmonate clades with terrestrial taxa diverged more recently (Ellobioidea: 140–160 Ma; Stylommatophora: 100–150 Ma) [3, 5].

#### Convergent adaptations related to realm shifts

The evolution of the lung in early panpulmonates, probably originating from the pallial cavity of an intertidal gilled-ancestor, was the key evolutionary innovation that allowed the diversification to non-marine realms [24]. Both Ellobioidea and Stylommatophora possess lungs, although they colonized the land in different times. Land invasions in Ellobioidea occurred at least twice, one within the genus *Pythia* (15–25 Ma) and the other in the subfamily Carychiinae (50–100 Ma) [8]; while terrestrialization in Stylommatophora appeared to be older (100–150 Ma) [5].

Different genes appeared to be under positive selection ( $\omega > 1$ ) in terrestrial panpulmonate branches. While *cob* and *nad5* are both part of the OXPHOS pathway, they belong to different complexes (complex III and I, respectively), suggesting that adaptations occurred in several molecular targets. The observed non-synonymous

mutations produced similar changes in the amino acid properties, albeit in different regions of each gene. Results from TreeSAAP for both *cob* and *nad5* showed that these changes alter the equilibrium constant (ionization of COOH) property (Fig. 2). This property may influence the protein efficiency reducing ROS production while increasing individual longevity [56]. Alterations in the equilibrium constant may have allowed organisms to better cope with abiotic stress conditions in the new hot, cold or dry habitats. For example, desiccation tolerance has been shown as a limiting factor for the invasion of dry habitats in thiarid freshwater snails [57]. Moreover, the activation of the antioxidant metabolism reducing ROS excess has been linked to desiccation tolerance in the algae *Mastocarpus stellatus* and *Porphyra columbina* occurring in the upper intertidal zone [58]. Since desiccation stress (or abiotic stress in general) is linked to metabolic activity and ROS production [59], this directly affects the invasion success of an evolutionary lineage.

The increase in metabolic efficiency has been also related with the terrestrial invasion in other animals, e. g., tetrapods, during the Devonian. Amphibians, lungfishes, and coelacanths presented significant changes in the same equilibrium constant (ionization of COOH) property suggesting an adaptation to increased oxygen levels and changing metabolic requirements [13]. These mutations affected both *cob* and *nad5* tetrapod genes. Also, it is noteworthy to mention that in the terrestrial panpulmonate *nad5* gene, the different approaches used by PAML and TreeSAAP found signatures of positive selection along with amino acid property changes in similar regions (Sites 308 and 512; Fig. 2, Table 3).

Changes at the molecular level could have also occurred in different taxa, so we compared the sites under positive selection in our data against previous studies to find sites with similar adaptive patterns (Tables 4 and 5). For *cob*, site 151 is located in an intermembrane domain. This site is homologous to site 158 in a previous cetacean-artiodactyl alignment (Table 4) and was revealed to be under selective pressure in cetaceans, also influencing the equilibrium constant of the cd2 intermembrane helix [20]. The authors proposed that non-

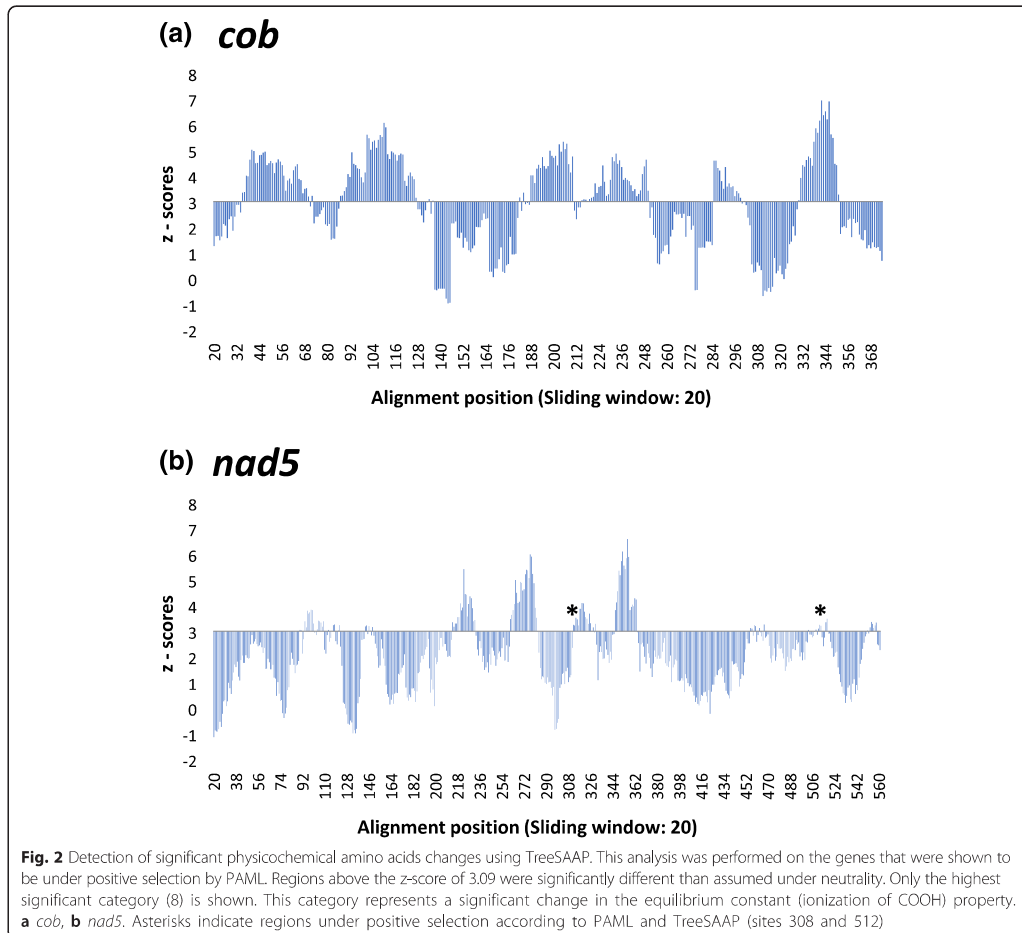
**Table 3** Branch-site test of positive selection on the mitochondrial genes

Gene	Proportion of site classes under model A				dN / dS ( $\omega$ ) in the foreground (terrestrial taxa)				Positively selected sites BEB (pp > 0.95)
	0	1	2a	2b	0	1	2a	2b	
<i>cox1</i>	88.37	9.98	1.49	0.17	0.0073	1.0000	1.0000	1.0000	-
<i>cox2</i>	79.08	14.54	5.39	0.99	0.0278	1.0000	1.0000	1.0000	188
<i>cob</i>	85.13	11.90	2.60	0.36	0.0243	1.0000	<b>2.4988</b>	<b>2.4989</b>	006, 151
<i>nad5</i>	71.79	21.13	5.48	1.61	0.0559	1.0000	<b>2.1521</b>	<b>2.1521</b>	169, 308, 474, 478, 479, 512

Only the genes that showed a significant difference from the null model are shown. Values in bold represent a  $\omega > 1$

BEB Bayes Empirical Bayes algorithm, pp posterior probability





synonymous changes in *cob* are related to increasing metabolic demands during cladogenesis. Similarly, adaptive evolution in *cob* has been related to the increase of energy metabolism in response to the evolution of flight in bats [60].

In case of *nad5*, site 474 is homologous to site 540 from the alignment of Garvin et al. [17] (Table 5). This site is positively selected in rorquals (*Balaenopteridae*) and salmon (*Salmonidae*). Site 474 is part of the biomechanical apparatus that generates the electrochemical gradient and it has been shown to be under positive selection in the Pacific salmon species [19] and in eutherians [14]. Also, site 474 corresponds to site 519 in subterranean rodents [49]. This codon position is positively selected only in lineages that independently colonized the subterranean niche, a habitat suggested to be

energetically demanding [61]. Mutations in the NADH complex, especially in the transmembrane domains, may affect the proton pump activity of this complex [14]. These changes could facilitate the proton flow and improve the efficiency of ATP production - characteristics associated to increased energetic requirements in non-marine habitats [62].

### Conclusions

We represent evidence of positive selection on several amino acid positions in the mitochondrial complexes I (*nad5*) and III (*cob*). These episodes of positive selection occurred in independent branches of panpulmonates with terrestrial taxa (*Ellobioidea* and *Stylommatophora*), indicating their possible role during the invasion of the land realm. Most of these sites have been shown to be

**Table 4** Alignment section of the *cob* gene from Garvin et al. [17]

	Position in the alignment from Garvin et al. [17]					
	156	157	158	159	160	161
<i>Anguilla</i>	V	G	D	T	L	V
Argentinoidei	V	G	E	A	L	V
Baleanopteridae	I	G	N	T	L	V
Caprinae	I	G	T	N	L	V
Delphinidae	I	G	T	T	L	V
<i>Hypsiglena</i>	L	G	T	S	L	T
<i>Oryzias</i>	V	G	N	A	L	V
Otariidae	I	G	A	N	L	V
Phocidae	I	G	T	D	L	V
Primates	I	G	T	D	L	V
Salmonidae	V	G	N	A	L	V
<i>Takifugu</i>	V	G	N	T	L	V
Ursidae	I	G	T	D	L	V
<i>Bos taurus</i>	I	G	T	N	L	V
McClellan et al. [20]			158			
da Fonseca et al. [14]			158			
This work			151			

The corresponding homologous position in *Bos taurus* is shown. Homologous positions found under positive selection in other studies are shown below the alignment

also under positive selection in several other taxa. Moreover, the general pattern suggests that non-synonymous mutations in both genes are probably linked to increased or altered metabolic requirements. An increased demand for energy due to the colonization of land and the necessity to cope with different abiotic stress conditions may have changed the physiological constraints in the early terrestrial panpulmonates and triggered functional adaptations at the mitochondrial level. Future studies can take into account the predicted codons and the information on the physicochemical changes to test whether these mutations also affect protein structure and function. New genomic information from panpulmonates will most likely reveal even more genes involved in metabolic and structural processes that were key to the colonization of the terrestrial realm.

## Methods

### Mitogenome sequencing

Our dataset comprises 50 complete mitochondrial genomes, and is representative of all described lineages within Euthyneura, except Acochlidia where no mitogenomes have been sequenced so far. We did not consider identical mitogenomes and mitogenomes that were not verified for biological accuracy by GenBank. Accession numbers for each sample are provided in the Additional file 6.

**Table 5** Alignment section of the *nad5* gene from Garvin et al. [17]

	Position in the alignment from Garvin et al. [17]					
	540	541	542	543	544	545
<i>Anguilla</i>	T	H	N	F	S	N
Argentinoidei	L	H	N	F	S	N
Baleanopteridae	<b>F</b>	S	K	F	S	T
Caprinae	T	F	K	F	S	N
Delphinidae	S	T	K	F	S	T
<i>Hypsiglena</i>	L	N	L	F	F	N
<i>Oryzias</i>	<b>T</b>	H	H	F	S	N
Otariidae	L	F	K	F	S	N
Phocidae	L	F	K	F	S	S
Hominidae	T	F	R	F	S	N
Salmonidae	T	H	N	F	S	N
<i>Takifugu</i>	P	H	H	F	S	N
Ursidae	P	F	K	F	S	N
<i>Escherichia coli</i>	NuoL527	NuoL528	NuoL529	NuoL530	NuoL531	NuoL532
Garvin et al. [19]	526					
Tomasco and Lessa [49]	519	520				
This work	474				478	479

Positions under positive selection according to Garvin et al. [17] are highlighted in bold. The corresponding homologous position in *Escherichia coli* is shown. Homologous positions found under positive selection in other studies are shown below the alignment

In addition, we used DNA previously isolated from specimens of *Arion rufus* (Linnaeus, 1758) [63], *Carychium tridentatum* (Risso, 1826) [64], and *Helicella itala* (Linnaeus, 1758) [65] for DNA shotgun sequencing. We followed the protocol described by Feldmeyer et al. [66] with some variations: 500 ng DNA per sample was used for library preparation, following the Roche GS FLX Titanium General Library Preparation specifications. Each sample was sequenced on 1/8 of a titanium plate on a 454 sequencer. It should be noted, that approximately 100 specimens of the microgastropod *C. tridentatum* originating from a single locality had to be pooled to obtain 500 ng of DNA.

#### Mitogenome assembly and annotation

Newbler v2.0.1 (Roche) was used for contig assembly, with standard settings. Then, we subjected contigs to BlastN and BlastX searches against the mitochondrial genomes of closely related taxa. In addition, to close the remaining gaps after the assembly, we designed flanking primers using Geneious R7 [67]. Primer sequences can be found in Additional file 7. Sequence amplification was conducted using the following PCR conditions: Each 25  $\mu$ L PCR mix included 1  $\mu$ L (10 pmol) of each primer, 2.5  $\mu$ L 10x PCR buffer, 2  $\mu$ L (100 mM) MgCl<sub>2</sub>, 0.2  $\mu$ L (20 mM) dNTPs, 0.3  $\mu$ L Taq-polymerase (Fermentas), 1.5  $\mu$ L (10 mg/mL) bovine serum albumin, 12.50  $\mu$ L ddH<sub>2</sub>O and 4  $\mu$ L template DNA in variable concentrations. Temperature conditions: 1 min at 95 °C, followed by 30 cycles of 30 s at 95 °C, 30 s at 52 °C and 30 s at 72 °C, and finally, 3 min at 72 °C. Visualization of PCR products was performed on a 1.4 % agarose gel. Amplicons were cleaned using the QIAquick PCR Purification Kit or the QIAquick Gel Extraction kit (Qiagen) whenever multiple bands were detected. Sanger sequencing was performed using the PCR primer pair (5 pmol) and the BigDye<sup>®</sup> Terminator v.3.1 Cycle Sequencing Kit (LifeTechnologies, Inc.) on an ABI 3730 capillary sequencer, using the facilities of the Senckenberg BiK-F Laboratory Centre, Frankfurt am Main.

Both, shotgun contigs and Sanger sequences were aligned in Geneious R7 to obtain the complete mitochondrial genomes. The complete mitogenome assemblies were annotated using the MITOS webserver [68]. This program also identified rRNA and tRNA genes. Additionally, we compared the results from MITOS to other annotation strategies like NCBI ORF Finder or Geneious R7, with similar results. Finally, we compared the new gene annotations against other panpulmonate mitochondrial genes to evaluate the length of the reading frames. Newly determined genomes were deposited into GenBank (accession numbers: KT626607, KT696545, KT696546).

#### Sequence alignments

The 13 protein-coding genes (PCG) were translated into amino acids in Geneious R7 using the invertebrate mitochondrial genetic code, and then aligned using the MAFFT [69]. Ambiguous aligned regions were removed using Gblocks [70]. For downstream analyses, we did not use alignments that had a length below 150 aa (~450 nt) after Gblocks trimming. Thus, nine genes were selected: *atp6*, *cox1*, *cox2*, *cox3*, *cob*, *nad1*, *nad2*, *nad4*, *nad5*. Then, nucleotide sequences of these genes were aligned using TranslatorX [71]. This software aligns the nucleotides by codons taking into account the information from the amino acid alignment. Gblocks was used again in the codon-based alignment with less stringent parameters to trim flanking regions and long gaps. The concatenated alignment length is 9711 nt.

#### Phylogenetic reconstructions

It has been shown that outgroup selection is important to conceal current hypothesis of euthyneuran phylogeny [21, 72]. Thus, we decided to follow Wägele et al. [28] choosing the “Lower Heterobranchia” clade (*Hydatina*, *Micromelo* and *Pupa*) as the outgroup. For tree reconstruction, we used only the first and second positions of the alignment in order to reduce saturation levels. The alignment length after removing the third position was 6474 nt. Data were partitioned by gene using the partition scheme suggested in PartitionFinder [73]. Maximum likelihood analyses were conducted in RAxML-HPC2 (8.0.9) [74, 75] implemented on XSEDE [76] (CIPRES Science Gateway). We followed the “hard and slow way” suggestions indicated in the manual and selected the best-likelihood tree after 1000 independent runs. Then, branch support was evaluated using bootstrapping with 1000 replicates, and confidence values were drawn in the best-scoring tree. Bayesian inference was conducted in MrBayes v3.2.2 [77] on XSEDE (CIPRES). Two simultaneous Monte Carlo Markov Chains (MCMC) were run, with the following parameters: eight chains of 50 million generations each, sampling every 1000 generations and a burn-in of 25 %. Tracer 1.6 [78] was used to evaluate effective sample sizes (ESS > 200). We assume that a bootstrap value of >70 % [79] and a posterior probability of > 0.95 [80] are evidence of significant nodal support.

#### Patterns of evolutionary rates

Relative evolutionary rates were calculated in the software MEGA6 [81], using the nucleotide and amino acid information following the procedure described by Merker et al. [82]. The rates are scaled such that the average evolutionary rate across all sites is 1. Sites with a rate lower than 1 evolve slower than the average while sites with a rate higher than 1 evolve faster than the average. The relative rates were estimated

under the General Time Reversible (GTR) model (+I) for nucleotide sequences (complete concatenated alignment: 9711 nt.) and under the mtREV model (+I) for amino acid sequences. The relative rates were scaled in windows with a size of 30 for nucleotides and 10 for amino acids, using the R package zoo [83]. Finally, we compared the amino acid divergence relative to the nucleotide divergence among main clades. Pairwise nucleotide and amino acid distances were calculated under the previously described substitution models.

#### Analysis of selective pressures

The CODEML program from PAML v4.8a [84] was used to analyze positive selection in each mitochondrial gene. PAML estimates the omega ratio ( $\omega = dN/dS$ ) using maximum likelihood. The omega ratio compares non-synonymous (dN) against synonymous (dS) substitutions per site. Assuming neutrality,  $\omega$  -values are equal to one; however,  $\omega > 1$  is expected if the gene undergoes adapting molecular evolution. In the latter scenario, non-synonymous mutations offer fitness advantages to the individual and have higher fixation probabilities than synonymous mutations [85].

The maximum likelihood topology was set as the guide tree. We re-estimated branch lengths on the tree using codon model M0 (one-ratio) and used them as fixed when fitting the site and branch-site models. In the site models, the  $\omega$ -ratio is allowed to vary among codons in the alignment [85], while in the branch-site models, the test focuses on the so-called foreground branches [50]. Specifically, we tested branches leading to the air-breathing land snails and slugs (Stylommatophora) and to the terrestrial *Carychium* (Ellobioidea).

We evaluated site models using a likelihood-ratio test (LRT). First, we compared the selection model (M2a; model = 0, NSites = 2, fix\_omega = 0, omega = 5) against the nearly neutral model (M1a; model = 0, NSites = 1, fix\_omega = 0, omega = 1) to detect signatures of  $\omega > 1$ . Then, we used models that calculate  $\omega$  from a beta distribution. Model M8 (model = 0, NSites = 8, fix\_omega = 0, omega = 5, 10 equal class proportions plus one class with  $\omega > 1$ ) was compared against either model M7 (model = 0, NSites = 7, fix\_omega = 0, omega = 1, 10 equal class proportions) or model M8a (model = 0, NSites = 8, fix\_omega = 0, omega = 1, 10 equal class proportions plus one class with  $\omega = 1$ ).

For the branch-site test we compared the model A (model = 2, NSites = 2, fix\_omega = 0, omega = 5) against the null model A (model = 2, NSites = 2, fix\_omega = 1, omega = 1). The LRT was calculated as follows,  $2 * (\ln L_{H1} - \ln L_{H0})$ . The Bayes Empirical Bayes (BEB) algorithm implemented in CODEML was used to calculate posterior probabilities of positive selected sites.

Genes detected to be positively selected in the branch-site test were then analyzed in TreeSAAP [86]. This software identifies significant physicochemical changes comparing the distribution of observed changes inferred from a phylogenetic tree against the random distribution of changes under neutrality. The magnitude of the change is rated from 1 (most conservative) to 8 (most radical). A highly significant z-score calculated in TreeSAAP ( $z > 3.09, p < 0.01$ ) indicates more non-synonymous substitutions than assumed under the neutral model [87]. We followed the suggestions from George and Bleck [13] to increase the accuracy of the test and lower the rate of false positives. Thus, we considered only the most radical changes (category 8,  $p < 0.01$ ) as significant. Moreover, we focused only on 20 of the 31 amino acid properties available in the software.

#### Alignment comparisons with previous studies

Almost all previous work on mitochondrial molecular adaptation has been done on vertebrates. We used the amino acid alignments from a recent review by Garvin et al. [17] to evaluate whether the positive selected sites found in our study are also present in a broader taxonomic context and could present a biological function. Mitochondrial sequences of *cob* and *nad5* reported by Garvin et al. [17] and references therein [14, 19, 20, 49] were downloaded and aligned using the global homology algorithm g-insi in MAFFT [69].

#### Additional files

**Additional file 1:** Evolutionary rates for nucleotides (NT; blue line) and amino acids (AA; red line) in the euthyneuran mitogenomes. Rates are scaled such that the average evolutionary rate across all sites is 1 (red line). The x-axis shows amino acid positions in the final concatenated alignment. (PDF 481 kb)

**Additional file 2:** Amino acid divergence versus nucleotide divergence in mitochondrial genomes of euthyneuran gastropods. Clades are differentiated by colors and symbols as shown in the legend. (PDF 205 kb)

**Additional file 3:** Comparison of models M0 (one-ratio), M1a (nearly neutral), and M2a (positive selection), and. Values in bold represent highly significant differences ( $p < 0.01$ ) from the null model. LRT: Likelihood ratio test, df: degrees of freedom, lnL: log likelihood, np: number of parameters. (XLSX 14 kb)

**Additional file 4:** Comparison of models M7, M8 and M8a. Values in bold represent highly significant differences ( $p < 0.01$ ) from the null model. LRT: Likelihood ratio test, df: degrees of freedom, lnL: log likelihood, np: number of parameters. (XLSX 14 kb)

**Additional file 5:** Branch-site test of positive selection on the mitochondrial genes (Model A vs. Null model). Values in bold represent highly significant differences ( $p < 0.01$ ) from the null model. LRT: Likelihood ratio test, df: degrees of freedom, lnL: log likelihood, np: number of parameters. (XLSX 10 kb)

**Additional file 6:** GenBank accession numbers of the 50 sequences used in this study. (XLSX 12 kb)

**Additional file 7:** Primer sequences used to close the gaps and complete the mitochondrial genomes of *Arion* and *Carychium*. (XLSX 9 kb)

#### Acknowledgements

We would like to thank Dr. Barbara Feldmeyer for providing suggestions to the manuscript and Claudia Nesselhauf for assistance in the laboratory.

**Funding**

The project was supported by the German funding program "LOEWE – Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz" of the Hessian State Ministry of Higher Education, Research and the Arts. PER also received a scholarship from "CONCYTEC/CIENCIACTIVA: Programa de becas de doctorado en el extranjero del Gobierno del Perú" (291-2014-FONDECYT).

**Availability of data and materials**

Mitogenome sequences are available at GenBank, accession numbers KT626607, KT696545, KT696546. Sequence information, multiple sequence alignments, partition schemes, and phylogenetic trees supporting the conclusions of this article are available in the FigShare repository. <https://dx.doi.org/10.6084/m9.figshare.c.3291377> [88].

**Consent for publication**

Not applicable.

**Ethics approval and consent to participate**

Not applicable.

**Author's contributions**

PER carried out the phylogenetic and molecular evolution analyses, participated in the mitogenome assembly, conceived the study and wrote the manuscript. AMW carried out the mitogenome assemblies, participated in the sequence alignment, and helped to draft the manuscript. MP participated in the design and coordination of the study and helped to draft the manuscript. All authors read and approved the final manuscript.

**Authors' information**

PER is a doctoral student the Faculty of Biosciences - Goethe University Frankfurt am Main and a member of the Molecular Ecology at the Senckenberg Biodiversity and Climate Research Centre. AMW is a post-graduate scientist and project coordinator in the Aquatic Ecosystem Research, University of Duisburg-Essen and a member of the Centre for Water and Environmental Research (ZWU) Essen, University of Duisburg-Essen. MP is a professor at the Faculty of Biosciences - Goethe University Frankfurt am Main and the leader of the Molecular Ecology group at the Senckenberg Biodiversity and Climate Research Centre.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup>Senckenberg Biodiversity and Climate Research Centre (BiK-F), Senckenberganlage 25, 60325 Frankfurt am Main, Germany. <sup>2</sup>Institute for Ecology, Evolution & Diversity, Faculty of Biological Sciences, Goethe University Frankfurt, Max-von-Laue-Straße 13, 60438 Frankfurt am Main, Germany. <sup>3</sup>Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, Av. Arenales 1256, Apartado 14-0434, Lima 14, Peru. <sup>4</sup>Aquatic Ecosystem Research, University of Duisburg-Essen, Universitätsstraße 5, 45141 Essen, Germany. <sup>5</sup>Centre for Water and Environmental Research (ZWU), University of Duisburg-Essen, Universitätsstraße 2, 45117 Essen, Germany.

Received: 29 February 2016 Accepted: 8 August 2016

Published online: 22 August 2016

**References**

- Laurin M. How Vertebrates Left Water. Berkeley: University of California Press; 2010.
- Little C. The Terrestrial Invasion: An Ecophysiological Approach to the Origins of Land Animals. Cambridge: Cambridge University Press; 1990.
- Zapata F, Wilson NG, Howison M, Andrade SC, Jorger KM, Schrödl M, et al. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. *Proc Biol Sci.* 2014;281(1794):20141739.
- Kameda Y, Kato M. Terrestrial invasion of pomatiopsid gastropods in the heavy-snow region of the Japanese Archipelago. *BMC Evol Biol.* 2011;11:18.
- Jorger KM, Stöger I, Kano Y, Fukuda H, Kneibelsberger T, Schrödl M. On the origin of Acochlidia and other enigmatic euthyneuran gastropods, with implications for the systematics of Heterobranchia. *BMC Evol Biol.* 2010;10:323.
- Kano Y, Neusser TP, Fukumori H, Jorger KM, Schrödl M. Sea-slug invasion of the land. *Biol J Linn Soc.* 2015;116(2):253–9.
- Klussmann-Kolb A, Dinapoli A, Kuhn K, Streit B, Albrecht C. From sea to land and beyond—new insights into the evolution of euthyneuran Gastropoda (Mollusca). *BMC Evol Biol.* 2008;8:57.
- Romero PE, Pfenninger M, Kano Y, Klussmann-Kolb A. Molecular phylogeny of the Ellobiidae (Gastropoda: Panpulmonata) supports independent terrestrial invasions. *Mol Phylogenet Evol.* 2016;97:43–54.
- Faddeeva A, Studer RA, Kraaijeveld K, Sie D, Ylstra B, Marien J, et al. Collembolan transcriptomes highlight molecular evolution of hexapods and provide clues on the adaptation to terrestrial life. *PLoS One.* 2015;10(6), e0130600.
- Vandebergh W, Bossuyt F. Radiation and functional diversification of alpha keratins during early vertebrate evolution. *Mol Biol Evol.* 2012;29(3):995–1004.
- Nikaïdo M, Noguchi H, Nishihara H, Toyoda A, Suzuki Y, Kajitani R, et al. Coelacanth genomes reveal signatures for evolutionary transition from water to land. *Genome Res.* 2013;23(10):1740–8.
- Amemiya CT, Alföldi J, Lee AP, Fan S, Philippe H, Maccallum I, et al. The African coelacanth genome provides insights into tetrapod evolution. *Nature.* 2013;496(7445):311–6.
- George D, Blicek A. Rise of the earliest tetrapods: an early Devonian origin from marine environment. *PLoS One.* 2011;6(7), e22136.
- da Fonseca RR, Johnson WE, O'Brien SJ, Ramos MJ, Antunes A. The adaptive evolution of the mammalian mitochondrial genome. *BMC Genomics.* 2008; 9(1):119.
- Moreno-Loshuertos R, Acín-Perez R, Fernandez-Silva P, Movilla N, Perez-Martos A, de Rodríguez C, et al. Differences in reactive oxygen species production explain the phenotypes associated with common mouse mitochondrial DNA variants. *Nat Genet.* 2006;38(11):1261–8.
- Dalziel AC, Moyes CD, Fredriksson E, Loughheed SC. Molecular evolution of cytochrome c oxidase in high-performance fish (teleostei: Scombroidei). *J Mol Evol.* 2006;62(3):319–31.
- Garvin MR, Bielawski JP, Sazanov LA, Gharrett AJ. Review and meta-analysis of natural selection in mitochondrial complex I in metazoans. *J Zool Syst Evol Res.* 2015;53(1):1–17.
- Pfenninger M, Lerp H, Tobler M, Passow C, Kelley JL, Funke E, et al. Parallel evolution of cox genes in H2S-tolerant fish as key adaptation to a toxic environment. *Nat Commun.* 2014;5:3873.
- Garvin MR, Bielawski JP, Gharrett AJ. Positive Darwinian selection in the piston that powers proton pumps in complex I of the mitochondria of Pacific salmon. *PLoS One.* 2011;6(9), e24127.
- McClellan DA, Palfreyman EJ, Smith MJ, Moss JL, Christensen RG, Sailsbery JK. Physicochemical evolution and molecular adaptation of the cetacean and artiodactyl cytochrome b proteins. *Mol Biol Evol.* 2005;22(3):437–55.
- Stöger I, Schrödl M. Mitogenomics does not resolve deep molluscan relationships (yet?). *Mol Phylogenet Evol.* 2013;69(2):376–92.
- White TR, Conrad MM, Tseng R, Balayan S, Golding R, de Frias Martins AM, et al. Ten new complete mitochondrial genomes of pulmonates (Mollusca: Gastropoda) and their impact on phylogenetic relationships. *BMC Evol Biol.* 2011;11:295.
- Grande C, Templado J, Zardoya R. Evolution of gastropod mitochondrial genome arrangements. *BMC Evol Biol.* 2008;8:61.
- Kocot KM, Halanych KM, Krug PJ. Phylogenomics supports Panpulmonata: opisthobranch paraphyly and key evolutionary steps in a major radiation of gastropod molluscs. *Mol Phylogenet Evol.* 2013;69(3):764–71.
- Schrödl M, Jorger KM, Klussmann-Kolb A, Wilson NG. Bye bye "Opisthobranchia"! A review on the contribution of mesopsammin sea slugs to euthyneuran systematics. *Thalassas.* 2011;27:101–12.
- Göbbeler K, Klussmann-Kolb A. Out of Antarctica?—new insights into the phylogeny and biogeography of the Pleurobranchomorpha (Mollusca, Gastropoda). *Mol Phylogenet Evol.* 2010;55(3):996–1007.
- Medina M, Lal S, Valles Y, Takaoka TL, Dayrat BA, Boore JL, et al. Crawling through time: transition of snails to slugs dating back to the Paleozoic, based on mitochondrial phylogenomics. *Mar Genomics.* 2011;4(1):51–9.
- Wägele H, Klussmann-Kolb A, Verbeek E, Schrödl M. Flashback and foreshadowing—a review of the taxon Opisthobranchia. *Organisms Div Evol.* 2014;14(1):133–49.
- Dayrat B, Conrad M, Balayan S, White TR, Albrecht C, Golding R, et al. Phylogenetic relationships and evolution of pulmonate gastropods (Mollusca): new insights from increased taxon sampling. *Mol Phylogenet Evol.* 2011;59(2):425–37.



30. Schrödl M. Time to say "Bye-bye Pulmonata"? *Spixiana*. 2014;37(2):161–4.
31. Thomaz D, Guiller A, Clarke B. Extreme divergence of mitochondrial DNA within species of pulmonate land snails. *Proceedings of the Royal Society B-Biological Sciences*. 1996;263(1368):363–8.
32. Hayashi M, Chiba S. Intraspecific diversity of mitochondrial DNA in the land snail *Euhadra peliomphala* (Bradybaenidae). *Biol J Linn Soc*. 2000;70(3):391–401.
33. Chiba S. Accelerated evolution of land snails *Mandarina* in the oceanic Bonin Islands: Evidence from mitochondrial DNA sequences. *Evolution*. 1999;53(2):460–71.
34. Pinceel J, Jordaens K, Backeljau T. Extreme mtDNA divergences in a terrestrial slug (Gastropoda, Pulmonata, Arionidae): accelerated evolution, allopatric divergence and secondary contact. *J Evol Biol*. 2005;18(5):1264–80.
35. Guiller A, Coutellec-Vreto MA, Madec L, Deunff J. Evolutionary history of the land snail *Helix aspersa* in the Western Mediterranean: preliminary results inferred from mitochondrial DNA sequences. *Mol Ecol*. 2001;10(1):81–7.
36. Pfenninger M, Posada D. Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. *Evolution*. 2002;56(9):1776–88.
37. Davison A, Clarke B. History or current selection? A molecular analysis of 'area effects' in the land snail *Cepaea nemoralis*. *Proceedings of the Royal Society B-Biological Sciences*. 2000;267(1451):1399–405.
38. Dépraz A, Hausser J, Pfenninger M. A species delimitation approach in the *Trochulus seniculus/hispidus* complex reveals two cryptic species within a sharp contact zone. *BMC Evol Biol*. 2009;9:171.
39. Goodacre SL. Population structure, history and gene flow in a group of closely related land snails: genetic variation in *Partula* from the Society Islands of the Pacific. *Mol Ecol*. 2002;11(1):55–68.
40. Parmakellis A, Kotsakiozi P, Rand D. Animal mitochondria, positive selection and cyto-nuclear coevolution: insights from pulmonates. *PLoS One*. 2013;8(4), e61970.
41. James JE, Piganeau G, Eyre-Walker A. The rate of adaptive evolution in animal mitochondria. *Mol Ecol*. 2016;25(1):67–78.
42. Wakeley J. The effects of subdivision on the genetic divergence of populations and species. *Evolution*. 2000;54(4):1092–101.
43. Thacker RW, Hadfield MG. Mitochondrial Phylogeny of Extant Hawaiian Tree Snails (Achatinellinae). *Mol Phylogenet Evol*. 2000;16(2):263–70.
44. Romero P, Ramirez R. Intraspecific divergence and DNA barcodes in *Systrophia helicycladae* (Gastropoda, Scolodontidae). *Rev Peru Biol*. 2011;18(2):201–8.
45. Sauer J, Hausdorf B. Reconstructing the evolutionary history of the radiation of the land snail genus *Xerocrassa* on Crete based on mitochondrial sequences and AFLP markers. *BMC Evol Biol*. 2010;10:299.
46. Nolan JR, Bergthorsson U, Adema CM. *Physella acuta*: atypical mitochondrial gene order among panpulmonates (Gastropoda). *J Molluscan Stud*. 2014; 80(4):388–99.
47. Pfenninger M, Cordellier M, Streit B. Comparing the efficacy of morphologic and DNA-based taxonomy in the freshwater gastropod genus *Radix* (Basommatophora, Pulmonata). *BMC Evol Biol*. 2006;6:100.
48. Zhai W, Nielsen R, Goldman N, Yang Z. Looking for Darwin in genomic sequences—validity and success of statistical methods. *Mol Biol Evol*. 2012; 29(10):2889–93.
49. Tomasco IH, Lessa EP. The evolution of mitochondrial genomes in subterranean caviomorph rodents: adaptation against a background of purifying selection. *Mol Phylogenet Evol*. 2011;61(1):64–70.
50. Zhang J, Nielsen R, Yang Z. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol*. 2005;22(12):2472–9.
51. Yang Z, dos Reis M. Statistical properties of the branch-site test of positive selection. *Mol Biol Evol*. 2011;28(3):1217–28.
52. Yang Z. *Molecular Evolution: A Statistical Approach*. Oxford: Oxford University Press; 2014.
53. Young JN, Rickaby RE, Kapralov MV, Filatov DA. Adaptive signals in algal Rubisco reveal a history of ancient atmospheric carbon dioxide. *Philos Trans R Soc Lond B Biol Sci*. 2012;367(1588):483–92.
54. Gharib WH, Robinson-Rechavi M. The branch-site test of positive selection is surprisingly robust but lacks power under synonymous substitution saturation and variation in GC. *Mol Biol Evol*. 2013;30(7):1675–86.
55. Berthelot C, Brunet F, Chalopin D, Juanchich A, Bernard M, Noel B, et al. The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nat Commun*. 2014;5:3657.
56. Beckstead WA, Ebbert MT, Rowe MJ, McClellan DA. Evolutionary pressure on mitochondrial cytochrome b is consistent with a role of Cytb17T affecting longevity during caloric restriction. *PLoS One*. 2009;4(6), e5836.
57. Facon B, Machinle E, Pointier JP, David P. Variation in desiccation tolerance in freshwater snails and its consequences for invasion ability. *Biol Invasions*. 2004;6:283–93.
58. Flores-Molina MR, Thomas D, Lovazzano C, Núñez A, Zapata J, Kumar M, et al. Desiccation stress in intertidal seaweeds: Effects on morphology, antioxidant responses and photosynthetic performance. *Aquat Bot*. 2014;113:90–9.
59. Mittler R. Oxidative stress, antioxidants and stress tolerance. *Trends Plant Sci*. 2002;7(9):405–10.
60. Shen YY, Liang L, Zhu ZH, Zhou WP, Irwin DM, Zhang YP. Adaptive evolution of energy metabolism genes and the origin of flight in bats. *Proc Natl Acad Sci U S A*. 2010;107(19):8666–71.
61. Da Silva C, Tomasco IH, Hoffmann F, Lessa EP. Genes and ecology: accelerated rates of replacement substitutions in the cytochrome b gene of subterranean rodents. *The Open Evolution Journal*. 2009;3:17–30.
62. Caballero S, Duchene S, Garavito MF, Slikas B, Baker CS. Initial evidence for adaptive selection on the NADH subunit Two of freshwater dolphins by analyses of mitochondrial genomes. *PLoS One*. 2015;10(5), e0123543.
63. Pfenninger M, Weigand A, Balint M, Klusmann-Kolb A. Misperceived invasion: the Lusitanian slug (*Arion lusitanicus* auct. non-Mabille or *Arion vulgaris* Moquin-Tandon 1855) is native to Central Europe. *Evol Appl*. 2014; 7(6):702–13.
64. Weigand AM, Jochum A, Pfenninger M, Steinke D, Klusmann-Kolb A. A new approach to an old conundrum—DNA barcoding sheds new light on phenotypic plasticity and morphological stasis in microsnails (Gastropoda, Pulmonata, Carychiidae). *Mol Ecol Resour*. 2011;11(2):255–65.
65. Steinke D, Albrecht C, Pfenninger M. Molecular phylogeny and character evolution in the Western Palaearctic Helicidae s.l. (Gastropoda: Stylommatophora). *Mol Phylogenet Evol*. 2004;32(3):724–34.
66. Feldmeyer B, Hoffmeier K, Pfenninger M. The complete mitochondrial genome of *Radix balthica* (Pulmonata, Basommatophora), obtained by low coverage shot gun next generation sequencing. *Mol Phylogenet Evol*. 2010; 57(3):1329–33.
67. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28(12): 1647–9.
68. Bernt M, Donath A, Juhling F, Externbrink F, Florentz C, Fritzsch G, et al. MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol*. 2013;69(2):313–9.
69. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013; 30(4):772–80.
70. Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol*. 2007;56(4):564–77.
71. Abascal F, Zardoya R, Telford MJ. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res*. 2010;38:W7–13.
72. Williams ST, Foster PG, Littlewood DT. The complete mitochondrial genome of a turbinid vetigastropod from MiSeq Illumina sequencing of genomic DNA and steps towards a resolved gastropod phylogeny. *Gene*. 2014;533(1):38–47.
73. Lanfear R, Calcott B, Kainer D, Mayer C, Stamatakis A. Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evol Biol*. 2014;14:82.
74. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 2006; 22(21):2688–90.
75. Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol*. 2008;57(5):758–71.
76. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: 2010 Gateway Computing Environments Workshop (GCE). New Orleans: IEEE; 2010. p. 1–8.
77. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 2012;61(3):539–42.
78. Rambaut A, Suchard MA, Xie D, Drummond AJ. Tracer v1.6. 2014.
79. Douady CJ, Delsuc F, Boucher Y, Doolittle WF, Douzery EJ. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Mol Biol Evol*. 2003;20(2):248–54.
80. Leaché AD, Reece TW. Molecular systematics of the Eastern Fence Lizard (*Sceloporus undulatus*): a comparison of Parsimony, Likelihood, and Bayesian approaches. *Syst Biol*. 2002;51(1):44–68.

81. Tamura K, Stecher G, Peterson D, Filipiński A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol.* 2013;30(12):2725–9.
82. Merker S, Thomas S, Volker E, Perwitasari-Farajallah D, Feldmeyer B, Streit B, et al. Control region length dynamics potentially drives amino acid evolution in tarsier mitochondrial genomes. *J Mol Evol.* 2014;79(1–2):40–51.
83. Zeileis A, Grothendieck G. zoo: S3 infrastructure for regular and irregular time series. *J Stat Softw.* 2005;14(6):1–17.
84. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 2007;24(8):1586–91.
85. Yang Z, Nielsen R, Goldman N, Pedersen AMK. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics.* 2000; 155(1):431–49.
86. Woolley S, Johnson J, Smith MJ, Crandall KA, McClellan DA. TreeSAAP: selection on amino acid properties using phylogenetic trees. *Bioinformatics.* 2003;19(5):671–2.
87. McClellan DA, McCracken KG. Estimating the influence of selection on the variable amino acid sites of the cytochrome b protein functional domains. *Mol Biol Evol.* 2001;18(6):917–25.
88. Romero PE, Weigand AM, Pfenninger M. Positive selection on panpulmonate mitogenomes provide new clues on adaptations to terrestrial life. *FigShare.* 2016. <https://dx.doi.org/10.6084/m9.figshare.c.3291377>.

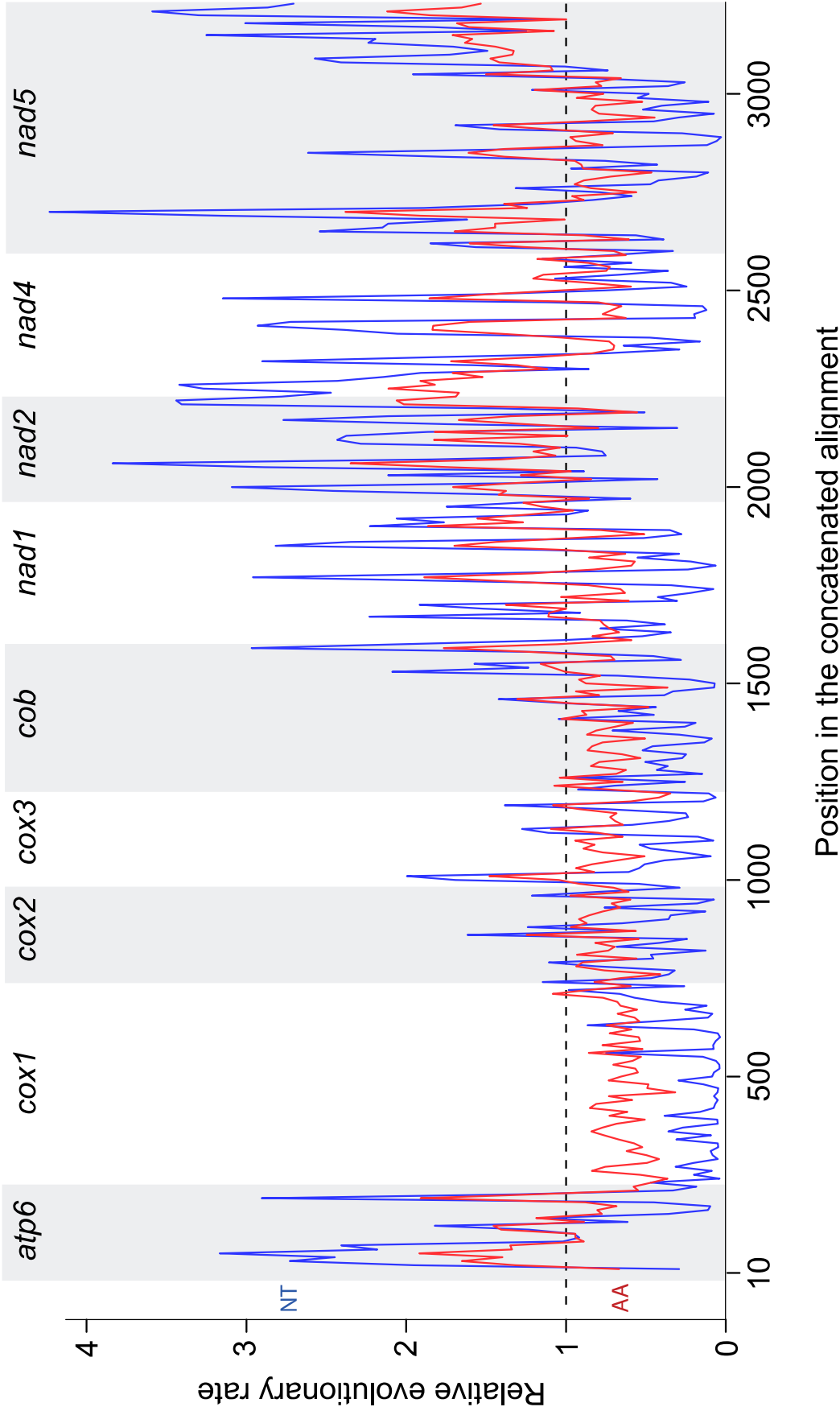
Submit your next manuscript to BioMed Central  
and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

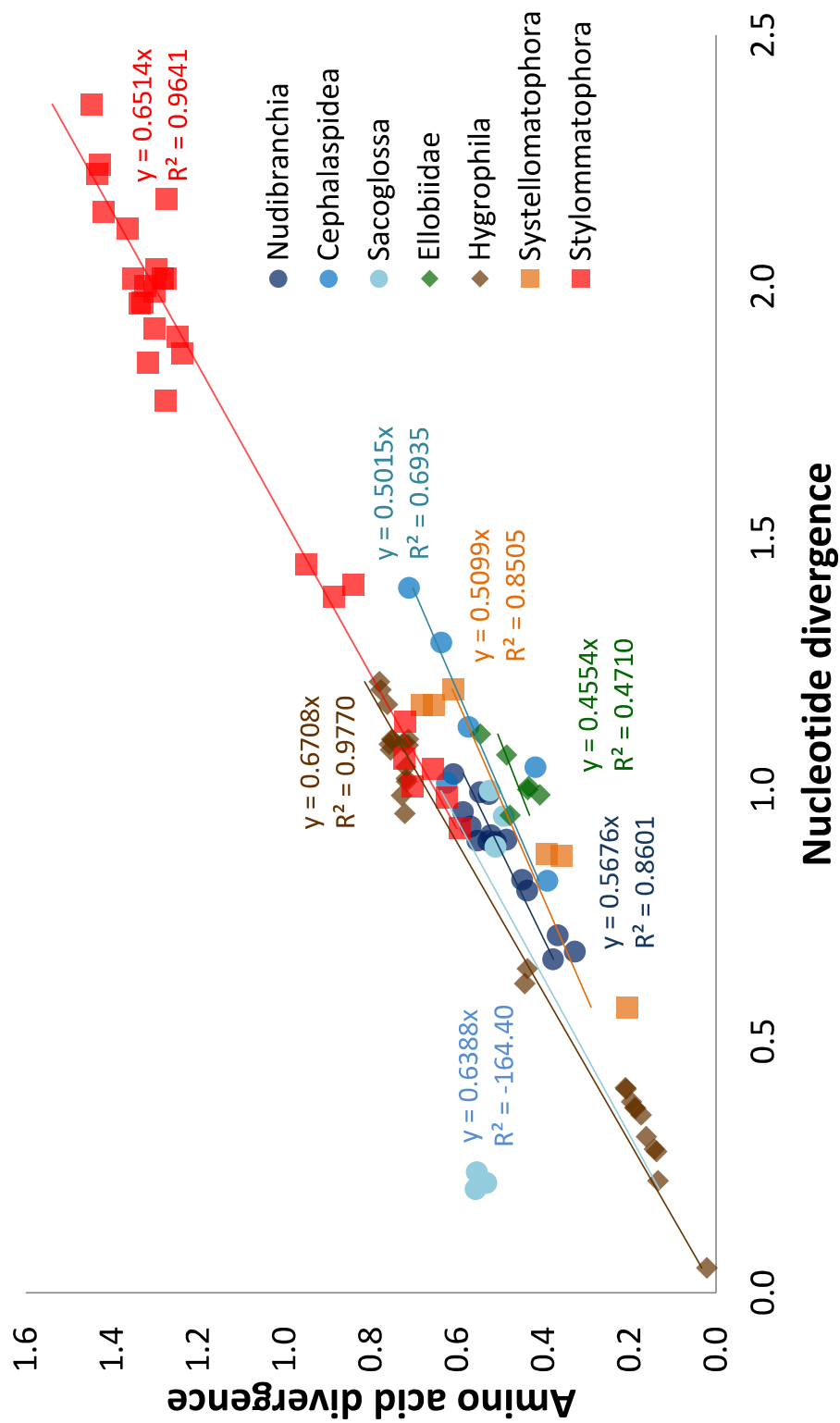


Additional File 1: Evolutionary rates for nucleotides (NT; blue line) and amino acids (AA; red line) in the euthyneuran mitogenomes. Rates are scaled such that the average evolutionary rate across all sites is 1 (red line). The x-axis shows amino acid positions in the final concatenated alignment.





Additional File 2: Amino acid divergence versus nucleotide divergence in mitochondrial genomes of euthyneuran gastropods. Clades are differentiated by colors and symbols as shown in the legend.



*Manuscript 1: Romero PE, Feldmeyer B, Pfenninger M. Panpulmonate transcriptomes reveal candidate genes involved in the adaptation to terrestrial and freshwater habitats in gastropods. Under review.*

### Anlage 3

#### Erklärung zu den Autorenanteilen

an der Publikation / des Manuskripts (Titel): Panpulmonate transcriptomes reveal candidate genes involved in the adaptation to terrestrial and freshwater habitats in gastropods.

Status (*submitted, under review, in revision, accepted*): Submitted, 03.09.2016.

Ab *accepted* Name der Zeitschrift:

Beteiligte Autoren (Namen mit eindeutigen Kürzeln):

- PER: Pedro Eduardo Romero
- BF: Barbara Feldmeyer
- MP: Markus Pfenninger

#### Was hat der Promovierende bzw. was haben die Koautoren beigetragen?

##### (1) zu Entwicklung und Planung

Promovierender PER: 50%

Co-Autor BF: 20%

Co-Autor MP: 30%

##### (2) zur Durchführung der einzelnen Untersuchungen und Experimente

Promovierender PER: 100%. RNA extraction and sequencing.

##### (3) zur Erstellung der Datensammlung und Abbildungen

Promovierender PER: 70%. Sampling in the Ryuku islands. Submission of raw RNA-seq data and transcriptomes to Genbank.

Co-Autor BF: 30%. Sampling of *Candidula* species. Submission of raw RNA-seq data and transcriptomes to Genbank.

##### (4) zur Analyse und Interpretation der Daten

Promovierender PER: 50%. RNA sequence analysis, multiple sequence alignments. Transcriptome assembly and annotation. Phylogenetic reconstruction. Orthogroups clustering. Positive selection analysis. Gene ontology.

Co-Autor BF: 40%. Transcriptome assembly and annotation. Orthogroups clustering. Interpretation of selection patterns. Gene ontology.

Co-Autor MP: 10%. Random association test. Interpretation of selection patterns.

##### (5) zum Verfassen des Manuskripts

Promovierender PER: 70%

Co-Autor BF: 15%

Co-Autor MP: 15%

Datum/Ort: 01.12.2016, Frankfurt am Main.

---

Unterschrift Promovend: \_\_\_\_\_

#### Zustimmende Bestätigungen der oben genannten Angaben

Unterschrift Betreuer: \_\_\_\_\_ Datum/Ort: 01.12.2016, Frankfurt am Main.

Ggfs. Unterschrift *corresponding author* PER: \_\_\_\_\_ Datum/Ort: 01.12.2016, Frankfurt am Main.

**Title**

Panpulmonate transcriptomes reveal candidate genes involved in the adaptation to terrestrial and freshwater habitats in gastropods

**Authors' names and affiliations**

Pedro E. Romero<sup>1,2</sup>

[pedro.romero@senckenberg.de](mailto:pedro.romero@senckenberg.de)

1 Senckenberg Biodiversity and Climate Research Centre (BiK-F), Senckenberganlage 25, 60325 Frankfurt am Main, Germany.

2 Institute for Ecology, Evolution & Diversity, Faculty of Biological Sciences, Goethe University Frankfurt, Max-von-Laue-Straße 13, 60438 Frankfurt am Main, Germany.

Barbara Feldmeyer<sup>1</sup>

[bfeldmeyer@senckenberg.de](mailto:bfeldmeyer@senckenberg.de)

1 Senckenberg Biodiversity and Climate Research Centre (BiK-F), Senckenberganlage 25, 60325 Frankfurt am Main, Germany.

Markus Pfenninger<sup>1,2</sup>

[pfenninger@bio.uni-frankfurt.de](mailto:pfenninger@bio.uni-frankfurt.de)

1 Senckenberg Biodiversity and Climate Research Centre (BiK-F), Senckenberganlage 25, 60325 Frankfurt am Main, Germany.

2 Institute for Ecology, Evolution & Diversity, Faculty of Biological Sciences, Goethe University Frankfurt, Max-von-Laue-Straße 13, 60438 Frankfurt am Main, Germany.

**Corresponding author**

Pedro E. Romero <sup>1,2</sup>

[pedro.romero@senckenberg.de](mailto:pedro.romero@senckenberg.de)

1 Senckenberg Biodiversity and Climate Research Centre (BiK-F), Senckenberganlage 25, 60325 Frankfurt am Main, Germany

2 Institute for Ecology, Evolution & Diversity, Faculty of Biological Sciences, Goethe University Frankfurt, Max-von-Laue-Straße 13, 60438 Frankfurt am Main, Germany.

## **Abstract**

The conquest of the land from aquatic habitats is a fascinating evolutionary event that happened multiple times in different phyla. Mollusks are among the organisms that successfully invaded the non-marine realm, resulting in the radiation of terrestrial panpulmonate gastropods. We compared panpulmonate transcriptomes to study the selective pressures that modeled the transitions from marine into terrestrial and freshwater realms in this molluscan lineage. *De novo* assembly of six panpulmonate transcriptomes resulted in 55,000 - 97,000 predicted open reading frames, of which 9 - 14% were functionally annotated. Adding published transcriptomes, we predicted 791 ortholog clusters shared among fifteen panpulmonate species, resulting in 700 amino acid and 736 codon-wise alignments. The branch-site test of positive selection applied to the codon-wise alignments showed seven genes under positive selection in the terrestrial lineages and twenty-eight genes, in the freshwater lineages. Gene ontology categories of these candidate genes include actin assembly, transport of glucose, and the tyrosine metabolism in the terrestrial lineages; and, DNA repair, metabolism of xenobiotics, mitochondrial electron transport, and ribosome biogenesis in the freshwater lineages. We identified candidate genes representing processes that may have played a key role during the water-to-land transition in Panpulmonata. These genes were involved in energy metabolism and gas-exchange surface development in the terrestrial lineages and in the response to the abiotic stress factors (UV radiation, osmotic pressure, xenobiotics) in the freshwater lineages. Our study expands the knowledge of possible adaptive signatures in genes and metabolic pathways related to the invasion of non-marine habitats in invertebrates.

## **Keywords**

Realm transition, Phylogenomics, Panpulmonata, Positive selection, Terrestrialization.

## 1. Introduction

The invasion from marine to non-marine habitats is one of the most enthralling events in the evolution of life on Earth. The transition from sea to land and freshwater environments occurred multiple times in different branches of the tree of life. Mollusks, along arthropods and vertebrates, are among the successful phyla that invaded the non-marine realm. Several branches from the molluscan class Gastropoda (Neritimorpha, Cyclophoroidea, Littorinoidea, Risssooidea, and Panpulmonata) have colonized terrestrial habitats multiple times [1, 2]. Especially, several independent land invasions in the Panpulmonata resulted in a significant adaptive radiation and explosive diversification that likely originated up to a third of the extant molluscan diversity [3]. Therefore, panpulmonate lineages are a promising system to study evolution of adaptations to non-marine habitats.

The habitat transition must have triggered several novel adaptations in behavior, breathing, excretion, locomotion, and osmotic and temperature regulation, to overcome problems that did not exist in the oceans such as dehydration, lack of buoyancy force, extreme temperature fluctuations and radiation damage [4-6]. Studies in vertebrates showed different genomic changes involved in the adaptation to the new habitats. Mudskippers, amphibious teleost fishes adapted to live on mudflats, possess unique immune genes to possibly counteract novel pathogens on land, and opsin genes for aerial vision and for enhancement of color vision [7]. Tetrapods showed adaptation signatures in the carbamoyl phosphate synthase I (CPS1) gene involved in the efficient production of hepatic urea [8]. Primitive sarcopterygians like the coelacanth *Latimeria* already possess various conserved non-coding elements (CNE) that enhance the development of limbs, and an expanded repertoire of genes related to the pheromone receptor VR1 that may have facilitated the adaptation to sense airborne chemicals during the water-to-land transition in tetrapods [9]. Also, vertebrate keratin genes responsible for skin rigidity underwent a functional diversification after the water-to-land transition, enhancing the protection against friction imposed by the new terrestrial lifestyle [10].

Conversely, information about the molecular basis of adaptation from marine to non-marine habitats in invertebrates is still scarce. Only one study reported adaptive signals in gene families (e. g. ATPases, DNA repair, and ribosomal proteins) that may have played a key role during terrestrialization in springtails and insects (Hexapoda) [11], clades that probably had a common pancrustacean ancestor living in a shallow marine environment [12, 13]. Mutations in the ATPases were suggested to provide the necessary energy to adapt to new high-energy demanding habitats [14], DNA repair genes would have helped reducing the damage produced by increased ultraviolet (UV) irradiation, and finally, as the ribosomal machinery is salt-sensitive, adaptive signs in the ribosomal proteins could have been a result of the different osmotic pressures within aquatic and terrestrial environments [15].

In a previous paper, we explored the adaptive signals in the mitochondrial genomes of panpulmonates [16]. We found that in the branches leading to lineages with terrestrial taxa (Ellobioidea and Stylommatophora), the mitochondrial genes *cob* and *nad5*, both involved in the oxidative phosphorylation pathway that finally produces ATP, appeared under positive selection. Moreover, the amino acid positions under selection have been related to an increased energy production probably linked to novel demands of locomotion [17, 18], and to changes in the equilibrium constant physicochemical property involved in the regulation of ROS production and thus, in the ability to tolerate new abiotic stress conditions [19].

Here, we expanded our search for candidate genes related to the adaptation to non-marine habitats, using transcriptome-wide data from several panpulmonate taxa, including marine, intertidal, freshwater and terrestrial lineages. We used a phylogenomic approach to reconstruct the evolutionary relationships of Panpulmonata and then tested for positive selection in the land and freshwater branches. Our approach aims to provide new insights into the selective pressures shaping the transition from marine to land and freshwater lifestyles.



## **2. Material and Methods**

### **2.1. Dataset collection**

The dataset from Zapata et al. [20] was used as a starting point for our study. We added to this dataset the transcriptome from *Radix balthica* [21] and retrieved additional freshwater specimens from the NCBI Sequence Read Archive (SRA) (<http://www.ncbi.nlm.nih.gov/sra>). We complemented the dataset with five intertidal and terrestrial specimens from Ellobioidea (*Carychium* sp., *Cassidula plecotremata*, *Melampus flavus*, *Pythia pachyodon*, *Trimusculus* sp.) and one terrestrial Stylommatophora (*Arion vulgaris*), collected in Japan (2013) and Germany (2014), respectively (see Supplementary File 10 in Ref. [22]). RNA was isolated following the RNeasy kit (QIAGEN) following the manufacturer's protocol. cDNA production and sequencing on the Illumina NextSeq500 platform (150 bp paired- end reads) was performed by StarSEQ GmbH (Mainz, Germany), according to their Illumina standard protocol. The final dataset comprised fifteen transcriptomes of panpulmonate species occurring in marine, intertidal, freshwater and terrestrial habitats (Table 1). Raw sequence data are deposited in the NCBI Sequence Read Archive as BioProject (PRJNA339817).

### **2.2. Read processing and quality checking**

FastQC [23] was used for initial assessment of reads quality. Then, Trimmomatic v0.33 [24] was used to remove and trim Illumina adaptor sequences and other reads with an average quality below 15 within a 4-base wide sliding window. In addition, we repeated the trimming analysis specifying a minimum length of 25 nt for further assembly comparisons. The same procedure was applied to all samples, except for *Radix* (454 reads). In this latter case, we got the transcriptome assembly directly from the author [21].

### **2.3. Transcriptome assembly**

*De novo* assembly was performed for all samples, except *Radix* (see last section), using Trinity v2.0.6 [25] with a minimum contig length of 100 amino acids, and Bridger v2014-12-

01 [26] with default options. Bridger required the trimmed set with the minimum length of 25 nt. We combined the results from Trinity and Bridger in a meta-assembly using MIRA [27] with default settings. Only sequences with longer than 100 aa were retained for further analyses. This step was done to improve the accuracy in ortholog determination and facilitate phylogenomic analyses [28]. Furthermore, we used the ORFpredictor server [29] to predict open reading frames (ORF) within the transcripts.

#### **2.4. Construction of ortholog clusters**

Ortholog clusters shared among protein sequences of the fifteen panpulmonate species were predicted using OrthoFinder [30] with default parameters. In case clusters contained more than one sequence per species, only a single sequence per species with the highest average similarity was selected using a homemade script. The predicted amino acid sequences from each ortholog cluster were aligned using MAFFT [31] with standard parameters. Nucleotide sequences in each orthogroup were aligned codon-wise using TranslatorX [32] taking into account the information from the amino acid alignments. Ambiguous aligned regions from the amino acid or codon alignments were removed using Gblocks [33] with standard settings. We used TrimAL [34] to remove poorly aligned or incomplete sequences in each ortholog cluster, using a minimum residue overlap score of 0.75.

#### **2.5. Phylogenomic analyses**

Phylogenetic relationships among the Panpulmonata were reconstructed based on a subset of 382 ortholog clusters. The subset selection was done using MARE [35], a tool designed to find informative subsets of genes and taxa within a large phylogenetic dataset of amino acid sequences. The concatenated amino acid alignment length resulted in 88622 positions. Data were partitioned by gene using the partition scheme suggested in Partition-Finder [36] using the *-rcluster* option (relaxed hierarchical clustering algorithm), suitable for phylogenomic data [37]. We reconstructed an unrooted tree to be used as an input for the

selection analyses. Maximum likelihood analyses were conducted in RAxML-HPC2 (8.0.9) [38]. Branch support was evaluated using 1000 rapid bootstraps, and confidence values were drawn in the best-scoring tree. Bayesian inference was conducted in MrBayes v3.2.2 [39]. Four simultaneous Monte Carlo Markov Chains (MCMC) were run, with the following parameters: eight chains of 20 million generations each, sampling every 20000 generations and a burn-in of 25%. Tracer 1.6 [40] was used to evaluate effective sample sizes (ESS > 200). We assume that a bootstrap value of >70% and a posterior probability of > 0.95 are evidence of significant nodal support.

## 2.6. Selection analyses

The test of positive selection was performed for 736 ortholog clusters (codon-wise alignments) in CODEML implemented in the software PAML v4.8 [41]. PAML estimated the omega ratio ( $\omega = dN: \text{non-synonymous sites} / dS: \text{synonymous sites}$ );  $\omega = 1$  indicates neutral evolution,  $\omega < 1$  purifying selection, and  $\omega > 1$  indicates positive selection [42]. To detect positive selection affecting sites along the terrestrial or freshwater branches (foreground) in comparison to the intertidal or marine lineages (background), the branch-site model A [43] in CODEML was applied (model = 2, NSsites = 2) for each ortholog cluster. The unrooted tree obtained using maximum likelihood was set as the guide tree. In order to avoid problems in convergence in the log-likelihood calculations, we ran three replicates of model A with different initial omega values ( $\omega = 0.5$ ,  $\omega = 1.0$ ,  $\omega = 5.0$ ). We also calculated the likelihood of the null model (model = 2, NSsites = 2, fixed  $\omega = 1.0$ ). Both models were compared in a likelihood ratio test ( $LRT = 2 * (\ln L \text{ model A} - \ln L \text{ null model})$ ). The Bayes Empirical Bayes (BEB) algorithm implemented in CODEML was used to calculate posterior probabilities of positive selected sites. We corrected p-values with a false discovery rate (FDR) cut-off value of 0.05 using the Benjamini and Hochberg method [44] implemented in R. The statistical significance of the overlap between positively selected genes from freshwater and terrestrial lineages was calculated using the R function *phyper*.

## 2.7. Functional annotation

The transcripts were annotated using BlastX [45]. We blasted the nucleotide sequences against the invertebrate protein sequence RefSeq database (release 73, November 2015), with an e-value cut-off of  $10^{-6}$ . We selected top hits with the best alignment and the lowest e-value. Gene ontology (GO) terms for each BLASTx search were obtained in the Blast2GO suite [46]. Functional annotation information was obtained from InterPro database [47] using the InterProScan [48]. GO terms were then assigned to each ortholog group that was found under positive selection. In addition, we added to this clusters the metabolic pathway information retrieved from the KAAS server [49]. This server assigns orthology identifiers from the KEGG database (Kyoto Encyclopedia of Genes and Genomes). Functional enrichment analysis using the Fisher exact test was also performed in Blast2GO comparing the genes under positive selection against all ortholog clusters.

## 3. Results

We generated approximately 2,100,000 - 3,400,000 Illumina for our six samples (five ellobiids and one stylommatophoran species, Table 1). The quality trimming eliminated 14 - 39% of short and low-quality fragments in our samples. *De novo* meta assembly with MIRA produced approximately 55,000 - 98,000 transcripts in our samples and 54,000 - 130,000 in the other additional samples (Table 1). For further analyses we used transcripts larger than 300 bp. This represented a reduction of less than 1% in our samples but a higher reduction in the public data (3 - 35%). The number of predicted open reading frames (55,000 - 97,000) was very similar to the number of transcripts > 300 bp in almost all cases, the only exception was *Radix balthica*, where only 57% of the transcripts obtained an ORF prediction. We obtained 9,000 - 30,000 single blast hits for our data, representing 5,000 - 13,000 single annotated genes. The percentage of annotated genes from our open reading frame data was 9 - 14%.

We predicted 791 ortholog clusters shared among all species, of which 700 ortholog clusters remained after removing spurious and poorly amino acid aligned sequences in trimAL. From this dataset, MARE selected 382 informative clusters to reconstruct the phylogeny of the panpulmonate species. The amount of missing data corresponds to 10.94% in the complete matrix, and 6.26% in the reduced matrix (see Supplementary Files 5 and 6 in Ref. [22], respectively).

Most branches in the panpulmonate tree received high support (Figure 1). The clade containing Stylommatophora and Systellommatophora was significantly supported (bootstrap: 94 / posterior probability: 1.0) and appeared as a sister of the monophyletic Ellobioidea (99/1.0). The Acochlidia clade was moderately supported (86/1.0). The association of the Acochlidia with the Ellobioidea, Stylommatophora, Systellommatophora clade had no significant bootstrap support but a high posterior probability (64/1.0). The Hygrophila clade was highly supported (100/1.0). The association of Amphiboloidea and Pyramidelloidea was also highly supported (100/1.0).

We detected selection signatures on genes (codon-wise alignments) across the terrestrial and freshwater lineages in Panpulmonata. The likelihood-ratio test (LRT) comparing the branch-site model A against the null model (neutral) showed seven ortholog clusters under positive selection in the land lineages and twenty-eight clusters in the freshwater lineages (see Supplementary File 7 in Ref. [22]). There was no overlapping within positively selected genes from the terrestrial and freshwater lineages. Table 2 shows examples of these candidate genes, their annotations, biological processes, molecular functions, and pathways involved. The BlastX annotations revealed candidate genes involved in the actin assembly, protein folding, transport of glucose, and vesicle transport in the terrestrial lineages. In the freshwater lineages, we found candidate genes associated to DNA repair, metabolism of xenobiotics, mitochondrial electron transport, protein folding, proteolysis, ribosome biogenesis, RNA processing and transport of lipids (see Supplementary Files 8 and 9 in Ref. [22]).

We found significant enriched GO (Gene ontology) terms neither in the terrestrial nor freshwater lineages.

Candidate genes under positive selection in the terrestrial lineages were involved in the carbohydrate digestion, endocytosis, focal adhesion, and the metabolism of lipids and tyrosine pathways. In case of the freshwater lineages, the candidate genes were involved in several metabolic pathways, for example, amino acid biosynthesis, focal adhesion, lysosome, oxidative phosphorylation, and protein signaling (Table 2, and Supplementary Files 8 and 9 in Ref. [22]).

#### **4. Discussion**

Panpulmonates transitioned from marine to terrestrial environments in at least four lineages: Acochlidia, Ellobioidea, Stylommatophora, and Systellomatophora, and to freshwater environments in at least three lineages: Acochlidia, Glacidorboidea, and Hygrophila [2, 50-53]. Thus, they are a very suitable model to study the invasion of non-marine realms. However, the phylogenetic relationships within this clade are yet to be resolved [50]. Our tree topology using 382 ortholog clusters resembles the one obtained from Jörger et al. [51], based on mitochondrial and nuclear markers. In addition, we found support for the Geophila: Stylommatophora (terrestrial) and Systellomatophora (intertidal/terrestrial) as sister groups. This clade has been proposed before based on the position of the eyes at the tip of cephalic tentacles [54]. Still, previous phylogenies using mitochondrial and nuclear markers failed to support this clade [16, 51, 52, 55]. We also found support for Eupulmonata (*sensu* Morton [55, 56]), a clade comprising Stylommatophora and Systellomatophora plus Ellobioidea (intertidal/terrestrial) [50], this clade was supported using a combination of mitochondrial and nuclear markers [51]. Generation of high-quality transcriptomic data for other panpulmonate clades (marine Sacoglossa and Siphonarioidea, freshwater Glacidorboidea), and additional data for terrestrial Stylommatophora and Systellomatophora, will definitively illuminate the evolutionary relationships in Panpulmonata.

Our study is the first genome-wide report on the molecular basis of adaptation to non-marine habitats in panpulmonate gastropods. In case of the terrestrial lineages, we found evidence that the different positively selected genes are involved in a general pattern of adaptation to increased energy demands. The adaptive signs found in a gene related to actin assembly (OG0001172, Table 2) can be related to the necessity to move (forage, hunt preys or escape from predators) in the terrestrial realm. Moreover, the displacement in an environment lacking the buoyancy force to float or swim requires more energy, which can be obtained by increasing the glucose uptake (OG0000137) to produce energy in form of ATP. The adaptive signatures we found before in two mitochondrial genes, *cob* and *nad5*, involved in energy production in the mitochondrion, also suggested a response to new metabolic requirements in the terrestrial realm, such as the increase of energy demands (to move and sustain the body mass).

One gene found under positive selection in the terrestrial genus *Pythia*, was involved in the metabolism of tyrosine (OG0000060). Tyrosine is the principal component of the thyroid hormones (TH). Despite invertebrates lack the thyroid gland responsible of the production of TH's; the synthesis of TH's has been demonstrated in mollusks and echinoderms. In these organisms, iodine is ligated to the tyrosine in the peroxisomes, producing thyroid hormones [57]. Notably, it has been suggested that iodinated tyrosine may have been essential in vertebrates during the transition to terrestrial habitats for TH's are required in the expression of transcription factors involved in the embryonic development and differentiation of the lungs [58]. Land snails adapted to breath air by losing their gills and transforming the inner surface of their mantle into a lung [5]. Therefore, we propose that the tyrosine pathway was also a key component in invertebrates probably promoting the development of novel gas exchange tissues in land snails.

A gene similar to cytochrome P450 was also found under positive selection (OG000120). Cytochrome P450s are proteins involved in the metabolism of xenobiotics. They were also under positive selection in the terrestrial Hexapoda lineages in comparison to other water-dwelling arthropods [11]. This result suggests that adaptations in these genes probably improve the response to new organic pollutants and toxins absent in the marine realm.

Another gene that showed adaptive signatures was the 40S ribosomal protein S3a (OG0002708). Likewise, ribosomal genes were also identified in a previous study on land-to-water transitions in hexapods [11] and plants [15]. In the latter study, it was suggested that the difference in the osmotic pressure from aquatic and terrestrial realms could affect the salt-sensitive ribosomal machinery, triggering adaptations to tolerate new salt conditions. This could also be the case for the freshwater animals (hypertonic) in comparison to the marine ones (hypotonic).

Finally, we found adaptive signatures in a DNA methyltransferase gene (OG0004116). This enzyme is part of the DNA repair system in the cell. Specifically, it removes methyl groups from O6-methylguanine produced by carcinogenic agents and it has been showed that its expression is regulated by the presence of ultraviolet B (UVB) radiation [59]. Positive selection on DNA repair genes has been found in hexapods [11], and in vertebrates living in high altitude environments (Tibetan antelopes) [14] or in mudflats (mudskippers) [7], suggesting an important role in the maintenance of the genomic integrity in response to the rise of temperature gradients or UV radiation in the terrestrial realms. In case of the aquatic environments, an extensive review has found an overall negative UVB effect on marine and freshwater animals [60]. However, the authors did not find a significant difference of the survival among taxonomic groups or levels of exposure in marine and freshwater realms, and suggested that the negative effects are highly variable among organisms and depends on several factors including cloudiness, ozone concentration, seasonality, topography, and behavior. Interestingly, it has been reported that survival in the freshwater snail *Physella*



*acuta* (Hygrophila) depends of the combination of a photoenzymatic repair system plus photoprotection provided by the shell thickness and active selection of locations below the water surface avoiding the sunlight [61]

## 5. Conclusions

We found that the positively selected genes in the terrestrial lineages were related to motility and to the development of novel gas-exchange tissues; while most of the genes in freshwater lineages were related to the response to abiotic stress such osmotic pressure, UV radiation and xenobiotics. These adaptations at the genomic level combined with novel responses in development and behavior probably facilitated the success during the transitions to the non-marine realm. Our results are very promising to understand the genomic basis of the adaptation during the sea-to-land transitions, and also highlight the necessity of more genome-wide studies especially in invertebrates, comparing marine, freshwater and terrestrial taxa, to unravel the evolution of the molecular pathways involved in the invasion of new realms.

## Acknowledgements

We thank Yasunori Kano and Annette Klussmann-Kolb for their assistance in the fieldwork and Tilman Schell for his assistance in the bioinformatics analyses. This work was supported by the German funding program “LOEWE – Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz” of the Hessen State Ministry of Higher Education, Research and the Arts. PER also received a scholarship from CONCYTEC/CIENCIACTIVA: Programa de becas de doctorado en el extranjero del Gobierno del Perú (291-2014-FONDECYT).

## 6. References

[1] Y. Kameda, M. Kato, Terrestrial invasion of pomatiopsid gastropods in the heavy-snow region of the Japanese Archipelago, *BMC Evol. Biol.* 11 (2011) 118.

- [2] P.E. Romero, M. Pfenninger, Y. Kano, A. Klussmann-Kolb, Molecular phylogeny of the Ellobiidae (Gastropoda: Panpulmonata) supports independent terrestrial invasions, *Mol. Phylogenet. Evol.* 97 (2016) 43-54.
- [3] K.M. Kocot, K.M. Halanych, P.J. Krug, Phylogenomics supports Panpulmonata: opisthobranche paraphyly and key evolutionary steps in a major radiation of gastropod molluscs, *Mol. Phylogenet. Evol.* 69 (2013) 764-771.
- [4] C. Little, *The terrestrial invasion: An ecophysiological approach to the origins of land animals*, Cambridge University Press, Cambridge, 1990.
- [5] M. Laurin, *How vertebrates left water*, University of California Press, Berkeley, 2010.
- [6] G. McGhee, *When the invasion of land failed: The legacy of the Devonian extinctions*, Columbia University Press, New York, 2013.
- [7] X. You, C. Bian, Q. Zan, X. Xu, X. Liu, J. Chen, J. Wang, Y. Qiu, W. Li, X. Zhang, Y. Sun, S. Chen, W. Hong, Y. Li, S. Cheng, G. Fan, C. Shi, J. Liang, Y. Tom Tang, C. Yang, Z. Ruan, J. Bai, C. Peng, Q. Mu, J. Lu, M. Fan, S. Yang, Z. Huang, X. Jiang, X. Fang, G. Zhang, Y. Zhang, G. Polgar, H. Yu, J. Li, Z. Liu, G. Zhang, V. Ravi, S.L. Coon, J. Wang, H. Yang, B. Venkatesh, J. Wang, Q. Shi, Mudskipper genomes provide insights into the terrestrial adaptation of amphibious fishes, *Nat Commun* 5 (2014) 5594.
- [8] C.T. Amemiya, J. Alföldi, A.P. Lee, S. Fan, H. Philippe, I. Maccallum, I. Braasch, T. Manousaki, I. Schneider, N. Rohner, C. Organ, D. Chalopin, J.J. Smith, M. Robinson, R.A. Dorrington, M. Gerdol, B. Aken, M.A. Biscotti, M. Barucca, D. Baurain, A.M. Berlin, G.L. Blatch, F. Buonocore, T. Burmester, M.S. Campbell, A. Canapa, J.P. Cannon, A. Christoffels, G. De Moro, A.L. Edkins, L. Fan, A.M. Fausto, N. Feiner, M. Forconi, J. Gamielien, S. Gnerre, A. Gnirke, J.V. Goldstone, W. Haerty, M.E. Hahn, U. Hesse, S. Hoffmann, J. Johnson, S.I. Karchner, S. Kuraku, M. Lara, J.Z. Levin, G.W. Litman, E. Mauceli, T. Miyake, M.G. Mueller, D.R. Nelson, A. Nitsche, E. Olmo, T. Ota, A. Pallavicini, S. Panji, B. Picone, C.P. Ponting, S.J. Prohaska, D. Przybylski, N.R. Saha, V. Ravi, F.J. Ribeiro, T. Sauka-Spengler, G. Scapigliati, S.M. Searle, T. Sharpe, O. Simakov, P.F. Stadler, J.J. Stegeman, K. Sumiyama, D. Tabbaa, H. Tafer, J. Turner-Maier, P. van Heusden, S. White, L. Williams, M. Yandell, H. Brinkmann, J.N. Volff, C.J. Tabin, N. Shubin, M. Scharl, D.B. Jaffe, J.H. Postlethwait, B. Venkatesh, F. Di Palma, E.S. Lander, A. Meyer, K. Lindblad-Toh, The African coelacanth genome provides insights into tetrapod evolution, *Nature* 496 (2013) 311-316.
- [9] M. Nikaido, H. Noguchi, H. Nishihara, A. Toyoda, Y. Suzuki, R. Kajitani, H. Suzuki, M. Okuno, M. Aibara, B.P. Ngatunga, S.I. Mzighani, H.W. Kalombo, K.W. Masengi, J. Tuda, S. Nogami, R. Maeda, M. Iwata, Y. Abe, K. Fujimura, M. Okabe, T. Amano, A. Maeno, T. Shiroishi, T. Itoh, S. Sugano, Y. Kohara, A. Fujiyama, N. Okada, Coelacanth genomes reveal signatures for evolutionary transition from water to land, *Genome Res* 23 (2013) 1740-1748.
- [10] W. Vandebergh, F. Bossuyt, Radiation and functional diversification of alpha keratins during early vertebrate evolution, *Mol. Biol. Evol.* 29 (2012) 995-1004.
- [11] A. Faddeeva, R.A. Studer, K. Kraaijeveld, D. Sie, B. Ylstra, J. Marien, H.J. op den Camp, E. Datema, J.T. den Dunnen, N.M. van Straalen, D. Roelofs, Collembolan transcriptomes highlight molecular evolution of hexapods and provide clues on the adaptation to terrestrial life, *PLoS. ONE* 10 (2015) e0130600.

- [12] B. Misof, S. Liu, K. Meusemann, R.S. Peters, A. Donath, C. Mayer, P.B. Frandsen, J. Ware, T. Flouri, R.G. Beutel, O. Niehuis, M. Petersen, F. Izquierdo-Carrasco, T. Wappler, J. Rust, A.J. Aberer, U. Aspöck, H. Aspöck, D. Bartel, A. Blanke, S. Berger, A. Böhm, T.R. Buckley, B. Calcott, J. Chen, F. Friedrich, M. Fukui, M. Fujita, C. Greve, P. Grobe, S. Gu, Y. Huang, L.S. Jermiin, A.Y. Kawahara, L. Krogmann, M. Kubiak, R. Lanfear, H. Letsch, Y. Li, Z. Li, J. Li, H. Lu, R. Machida, Y. Mashimo, P. Kapli, D.D. McKenna, G. Meng, Y. Nakagaki, J.L. Navarrete-Heredia, M. Ott, Y. Ou, G. Pass, L. Podsiadlowski, H. Pohl, B.M. von Reumont, K. Schütte, K. Sekiya, S. Shimizu, A. Slipinski, A. Stamatakis, W. Song, X. Su, N.U. Szucsich, M. Tan, X. Tan, M. Tang, J. Tang, G. Timelthaler, S. Tomizuka, M. Trautwein, X. Tong, T. Uchifune, M.G. Walz, B.M. Wiegmann, J. Wilbrandt, B. Wipfler, T.K. Wong, Q. Wu, G. Wu, Y. Xie, S. Yang, Q. Yang, D.K. Yeates, K. Yoshizawa, Q. Zhang, R. Zhang, W. Zhang, Y. Zhang, J. Zhao, C. Zhou, L. Zhou, T. Ziesmann, S. Zou, Y. Li, X. Xu, Y. Zhang, H. Yang, J. Wang, J. Wang, K.M. Kjer, X. Zhou, Phylogenomics resolves the timing and pattern of insect evolution, *Science* 346 (2014) 763-767.
- [13] B.M. von Reumont, R.A. Jenner, M.A. Wills, E. Dell'ampio, G. Pass, I. Ebersberger, B. Meyer, S. Koenemann, T.M. Iliffe, A. Stamatakis, O. Niehuis, K. Meusemann, B. Misof, Pancrustacean phylogeny in the light of new phylogenomic data: support for Remipedia as the possible sister group of Hexapoda, *Mol. Biol. Evol.* 29 (2012) 1031-1045.
- [14] R.L. Ge, Q. Cai, Y.Y. Shen, A. San, L. Ma, Y. Zhang, X. Yi, Y. Chen, L. Yang, Y. Huang, R. He, Y. Hui, M. Hao, Y. Li, B. Wang, X. Ou, J. Xu, Y. Zhang, K. Wu, C. Geng, W. Zhou, T. Zhou, D.M. Irwin, Y. Yang, L. Ying, H. Bao, J. Kim, D.M. Larkin, J. Ma, H.A. Lewin, J. Xing, R.N. Platt, 2nd, D.A. Ray, L. Auvil, B. Capitanu, X. Zhang, G. Zhang, R.W. Murphy, J. Wang, Y.P. Zhang, J. Wang, Draft genome sequence of the Tibetan antelope, *Nat Commun* 4 (2013) 1858.
- [15] L. Wissler, F.M. Codoner, J. Gu, T.B. Reusch, J.L. Olsen, G. Procaccini, E. Bornberg-Bauer, Back to the sea twice: identifying candidate plant genes for molecular evolution to marine life, *BMC Evol. Biol.* 11 (2011) 8.
- [16] P.E. Romero, A.M. Weigand, M. Pfenninger, Positive selection on panpulmonate mitogenomes provide new clues on adaptations to terrestrial life, *BMC Evol. Biol.* 16 (2016) 164.
- [17] Y.Y. Shen, L. Liang, Z.H. Zhu, W.P. Zhou, D.M. Irwin, Y.P. Zhang, Adaptive evolution of energy metabolism genes and the origin of flight in bats, *Proc. Natl. Acad. Sci. U. S. A* 107 (2010) 8666-8671.
- [18] D.A. McClellan, E.J. Palfreyman, M.J. Smith, J.L. Moss, R.G. Christensen, J.K. Sailsbery, Physicochemical evolution and molecular adaptation of the cetacean and artiodactyl cytochrome b proteins, *Mol. Biol. Evol.* 22 (2005) 437-455.
- [19] R. Mittler, Oxidative stress, antioxidants and stress tolerance, *Trends Plant Sci* 7 (2002) 405-410.
- [20] F. Zapata, N.G. Wilson, M. Howison, S.C. Andrade, K.M. Jorger, M. Schrödl, F.E. Goetz, G. Giribet, C.W. Dunn, Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda, *Proc Biol Sci* 281 (2014) 20141739.
- [21] B. Feldmeyer, C.W. Wheat, N. Krezdorn, B. Rotter, M. Pfenninger, Short read Illumina data for the de novo assembly of a non-model snail species transcriptome (*Radix balthica*, Basommatophora, Pulmonata), and a comparison of assembler performance, *BMC Genomics* 12 (2011) 317.

- [22] P.E. Romero, B. Feldmeyer, M. Pfenninger, Dataset for phylogenomics and detection of positive selection among marine, freshwater and terrestrial panpulmonate mollusks, *Data in Brief*.
- [23] S. Andrews, FastQC, in.
- [24] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinformatics* 30 (2014) 2114-2120.
- [25] M.G. Grabherr, B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B.W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, A. Regev, Full-length transcriptome assembly from RNA-Seq data without a reference genome, *Nat Biotechnol* 29 (2011) 644-652.
- [26] Z. Chang, G. Li, J. Liu, Y. Zhang, C. Ashby, D. Liu, C.L. Cramer, X. Huang, Bridger: a new framework for de novo transcriptome assembly using RNA-seq data, *Genome. Biol* 16 (2015) 30.
- [27] B. Chevreux, T. Wetter, S. Suhai, Genome sequence assembly using trace signals and additional sequence information in: *Comput Sci Biol Proc Ger Conf Bioinformatics (GCB)*, 1999, pp. 45-56.
- [28] L. Zhao, N. Zhang, P.F. Ma, Q. Liu, D.Z. Li, Z.H. Guo, Phylogenomic analyses of nuclear genes reveal the evolutionary relationships within the BEP clade and the evidence of positive selection in Poaceae, *PLoS ONE* 8 (2013) e64642.
- [29] X.J. Min, G. Butler, R. Storms, A. Tsang, OrfPredictor: predicting protein-coding regions in EST-derived sequences, *Nucleic Acids Res.* 33 (2005) W677-680.
- [30] D.M. Emms, S. Kelly, OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy, *Genome. Biol* 16 (2015) 157.
- [31] K. Katoh, D.M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability, *Mol. Biol. Evol.* 30 (2013) 772-780.
- [32] F. Abascal, R. Zardoya, M.J. Telford, TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations, *Nucleic Acids Res.* 38 (2010) W7-13.
- [33] G. Talavera, J. Castresana, Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments, *Syst. Biol* 56 (2007) 564-577.
- [34] S. Capella-Gutierrez, J.M. Silla-Martinez, T. Gabaldon, trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses, *Bioinformatics* 25 (2009) 1972-1973.
- [35] B. Misof, B. Meyer, B.M. von Reumont, P. Kuck, K. Misof, K. Meusemann, Selecting informative subsets of sparse supermatrices increases the chance to find correct trees, *BMC Bioinformatics* 14 (2013) 348.
- [36] R. Lanfear, B. Calcott, S.Y. Ho, S. Guindon, Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses, *Mol. Biol. Evol.* 29 (2012) 1695-1701.
- [37] R. Lanfear, B. Calcott, D. Kainer, C. Mayer, A. Stamatakis, Selecting optimal partitioning schemes for phylogenomic datasets, *BMC Evol. Biol.* 14 (2014) 82.
- [38] A. Stamatakis, RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models, *Bioinformatics* 22 (2006) 2688-2690.

- [39] F. Ronquist, M. Teslenko, P. van der Mark, D.L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M.A. Suchard, J.P. Huelsenbeck, MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space, *Syst. Biol.* 61 (2012) 539-542.
- [40] A. Rambaut, M.A. Suchard, D. Xie, A.J. Drummond, Tracer v1.6, 2014.
- [41] Z. Yang, PAML 4: phylogenetic analysis by maximum likelihood, *Mol. Biol. Evol.* 24 (2007) 1586-1591.
- [42] Z.H. Yang, R. Nielsen, N. Goldman, A.M.K. Pedersen, Codon-substitution models for heterogeneous selection pressure at amino acid sites, *Genetics* 155 (2000) 431-449.
- [43] J. Zhang, R. Nielsen, Z. Yang, Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level, *Mol. Biol. Evol.* 22 (2005) 2472-2479.
- [44] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate - a practical and powerful approach to multiple testing, *J Roy Stat Soc B Met* 57 (1995) 289-300.
- [45] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.* 25 (1997) 3389-3402.
- [46] S. Gotz, J.M. Garcia-Gomez, J. Terol, T.D. Williams, S.H. Nagaraj, M.J. Nueda, M. Robles, M. Talon, J. Dopazo, A. Conesa, High-throughput functional annotation and data mining with the Blast2GO suite, *Nucleic Acids Res.* 36 (2008) 3420-3435.
- [47] A. Mitchell, H.Y. Chang, L. Daugherty, M. Fraser, S. Hunter, R. Lopez, C. McAnulla, C. McMenamin, G. Nuka, S. Pesseat, A. Sangrador-Vegas, M. Scheremetjew, C. Rato, S.Y. Yong, A. Bateman, M. Punta, T.K. Attwood, C.J. Sigrist, N. Redaschi, C. Rivoire, I. Xenarios, D. Kahn, D. Guyot, P. Bork, I. Letunic, J. Gough, M. Oates, D. Haft, H. Huang, D.A. Natale, C.H. Wu, C. Orengo, I. Sillitoe, H. Mi, P.D. Thomas, R.D. Finn, The InterPro protein families database: the classification resource after 15 years, *Nucleic Acids Res.* 43 (2015) D213-221.
- [48] P. Jones, D. Binns, H.Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, G. Nuka, S. Pesseat, A.F. Quinn, A. Sangrador-Vegas, M. Scheremetjew, S.Y. Yong, R. Lopez, S. Hunter, InterProScan 5: genome-scale protein function classification, *Bioinformatics* 30 (2014) 1236-1240.
- [49] Y. Moriya, M. Itoh, S. Okuda, A.C. Yoshizawa, M. Kanehisa, KAAS: an automatic genome annotation and pathway reconstruction server, *Nucleic Acids Res.* 35 (2007) W182-185.
- [50] M. Schrödl, Time to say "Bye-bye Pulmonata"?, *Spixiana* 37 (2014) 161-164.
- [51] K.M. Jörger, I. Stöger, Y. Kano, H. Fukuda, T. Knebelberger, M. Schrödl, On the origin of Acochlidia and other enigmatic euthyneuran gastropods, with implications for the systematics of Heterobranchia, *BMC Evol. Biol.* 10 (2010) 323.
- [52] B. Dayrat, M. Conrad, S. Balayan, T.R. White, C. Albrecht, R. Golding, S.R. Gomes, M.G. Harasewych, A.M. de Frias Martins, Phylogenetic relationships and evolution of pulmonate gastropods (Mollusca): New insights from increased taxon sampling, *Mol. Phylogenet. Evol.* 59 (2011) 425-437.
- [53] Y. Kano, T.P. Neusser, H. Fukumori, K.M. Jörger, M. Schrödl, Sea-slug invasion of the land, *Biol. J. Linn. Soc.* 116 (2015) 253-259.
- [54] B. Dayrat, S. Tillier, Evolutionary relationships of euthyneuran gastropods (Mollusca): a cladistic re-evaluation of morphological characters, *Zool. J. Linn. Soc.* 135 (2002) 403-470.

- [55] T.R. White, M.M. Conrad, R. Tseng, S. Balayan, R. Golding, A.M. de Frias Martins, B.A. Dayrat, Ten new complete mitochondrial genomes of pulmonates (Mollusca: Gastropoda) and their impact on phylogenetic relationships, *BMC Evol. Biol.* 11 (2011) 295.
- [56] J.E. Morton, The evolution of the Ellobiidae with a discussion on the origin of the Pulmonata, *Proc. Zool. Soc. Lond.* 125 (1955) 127-168.
- [57] A. Heyland, D.A. Price, M. Bodnarova-Buganova, L.L. Moroz, Thyroid hormone metabolism and peroxidase function in two non-chordate animals, *J Exp Zool B Mol Dev Evol* 306 (2006) 551-566.
- [58] S.J. Crockford, Evolutionary roots of iodine and thyroid hormones in cell-cell signaling, *Integr. Comp. Biol* 49 (2009) 155-166.
- [59] J. Lee, K.-J. Rhee, S.H. Kim, Y.W. Kang, Y. Cho, S.-J. Yang, C.-H. Pan, Y.S. Kim, O6-Methylguanine-DNA methyltransferase (MGMT) gene expression is associated with ultraviolet B (UVB)-induced cell growth inhibition and recovery, *Genes & Genomics* 37 (2015) 789-796.
- [60] B.A. Bancroft, N.J. Baker, A.R. Blaustein, Effects of UVB radiation on marine and freshwater organisms: a synthesis through meta-analysis, *Ecol. Lett.* 10 (2007) 332-345.
- [61] M.H. Olson, N.E. Barbieri, Mechanisms of ultraviolet radiation tolerance in the freshwater snail *Physa acuta*, *Freshwater Science* 33 (2014) 66-72.

## Tables titles and legends

Table 1. Descriptive statistics of the assembled transcriptomes. *De novo* assemblies of new panpulmonate transcriptomes are highlighted in bold.

Clade	Species	Habitat	SRA Accession Number	Number of raw reads	Trimmed reads		MIRA meta-assembly		BlastX			
					For Trinity	For Bridger	All transcripts	Transcripts > 300 (bp)	N50 (bp)	Predicted ORF's	Single hits	Single genes
Acochlidia	<i>Microhedyale glandulifera</i>	Marine	SRX644682	6194970	6103287	6028350	129453	84468	742	83983	28750	12772
Acochlidia	<i>Strubellia wawrai</i>	Freshwater	SRX644701	24132673	24049385	23436737	82681	79947	1582	79735	27911	13735
Amphiboloidea	<i>Phallomedusa solida</i>	Intertidal	SRX644691	25685273	25496722	24822510	68633	65500	1424	64978	24560	12722
Ellobioidea	<b><i>Carychium</i> sp.</b>	Terrestrial	<b>SRX2069652</b>	<b>33608344</b>	<b>23461211</b>	<b>23461211</b>	<b>87994</b>	<b>87719</b>	<b>2035</b>	<b>87242</b>	<b>23586</b>	<b>11672</b>
Ellobioidea	<b><i>Cassidula plecotremata</i></b>	Intertidal	<b>SRX2069709</b>	<b>26316221</b>	<b>18928297</b>	<b>17989294</b>	<b>62318</b>	<b>62031</b>	<b>1533</b>	<b>61569</b>	<b>9583</b>	<b>5717</b>
Ellobioidea	<b><i>Melampus flavus</i></b>	Intertidal	<b>SRX2069588</b>	<b>21068142</b>	<b>16919851</b>	<b>16052613</b>	<b>97629</b>	<b>97334</b>	<b>1640</b>	<b>96816</b>	<b>22187</b>	<b>10778</b>
Ellobioidea	<i>Ophicardelus sulcatus</i>	Intertidal	SRX644688	16026272	15737826	15425623	74467	71073	1499	70573	20851	10937
Ellobioidea	<b><i>Pythia pachyodon</i></b>	Terrestrial	<b>SRX2069631</b>	<b>24016251</b>	<b>20555474</b>	<b>19399330</b>	<b>92548</b>	<b>92264</b>	<b>2035</b>	<b>91759</b>	<b>24077</b>	<b>11776</b>
Ellobioidea	<b><i>Trimusculus</i> sp.</b>	Intertidal	<b>SRX2069510</b>	<b>25613160</b>	<b>16343735</b>	<b>15696090</b>	<b>55728</b>	<b>55391</b>	<b>1240</b>	<b>55213</b>	<b>9310</b>	<b>5547</b>
Hygrophila	<i>Planorbarius comeus</i>	Freshwater	SRX482734	28040804	27417761	26569686	86067	82815	2587	78934	39051	16959
Hygrophila	<i>Biomphalaria glabrata</i>	Freshwater	SRX327185	172317158	170810180	164139487	70439	67493	2070	66689	33757	17015
Hygrophila	<i>Radix balthica</i>	Freshwater	SRX037214	16923850	NA	NA	54450	54450	679	30883	14614	8140
Pyramidelloidea	<i>Turbonilla</i> sp.	Marine	SRX644703	26619896	26219791	25806515	132978	127301	1023	126409	34180	14316
Stylommatophora	<b><i>Arion vulgaris</i></b>	Terrestrial	<b>SRX2069745</b>	<b>24874185</b>	<b>19179005</b>	<b>17977774</b>	<b>92316</b>	<b>91984</b>	<b>1454</b>	<b>91417</b>	<b>29208</b>	<b>13134</b>
Stylommatophora	<i>Onchidella floridana</i>	Intertidal	SRX644687	14872953	14797399	14528352	79502	74540	1057	74146	28257	13679

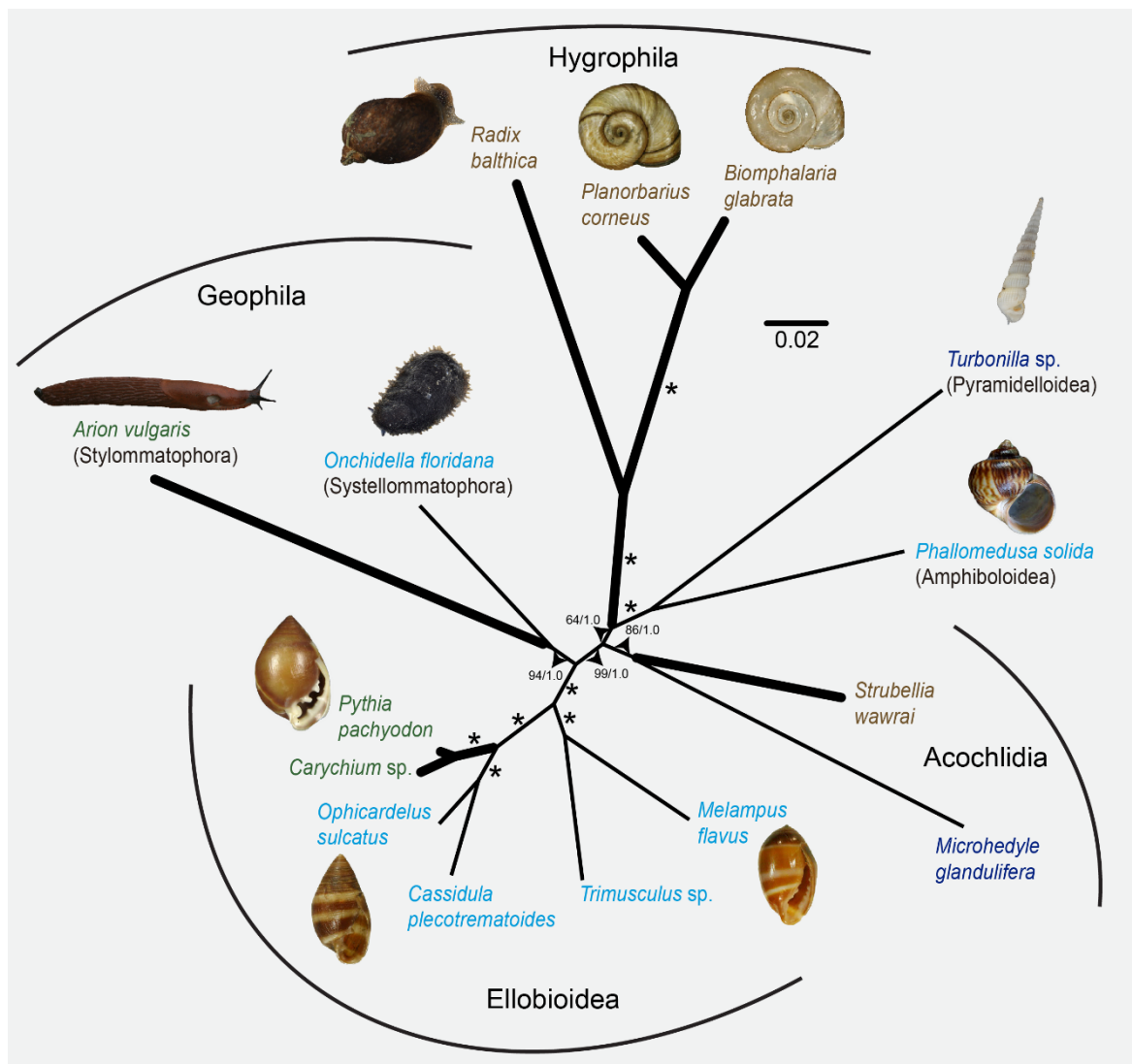
Table 2. Examples of ortholog clusters under positive selection in the terrestrial and freshwater lineages. The complete information can be found in the Supplementary Files 8 and 9 in Ref. [22].

Ortholog cluster	BlastX annotation	Molecular function	Biological process	KEGG pathway
<b>Terrestrial</b>				
OG0000060	tyramine beta-hydroxylase-like	Copper ion binding, oxidoreductase activity	Oxidation-reduction process	Tyrosine metabolism
OG0000137	sodium glucose co-transporter 4-like	Transmembrane transport	Transporter activity	Carbohydrate digestion and absorption
OG0001172	alpha-sarcomeric-like isoform X2	Actin filament binding, calcium ion binding	Actin crosslink formation, actin filament bundle assembly	Focal adhesion
<b>Freshwater</b>				
OG0000120	cytochrome P450 3A7-like	Monoxygenase activity, iron ion binding	Xenobiotic metabolic process	Aminobenzoate degradation, steroid hormone biosynthesis
OG0004116	methylated-DNA--cysteine methyltransferase-like isoform X2	methylated-DNA-[protein]-cysteine S methyltransferase activity	DNA repair	-
OG0004174	cytochrome c oxidase subunit 4 isoform mitochondrial-like	Cytochrome c oxidase activity	Proton transport, mitochondrial electron transport, cytochrome c to oxygen	Oxidative phosphorylation
OG0002708	40S ribosomal protein S3a	-	RNA binding, protein binding	rRNA processing, translation



## Figure title and legend

Figure 1. Unrooted phylogeny of Panpulmonata. A bootstrap value and posterior probability equal to 100/1.0 is represented with \*. Colors represent the habitat where species occur: Dark blue, marine; light blue, intertidal; brown, freshwater; green, terrestrial. Freshwater and terrestrial branches tested for positive selection are highlighted with a thick bar. Photo credits: Natural Museum Rotterdam (*B. glabrata* NMR 81004, *M. flavus* NMR 82913, *O. sulcatus* NMR7769, *P. corneus* NMR78762); J. Klawiter (*A. vulgaris*), CC-BY-3.0; Bishogai database (*P. pachyodon*); Moorea Biocode (*Turbonilla* sp.), CC-BY-NC-SA 3.0; A. Duppont (*O. floridana*); B. Feldmeyer (*R. balthica*); S. Grove (*P. solida*).



# CURRICULUM VITAE

## Personal details

Name: Pedro Eduardo Romero Condori

Address: Eschersheimer Landstrasse 248E, 60320 Frankfurt am Main

Email: quipu.romero@gmail.com

Date of birth: 16 December 1986

Place of birth: Lima, Peru

## Education

Since 2012. PhD-studies in Biological Sciences. Goethe University, Frankfurt am Main, Germany. Thesis: "Phylogeny and evolution in Panpulmonata (Gastropoda, Mollusca)", Goethe University. Supervisor: Prof. Dr. Markus Pfenninger.

2007-2010. Master in Molecular Biology. Universidad Nacional Mayor de San Marcos, Lima, Peru. Thesis: "Phylogeography of *Systrophia helicycloides*: How is the tropical rainforest dynamics reflected in land snail mitochondrial genes?". Supervisor: Prof. Dr. Rina Ramírez.

2008. *Título Profesional* in Biological Sciences with mention in Genetics and Cell Biology. Universidad Nacional Mayor de San Marcos, Lima, Peru. Thesis: "Genetic diversity and structure of *Bostryx scalariformis* based on 16S rRNA mitochondrial gene polymorphisms". Supervisor: Prof. Dr. Rina Ramírez.

2002-2006. Bachelor in Biological Sciences. Universidad Nacional Mayor de San Marcos, Lima, Peru.

## Research experience

Since 2013. Molecular Ecology Group. Senckenberg Biodiversity and Climate Research Centre. Frankfurt am Main, Germany. Supervisor: Prof. Dr. Markus Pfenninger.

- 2012-2013. Phylogeny and Systematics Group. Institute for Ecology, Evolution and Diversity. Goethe University. Frankfurt am Main, Germany. Supervisor: Prof. Dr. Annette Klussmann-Kolb.
2012. Fish Evolution Group. Institute of Evolutionary Sciences. University of Montpellier II. Montpellier, France. Supervisor: Dr. Christelle Tougard.
2011. Molecular Biology and Biotechnology Laboratory. Peruvian National Institute of Health (INS). Lima, Peru. Supervisor: Dr. Heinner Guio.
2010. Department of Terrestrial Zoology. Naturalis Biodiversity Center. Leiden, the Netherlands. Supervisor: Dr. Bram Breure.
2009. Laboratory of Genetics and Molecular Evolution of Birds. Institute of Biosciences. São Paulo University. São Paulo – Brazil. Supervisor: Prof. Dr. Cristina Miyaki.
- 2006 – 2012. Department of Malacology, Museum of Natural History and Laboratory of Molecular Systematics and Phylogeography, Faculty of Biology. San Marcos University. Lima, Peru. Supervisor: Prof. Dr. Rina Ramírez.

## **Publications**

- Romero PE**, Weigand AM, Pfenninger M. 2016. Positive selection on panpulmonate mitogenomes provide new clues on adaptations to terrestrial life. *BMC Evol Biol.* 2016;16(1):164.
- Romero PE**, Pfenninger M, Kano Y, Klussmann-Kolb A. Molecular phylogeny of the Ellobiidae (Gastropoda: Panpulmonata) supports independent terrestrial invasions. *Mol Phylogenet Evol.* 2016;97:43-54.
- Romero PE**, Breure ASH. Support and surprises: molecular phylogeny of the land snail superfamily Orthalicoidea using a three-locus gene analysis with a divergence time analysis and ancestral area reconstruction (Gastropoda: Stylommatophora). *Arch. Molluskenkunde* 2012;141(1):1-20.

- Ramírez R, Borda V, **Romero PE** et al. Biodiversity and endemism of the western Amazonia land snails *Megalobulimus* and *Systrophia*. Rev. Peru. Biol. 2012;19(1):59-74.
- Romero PE**, Ramirez R. Intraspecific divergence and DNA barcodes in *Systrophia helicycloides* (Gastropoda, Scolodontidae). Rev Peru Biol. 2011;18(2):201-8.
- Borda V, Ramírez R, **Romero PE**. Pediose gland in land snails and its evolutionary implications, with emphasis on *Megalobulimus*. Rev. Peru. Biol. 2010;17(1):43-52.
- Ramirez J, Ramírez R, **Romero PE** et al. Evolutionary position of Peruvian land snails (Orthalicidae) among Stylommatophora (Mollusca: Gastropoda). Rev. Peru. Biol. 2009;16(1):51-6.

### **Conference presentations**

- Romero PE et al. 2016. Molecular phylogeny of the Ellobiidae (Gastropoda: Panpulmonata) supports independent terrestrial invasions. World Congress of Malacology 2016. Penang, Malaysia.
- Romero PE. 2016. From sea to land and beyond: Mitochondrial evolution and habitat transitions in land snails. 1<sup>st</sup> SINAPSIS. Paris, France.
- Romero PE et al. 2015. The molecular phylogeny of Ellobiidae (Gastropoda: Panpulmonata) supports independent terrestrial invasions. Evolution 2015. Guaruja, Brazil.
- Romero PE et al. 2013. Towards a molecular phylogeny of Ellobiidae (Gastropoda, Pulmonata). World Congress of Malacology 2013. Azores, Portugal.
- Romero PE et al. 2012. A preliminary molecular phylogeny reveals population dynamics and divergence times in Peruvian Dengue virus serotypes 3 and 4. International Scientific Meeting /ECI. Lima, Peru.
- Romero PE et al. 2011. A new molecular phylogeny of the superfamily Orthalicoidae. XX ICBAR scientific meeting. Lima, Peru.

Romero PE et al. 2008. Genetic diversity and differentiation in the Peruvian coastal desert: The case of *Bostryx scalariformis* (Gastropoda. Bulimulidae). VII Latin American Congress of Malacology. Valdivia, Chile.

## **Teaching experience**

2012 Animal phylogeny and systematics. Goethe University Frankfurt am Main. Faculty of Biosciences. Germany

2006-2012. Molecular genetics, Bioinformatics, Phylogeography. Universidad Nacional Mayor de San Marcos. Faculty of Biology. Peru.

2011. Biology for engineers. Universidad Nacional de Ingeniería. Faculty of Mechanical Engineering. Peru.

2015, 2012. Organizer and lecturer in the workshop: "Bioinformatics tools to study biodiversity". Museum of Natural History. Universidad Nacional Mayor de San Marcos. Peru.

## **Scholarships**

2015. CONCYTEC. Peru. Scholarship to fund the final year of the PhD studies at Goethe University, Frankfurt am Main.

2008. CONCYTEC, Peru. Scholarship to fund the final year of the Master studies in Molecular Biology at UNMSM, Lima.