Parallel duplication and loss of aquaporin-coding genes during the “out of the sea” transition as potential key drivers of animal terrestrialization

Gemma I. Martínez-Redondo | Carolina Simón Guerrero | Leandro Aristide | Pau Balart-García | Vanina Tonzo | Rosa Fernández

Abstract
One of the most important physiological challenges animals had to overcome during terrestrialization (i.e., the transition from sea to land) was water loss, which alters their osmotic and hydric homeostasis. Aquaporins are a superfamily of membrane water transporters heavily involved in osmoregulatory processes. Their diversity and evolutionary dynamics in most animal lineages remain unknown, hampering our understanding of their role in marine–terrestrial transitions. Here, we interrogated aquaporin gene repertoire evolution across the main terrestrial animal lineages. We annotated aquaporin-coding genes in genomic data from 458 species from seven animal phyla where terrestrialization episodes occurred. We then explored aquaporin gene evolutionary dynamics to assess differences between terrestrial and aquatic species through phylogenomics and phylogenetic comparative methods. Our results revealed parallel aquaporin-coding gene duplications during the ecological transition from marine to nonmarine environments (e.g., brackish, freshwater and terrestrial), rather than from aquatic to terrestrial ones, with some notable duplications in ancient lineages. In contrast, we also recovered a significantly lower number of superaquaporin genes in terrestrial arthropods, suggesting that more efficient oxygen homeostasis in land arthropods might be linked to a reduction in this type of aquaporin. Our results thus indicate that aquaporin-coding gene duplication and loss might have been one of the key steps towards the evolution of osmoregulation across animals, facilitating the "out of the sea" transition and ultimately the colonization of land.

KEYWORDS
evolution, genomics/proteomics, invertebrates, macroevolution, molecular, transcriptomics
1 | INTRODUCTION

The ecological transitions from water to land (i.e., terrestrialization) are one of the most remarkable events shaping the evolution of biodiversity as we know it today, with 77 and 93% of extant animal and plant species, respectively, inhabiting terrestrial habitats (Román-Palacios et al., 2022). Animals have successfully terrestrialized independently in nine different phyla: chordates, nematodes, tardigrades, arthropods, onychophorans, platyhelminthes, molluscs, annelids, and nemerteans (Figure S1). Within most phyla, more than one independent habitat shift has occurred, ranging from only one in the onychophorans to more than 30 in nematodes (Holterman et al., 2019), at different times along Earth’s history. The most studied terrestrialization event is, beyond doubt, the one that gave rise to terrestrial tetrapods during the Devonian period from euryhaline aquatic ancestors (Goedert et al., 2018; Qvarnström et al., 2018). Nevertheless, the most ancient terrestrialization events occurred in arthropods (arachnids, myriapods and hexapods) from marine ancestors during the late Cambrian-early Ordovician (Rota-Stabelli et al., 2013). More than four additional independent terrestrialization events occurred in the following millions of years within Pancrustacea (isopods, amphipods and anomuran and brachyuran decapods) (Sharma, 2017; Watson-Zink, 2021). Other terrestrialization events that have recently gained interest are the ones in molluscs and annelids. In the case of molluscs, gastropods have colonized land more than nine times (Kameda & Kato, 2011), with the latest estimates suggesting up to 30 times (Vermeij & Watson-Zink, 2022), either following a freshwater route or through marginal environments such as brackish water (Krug et al., 2022; van Straalen, 2021; Vermeij & Watson-Zink, 2022). Among them, the most successful terrestrial gastropod lineage is the Stylommatophora, an order of air-breathing pulmonary land gastropods that contain most living terrestrial snails and slugs. Regarding annelids, earthworms (Crassiclitellata) derive from a single terrestrialization event, while other events have occurred within Hirudinida and Enchytraeidae (Érséus et al., 2020; van Straalen, 2021). Another example is that of onychophorans, or velvet worms, which are the only entirely terrestrial animal phylum. This lineage closely related to arthropods has been hypothesised to have a marine origin based on the similar osmotic pressure of their haemolymph to terrestrial arthropods (Little, 1990). Together, these independent colonization events (Figure 1a) make animal terrestrialization a complex and intriguing area of study, especially since the genomic underpinnings facilitating these transitions remain highly unexplored.

Terrestrialization required many complex physiological changes in different biological processes, such as respiration and locomotion or reproduction, that allowed to overcome the environmental barriers and adapt to life on land. One of the most important challenges species faced during this ecological transition was avoiding water loss, which alters their osmotic and hydric homeostasis. Most marine animals are osmoconformers, which means that they cannot regulate their internal osmotic pressure under environmental shifts. When marine animals transitioned to euryhaline or freshwater habitats, they faced osmotic stress due to the differences in the ionic composition between their extracellular fluids and the environment. The development of the ability to osmoregulate, that is, to maintain the osmotic pressure of their extracellular fluids (Rankin & Davenport, 1981), allowed them to overcome this osmotic shock. Accordingly, terrestrial animals have acquired several physiological adaptive mechanisms to regulate their internal osmotic balance and water loss.

At the molecular level, animals can modify the osmotic composition of their fluids by regulating the transport of ions or water through biological membranes (Takvam et al., 2021; Uchiyama & Konno, 2006). Previous experiments have shown that cells of some osmoregulatory organs express pore-forming proteins called aquaporins in their membranes (Cabero et al., 2020; Finn et al., 2014; Kaufmann et al., 2005; Spring et al., 2009; Wittekindt & Dietl, 2019; Yi et al., 2011). Aquaporins are a superfamily of small membrane intrinsic proteins (24–36 kDa) that are heavily involved in osmoregulatory processes by forming pores in the cell membrane that facilitate the transport of water and other small molecules (Abascal et al., 2014; Kruse et al., 2006). Structurally, the basic units of aquaporins are monomers characterized by six transmembrane domains that surround a narrow pore, and two selectivity filters: two Asn-Pro-Ala (NPA) motifs, responsible for proton exclusion, and an aromatic/arginine (ar/R) region that determines the pore size (Figure 1b). Variations in the amino acid sequences of these motifs can slightly change pore size, structure and hydrophobicity, and ultimately determine different subtypes of aquaporins and their ability to transport molecules other than water (Figure 1c) (de Maré et al., 2020). This general structure is conserved between bacteria and eukaryotic aquaporins, as well as between different subtypes of aquaporins, despite the low sequence similarity of these proteins (<30%) (Kruse et al., 2006).

Initial phylogenetic analyses of aquaporins across all organisms, together with the permeability differences of aquaporins, allowed their classification into two main subtypes: aquaporins that transport water (Aqps), and aquaglyceroporins that are also permeable to glycerol and other molecules (Gips) (Park & Saier Jr., 1996; Zardoya & Villalba, 2001). Further studies showed that both are also permeable to other molecules (CO₂, H₂O₂ and ammonia, among others) (Finn & Cerdà, 2015). In animals, Aqps can be further divided into three subtypes: classical aquaporins, aquaammoniapiorins, and superaquaporins (names that we will use hereafter instead of unorthodox aquaporin subtypes). Each of these subtypes contains highly conserved motifs (Figure 1d) (Abascal et al., 2014).

Previous studies have analysed the repertoire of aquaporin-coding genes (Aqp-coding genes hereafter) in selected animal species, for example, Danio rerio (Tingaud-Sequeira et al., 2010), Caenorhabditis elegans (Huang et al., 2007), Milnesium tardigradum (Grohne et al., 2013), blood-feeding arthropods (Benoit et al., 2014), dust mites (Peng et al., 2018) or the salmon louse and its host (Stavang et al., 2015), to name a few. Meanwhile, others have focused on the diversity and evolution of animal aquaporins in groups of human interest, such as vertebrate (Finn et al., 2014) and hemipteron
aquaporins (Van Ekert et al., 2016), or insect entomoglyceroporins (Finn et al., 2015). Recent studies have leveraged newly available transcriptomic and genomic resources to expand taxon sampling and dig deeper into aquaporin (or aquaglyceroporin) diversity in groups of animals that are not widely studied, such as copepods and crustaceans (Catalán-García et al., 2021), gastropods (Colgan & Santos, 2018), annelids (Mucciolo et al., 2021), and nonvertebrate deuterostomes (Yilmaz et al., 2020). However, the aquaporin gene repertoire in other animal groups, such as molluscs and less studied arthropod orders (i.e., chelicerates other than spiders, scorpions and mites), has not yet been explored, let alone in a comparative framework within the context of animal terrestrialization. Only a couple of previous studies have explored the potential key role of aquaporin gene duplications in the adaptation of vertebrates to life on land, with the identification of three tetrapod-exclusive classical aquaporins functionally related to water absorption mechanisms (Finn et al., 2014), and some signals of positive selection in semi-terrestrial mudskipper aquaporins (Lorente-Martínez et al., 2018). Nevertheless, even though the importance of gene duplication in originating new functions and driving adaptation has long been reported (Kondrashov, 2012), the potential role of aquaporin gene duplication and loss in the terrestrialization of lineages other than tetrapods has not been explored yet.

Here, we test the hypothesis that gene duplication and loss in Aqp-coding genes may have played a key role in animal terrestrialization and suggest that they may have facilitated marine-to-land habitat transition
in other nonchordate phyla. To show this, we mined and annotated more than 6300 Aqp-coding gene sequences from 458 genomic and transcriptomic data sets, maximizing taxon sampling to ensure a good representation of most animal lineages with terrestrial representatives. To assess whether aquaporins may have had a role during animal terrestrialization, we explored aquaporin gene evolutionary dynamics to assess differences between terrestrial and aquatic species through phylogenomics and phylogenetic comparative methods. We interpret and discuss our results in the context of animal terrestrialization.

2 | MATERIALS AND METHODS

2.1 | Obtaining genomic and transcriptomic data and habitat assignment

We analysed the aquaporin gene repertoire of 485 species from seven animal phyla where there has been at least one ecological transition to land: Arthropoda, Onychophora, Tardigrada, Nematoda, Mollusca, Annelida and Vertebrata (Figure 1a, Figure S1). The full list of species included and the sources from which they were obtained is available as Data S1. Aquaporins from vertebrates, M. tardigradum, and C. elegans were directly obtained from the literature (Finn et al., 2014; Grohme et al., 2013; Huang et al., 2007, respectively). The only exceptions were the aquaporins from the amphibian Leptobrachium leishanense, for which we downloaded the genome annotation (Cunningham et al., 2021). In the case of annelids, Mucciolo et al. (2021) provided a fairly good representation of aquaporins in Pleistoannelida, the most abundant clade within annelids. However, as noted by the authors, the methodology followed could have underestimated the identification of aquaporins due to the incompleteness or low quality of data within the source database. For instance, an example of this is Eisenia andrei, for which only one Aqp was identified. To verify this information, we downloaded the genome annotation of E. andrei (Shao et al., 2020) to annotate its Aqp-coding genes ourselves as described below. We identified a total of 27 Aqp-coding genes, in contrast with the one reported by Mucciolo et al. (2021). Data from the remaining 415 invertebrate species were obtained from MATEdb (Fernández et al., 2022), which provides standard comparable proteomes for phylogenomic studies. We downloaded the gene annotation of those species (longest isoform for each gene), and their taxonomic and habitat information. Habitat information for species was obtained from the World Register of Marine Species (WoRMS) database (WoRMS Editorial Board, 2022). A species was considered to be aquatic if it spent at least one stage of its life cycle in an aquatic environment.

2.2 | Annotating genes encoding for aquaporins in downloaded proteomes

From the whole set of proteomes, genes encoding for aquaporins were annotated using BITACORA version 1.3 (Vizueta et al., 2020) in "protein mode". Briefly, BITACORA combines BLAST and HMMER (Eddy, 2011) searches against a custom curated database of genes belonging to the gene family of interest. Due to the conservation of aquaporins across all animals, our aquaporin curated database contained a list of 123 metazoan aquaporin sequences that have been experimentally confirmed in previous studies (Data S2). To benchmark this methodology, we used Drosophila melanogaster to compare the number of aquaporins identified using BITACORA with the number of aquaporins annotated in its genome, obtaining the same result. In addition, it should be noted that the annotation of aquaporins is dependent on the data used and their absence in a species does not mean that it may not be actually present in its genome. We have used as many representative species of the same clade as possible to try to reduce this bias.

2.3 | Aquaporin-coding genes nomenclature

Aquaporin-coding gene nomenclature between animal phyla is challenging due to the lack of consistency across different studies (Benoit et al., 2014; Kosicka et al., 2016, 2020; Mucciolo et al., 2021; Stavang et al., 2015; Yilmaz et al., 2020). This conflict worsens when considering papers on the identification of aquaporins of isolated species without taking their evolutionary history into account (Grohme et al., 2013; Huang et al., 2007). Here, we follow the system used by Mucciolo et al. (2021) and Kosicka et al. (2020) but taking into account the gene nomenclature rules, and name each of the aquaporin subtypes as “AqpX-like”, X being the lowest nonzero number of the vertebrate aquaporin-coding-gene that forms part of it. Thus, we define the four subtypes as Aqp3-like (aquaglyceroporins), Aqp1-like (classical aquaporins), Aqp8-like (aquaanmoniaporins) and Aqp11-like (superaquaporinss), and this nomenclature is followed hereafter.

2.4 | Phylogenetic tree inference of aquaporin-coding genes

To identify which aquaporins belonged to which subtype (i.e., Aqp1-like, Aqp3-like, Aqp8-like or Aqp11-like) for posterior analyses, we first independently classified aquaporins of all animal species included in this study, except for the ones that had already been classified in previous publications (Finn et al., 2014; Mucciolo et al., 2021). We based this classification on the description of the aquaporin phylogenetic clusters and their relative position in the phylogenetic tree with respect to the already classified sequences, which were considered as a backbone.

For arthropods, onychophorans and molluscs (i.e., lineages whose aquaporins we identified de novo in this study) we manually curated the classification by checking for unclustered aquaporin sequences and long branches that could represent outliers using BLASTP against the nonredundant NCBI database. This was done to reclassify identified aquaporins that had mistakenly fallen into the
wrong subtype and remove contaminants from species outside of the studied phylum that had escaped filtering in the original data sets. In total, 41 arthropod and mollusc sequences were removed from the initial set.

To build all phylogenetic trees, amino acid sequences were aligned using PASTA version 1.8.6 (Mirarab et al., 2015). The alignments included 202 sequences previously described as aquaporins (Finn et al., 2014, 2015) to use as a backbone for identifying aquaporin subfamilies in our data set (Data S3). In addition, two previously identified malaglyceroporins and one malacoaquaporin (Kosicka et al., 2016) were also included when building aquaporin phylogenetic trees. All resultant alignments were trimmed using trimAL version 1.4.1 (Capella-Gutiérrez et al., 2009). We tested different trimming methods and gap thresholds to select the best one in terms of keeping fewer alignment positions to reduce computational time while keeping biologically informative sites (Figure S2). Final alignments were trimmed with a gap threshold of 0.01.

We created guiding trees for each of the alignments using the LG model in FastTree version 2.1.11 (Price et al., 2010), which were later used as starting trees for the ML inference of aquaporin subfamilies in our data set (Data S3). We applied an LG protein matrix with a profile mixture model of 10 classes (C10), plus an additional class of empirical amino acid profile (F), accounting for across-site compositional heterogeneity using a discrete Gamma model with 4 categories (G). We used this mixture model as it does not involve a very high number of parameters to avoid overfitting. We also included ultrafast bootstrap (UFBoot) (Minh et al., 2013) and SH-aLRT test (Guindon et al., 2010) to obtain support values of resultant trees. The AqpM sequence of Methanococcus maripaludis (NCBI accession NP_988083), included in the backbone, was always used as an outgroup for tree rooting, as suggested in previous studies (Finn et al., 2014). To avoid long branch attraction (LBA) due to their higher divergence and because our preliminary results using all aquaporins showed that they formed a monophyletic clade, entomoglyceroporins were analysed separately and excluded from the phylogeny that combined all animal Aqp1-like sequences. We visualized the trees using iTOL (Letunic & Bork, 2021) and created the final figures with the R packages ggtree (Yu et al., 2017) and ggtreeExtra (Yu et al., 2021). A full list of packages used is included in the code available on GitHub (see Data Availability section).

2.5 Inference of a combined dated species tree

To obtain the phylogenetic relatedness information required for the phylogenetic comparative methods, we combined the time-scaled ultrametric species trees for each phylum. Vertebrate and onychophoran time trees were obtained from the literature (Baker et al., 2021; Irisarri et al., 2017) and pruned to keep only a selection of 13 vertebrate species whose aquaporins have previously been studied (Finn et al., 2014) and the four onychophorans (or closely-related species) that were included in MATEdb (Fernández et al., 2022), respectively. For the rest of the phyla, arthropods, molluscs and annelids, time-scaled ultrametric trees were inferred independently by using a node-dating approach on species trees that followed the most recent topologies (Anderson & Lindgren, 2021; Ballesteros et al., 2022; Colgan et al., 2007; Cunha & Giribet, 2019; Doğan et al., 2020; Erséus et al., 2020; Fernández et al., 2018; González et al., 2015; Johnson et al., 2018; Kallal et al., 2021; Kawahara et al., 2019; Kocot et al., 2011, 2013, 2020, 2017, 2019; Lee et al., 2019; Lindgren, 2010; Lozano-Fernandez et al., 2019; McKenna et al., 2019; Misof et al., 2014; Moles & Giribet, 2021; Ontano et al., 2021; Pabst & Kocot, 2018; Peters et al., 2017; Razkin et al., 2015; Richter, 2001; Salvi & Mariottini, 2016; Schwentner et al., 2018; Smith, 2021; Smith et al., 2011; Sun et al., 2020; Uribe et al., 2016; Uribe & Zardoya, 2017; Wang et al., 2017; Wiegmann et al., 2011; Yang et al., 2018). In the case of annelids, we dated the species tree using the Aqp1-like gene sequences in MCMCtree (Yang, 2007). We first inferred the annelid Aqp1-like gene tree using aquaporins annotated in Mucciolio et al. (2021) and our inferred aquaporins for E. andrei, as our preliminary analyses showed that the aquaporin phylogeny is broadly concordant with the species tree topology at the level of aquaporin subclusters. Second, from this gene tree, we derived 16 sets of 1:1 orthologues with at least four species represented using PhyloPyPruner (Thalen, 2019), considering the phylogenetic interrelationships of annelid species as described in Erséus et al. (2020). These orthologous sequences were then concatenated into a single matrix of 2694 sites for their analysis and split into two partitions based on their evolutionary rate (measured as the sum of branch lengths divided by the number of tips in the respective gene tree). Prior distributions for node ages (see Appendix S1 in Github) were based on Chen et al. (2020) for the annelid MRCA and in the divergence time intervals and fossil calibrations in Erséus et al. (2020). The species Romanchella perrieri was removed by PhyloPyPruner and was therefore not included in the tree. Dating analysis of arthropods and molluscs was based on 50 genes each, selected from a set of hundreds of phylogenetically informative genes. Full details are provided elsewhere (Aristide & Fernández, 2022; Methods S1). Finally, the ultrametric trees of all five phyla were drafted into a manually-built backbone topology depicting the relationships among them using the RRphylo R package (Castiglione et al., 2018). Divergence times for the ancestral nodes in the backbone tree were obtained from Dohrmann and Wörheide (2017).

2.6 Phylogenetic comparative methods to explore aquaporin evolution

In order to test whether there were differences in the mean total number of aquaporins between terrestrial and aquatic species, we implemented a phylogenetic generalized linear model using the phyloglm R package (Ho & Ané, 2014). Briefly, we fitted a phylogenetic Poisson regression model to our data describing the relationship between the number of aquaporins and habitat, accounting for the phylogenetic history of the species using the combined ultrametric species tree. For habitat, we classified species in two different ways:
marine versus nonmarine (i.e., freshwater, brackish and terrestrial), and terrestrial versus nonterrestrial (i.e., freshwater, brackish and marine). This classification allowed us to identify shifts in aquaporin number during marine to freshwater and terrestrial transitions, and marine and freshwater to terrestrial transitions. We tested for differences in each of the four main subtypes of aquaporins and in all aquaporins together.

3 | RESULTS

3.1 | An expanded catalogue of aquaporins in arthropods, molluscs and onychophorans unveils the complex dynamics of aquaporin evolution

Our knowledge of animal aquaporin evolution is patchy. To address this limitation and achieve a better understanding of their potential involvement in terrestrialization, we first explored the aquaporin gene repertoire in arthropods, molluscs and onychophorans, since we had an extensive taxon sampling for each phylum. We identified a total of 2303 arthropodan, 1225 molluscan, and 52 onychophoran Aqps (Data S4) after annotating 310 arthropodan, 101 molluscan and four onychophoran genomic data sets. Only some Aqps from a few arthropod species and even fewer molluscs species had been previously annotated as aquaporins (Catalán-García et al., 2021; Colgan & Santos, 2018; Finn et al., 2015; Stavang et al., 2015). We also report here the first identification of Aqps in the phylum Onychophora.

As we wanted to include additional phyla to have a better understanding of the evolution of Aqp-coding genes in animals that have transitioned from the sea to the land, we searched the literature for additional aquaporin sequences in representatives from other phyla that underwent terrestrialization episodes (Nematoda, Tardigrada, Craniata and Annelida), obtaining a total of 4230 aquaporins. Aquaporins of C. elegans (Huang et al., 2007) and M. tardigradum (Grohme et al., 2013) were classified into the four main subtypes by building a phylogenetic tree with the backbone sequences. For annelids and vertebrates, aquaporins had already been classified into the main groups, which were later confirmed in our phylogeny.

Our inferred phylogeny of all animal aquaporins (Figures S3 and S4) supports the four aquaporin clades previously defined in the literature (Abascal et al., 2014): aquaglyceroporins (Aqp3-like), classical aquaporins (Aqp1-like), aquaammoniaporins (Aqp8-like) and superaquaporins (Aqp11-like). Moreover, the loss of Aqp8-like in arthropods and Aqp1-like in C. elegans is recovered, as previously reported (Finn & Cerdà, 2015; Stavang et al., 2015), together with the identification of Aqp8-like loss in onychophorans, suggesting that the loss of Aqp8 may have occurred before the split between Onychophora and Arthropoda. Hexapod-specific glyceroporins (i.e., entomoglyceroporins or Egps) clustered within arthropod Aqp1-like, showing their Prip-like origin as previously reported (Finn et al., 2015).

3.2 | Phylogenetic comparative methods unveil parallel aquaporin-coding gene duplications in nonmarine animals and gene loss in terrestrial animals

The presence of Aap clades specific to terrestrial animals, such as the Egps in insects (Finn et al., 2015) or AQP2, AQP5 and AQP6 in tetrapods (Finn et al., 2014; Lorente-Martínez et al., 2018), suggests that the number of aquaporins found in terrestrial animals might be different than the number found in aquatic ones. To test whether there is a correlation between the number of aquaporins and the habitat in which extant species live, we obtained Aqp-coding genes identified in each species and performed a phylogenetic Poisson regression using the habitat as an independent variable. We tested two different hypotheses using this model. The first one is that the number of Aqp-coding genes in terrestrial species is different to that of aquatic ones, supporting that a copy number variation of Aqp-coding genes through gene gain, duplication or loss may have facilitated terrestrialization during aquatic-terrestrial transitions. The second hypothesis tested is that the number of Aqp-coding genes in marine species is different to that in nonmarine species.

We first explored two independent tests with genomic data of the five phyla (Onychophora, Craniata (Vertebrata), Mollusca, Annelida and Arthropoda; five-phyla data set hereafter; Figure 2a). Tardigrades and nematodes were not included as we only analysed one species per phylum and lacked the representation we had in the others, which could bias the results. Moreover, since we had a dense taxon representation of most main lineages in arthropods and many molluscs, we performed the tests independently for each of these two phyla as well. These tests were explored both taking into account all Aqp-coding genes and in each type of Aqp-coding genes independently, following the main four types described above (i.e., Aqp1-like, Aqp3-like, Aqp8-like and Aqp11-like).

Overall, our results showed statistical differences in the number of several aquaporin-coding gene subfamilies when we compared marine and nonmarine species, pointing to a significantly higher number in nonmarine species than in marine ones in most cases (Figure 2b, c). First, in the analyses of the five-phyla data set we observed that nonmarine animals have a significantly higher number of Aqp3-like compared to marine species (Figure 2b, left). A new nomenclature for Aqp3-like subclades in the different animal phyla and the implications of its phylogenetic history in the context of animal terrestrialization is discussed in detail below (see also Table 1).

Regarding the phylum-specific analyses, we recovered in molluscs a significantly higher number of Aqp-coding genes in nonmarine compared to marine species (Figure 2c). Independent analyses for each type of Aqp-coding gene revealed a significantly higher number in nonmarine molluscs of Aqp1-like and Aqp8-like (Figure 2b, right). A description of the phylogenetic relationships of Aqp1-like and Aqp8-like as well as their potential role in mollusc terrestrialization is discussed below. In addition, our results indicated that terrestrial arthropods have a significantly lower mean number of Aqp11-like aquaporins than aquatic ones (Figure 2b, centre).
Phylogenetic relationships among Aqp11-like aquaporins are represented in Figure S5, showing some lineage-specific duplications in secondarily aquatic species.

3.3 | Ancient independent duplications in Aqp-coding genes are generally correlated with the “out of the sea” transition across animal phyla

3.3.1 | Classical aquaporins (Aqp1-like)

The phylogeny of protostome classical aquaporins (Figure 3, Figure S6) reveals a complex pattern of gains and losses in the different aquaporin-coding gene clusters, with many originating early during animal evolution. We highlight two clades of Aqp1-like, that we designate as Aqp1-bila1 and Aqp1-bila2, and that include the majority of Aqp1-like animal sequences (except for Aqp1-ann1, Prip and Drip). Since deuterostomes (mostly vertebrates but also a tunicate and a sea urchin, Figure 3, in turquoise) cluster inside Aqp1-bila2, while being absent in Aqp1-bila1, we can infer that the duplication that gave rise to both Aqp1-bila clades could have occurred at the base of Bilateria, before the diversification into Deuterostomia and Protostomia, followed by the loss of Aqp1-bila1 in the lineage that gave rise to deuterostomes. Both Aqp1-bila copies were further duplicated or lost independently across protostome evolution. Aqp1-bila1 contains mainly genes from species within Spiralia, together with onychophorans and tardigrades, while Aqp1-bila2 includes sequences from all animal phyla included in this study, except for nematodes, for which a lack of Aqp1-like sequences had previously been reported (Finn & Cerdà, 2015). It is worth mentioning that phyla that have the most profound adaptations to life on land (i.e., chordates and arthropods) underwent the independent loss of Aqp1-bila1 and only retained Aqp1-bila2.

Classical aquaporins were also independently duplicated in terrestrial lineages (Figure 2b, c). Onychophorans, the only fully-terrestrial animal phylum, have three clusters of Aqp1-like sequences
| Phylum         | Aqp1-like Evidence | Aqp3-like Evidence | Aqp8-like Evidence | Aqp11-like Evidence | Putative role in terrestrialization | Putative role in terrestrialization | Putative role in terrestrialization | Putative role in terrestrialization |
|---------------|-------------------|-------------------|--------------------|---------------------|-------------------------------------|--------------------------------------|--------------------------------------|-------------------------------------|
| Arthropoda    | Independent duplications | Increased efficiency in excretion | Ancestral duplication and differential retention | Water absorption in intestine | – | Reduction in terrestrial species | Redox homeostasis in ER |
| Onychophora   | Expansion          | Unknown            | Ancient duplication and independent lineage-specific expansions | Resistance to desiccation | – | – | – |
| Tardigrada    | –                  | –                  | Anhydrobiosis      | –                   | – | – | – |
| Annelida      | Duplication in Clitellata and loss in secondarily marine species | Sea-to-freshwater transition that predated terrestrialization of Enchytreida and Crassiclitellata | Duplication in Crassiclitellata | Accumulation of polyols in cocoons when exposed to cold or drought | – | – | – |
| Mollusca      | Evolution of glycerol transporters and duplication in Stylommatophora | Aestivation, Water absorption in the intestine | Duplication in Caenogastropoda | Sea-to-freshwater transition that predated terrestrialization of Pomatidae | Expansion in Panpulmonata | Aestivation, redox stress | – |
| Chordata\(^{a}\) | Tetrapod-specific aquaporins | Evolution of water conservation mechanisms | Positive selection in mudskippers | Gill function\(^{c}\) | – | Positive selection in mudskippers | Redox homeostasis in ER |

\(^{a}\) Information from the phylum Nematoda is not discussed as only one terrestrial species was included in our analyses.

\(^{b}\) Chordate results are proposed by previous studies (Finn et al., 2014; Lorente-Martínez et al., 2018).

\(^{c}\) The putative function of positively selected Aqp10a (Aqp3-like) in mudskippers' gills is proposed here based on the differential gene expression of its orthologue Aqp10a and paralogue Aqp10b in zebrafish gills (Finn & Cerdà, 2011), and the enlarged gill chambers mudskippers have whose function is related to their semiterrestrial lifestyle (i.e., they close the gill chambers when exposed to air to keep gills moist). Nevertheless, no functional studies have looked at Aqps in mudskippers' gills.
FIGURE 3 Maximum likelihood phylogenetic tree of 2202 Aqp1-like plus 63 backbone sequences in six studied phyla. Duplications are marked with a green dot. Bootstrap values are included only in nodes with support values higher than 70, and represented in a grey scale in the central phylogeny. (a) Both copies of malacoaquaporins (Maqp) in Panpulmonata. (b) Prip-like and Drip clades in arthropods. (c) Malaglyceroporins' duplication in Panpulmonata. (d) Aqp1-ann2 expansion within Crassicillettata. (e) Aqp1-ann6 clade expansion in Crassicillettata. (f) Aqp1-mol1 duplication in Stylommatophora. (g) Aqp1-ann4 duplication in Clitellata.
(Figure 3, in purple). One of them, Aqp1-onych2, harbours a gene duplication at the level of phylum, followed by several duplications and gene losses. In addition, our data show a not fully resolved clade containing the insect-exclusive Drip and arthropod Prip-like clusters (Figure 3b). It appears that there were a series of expansions in the arthropod ancestor followed by a differential loss and independent duplications in the different arthropod lineages. The Prip-like paphropodan, symphylan and chilopodan myriapod classes retain two ancestral copies, while the origin of the second copy in diplopod myriapods that retain more than one is more recent. Although previous studies have suggested that there has been an ancient whole-genome duplication (WGD) in Arachnopulmonata (Schwager et al., 2017), the Prip-like duplications that we recover do not suggest that as the origin of the Arachnopulmonata clades. On the contrary, the high clustering of all xiphosuran Prip-like duplicate sequences is consistent with them being ohnologues (duplicated genes originated by a WGD event and named after Susumu Ohno by Wolfe, 2000) that originated after the three rounds of WGD suggested by Nong et al. (2021). Nevertheless, an independent phylogenetic tree we inferred using only arthropod sequences recovered a different topology, with hexapod Drip (including the previously-mentioned Drip insect-exclusive cluster and some additional early-splitting hexapod sequences) and Prip-like having emerged after the divergence of Pancrustacea from Myriapoda. However, this topology suggesting independent duplications in each of the three terrestrial arthropod groups is not supported by bootstrapping, and therefore these interrelationships should be treated with caution (Figure S7). In addition to Drip and Prip-like clades, it is noteworthy that the Aqp1-arth cluster (Figure 3) contains chelicerate and myriapod sequences, while it lacks pancrustacean sequences (except for some sequences from Malacostraca, Ostracoda, Cirripedia and a couple of earliest insect-splitting species – Figure S6).

Regarding annelids, we observe six different clusters named Aqp1-ann1 to Aqp1-ann6 (Figure 3, in yellow). Among them are the Aqp1-ann6, Aqp1-ann4 and Aqp1-ann2 clusters (Figure 3e, g, d, respectively). Aqp1-ann6 include a nonsupported ancient duplication followed by a complex pattern of gene gain and losses (Figure 3e). We highlight the wide expansions in terrestrial Crassiclitellata (i.e., earthworms), putatively related to their terrestrialization (Figure 3e). In addition, we observe a duplication of Aqp1-ann4 in Clitellata with a subsequent loss of one copy in secondarily marine clitellates (Olavius algarensis), as shown in Figure 3g. Regarding Aqp1-ann2 (Figure 3d), there are two subsequent duplications at the level of Clitellata followed by the loss of all copies but one in non-Crassiclitellata annelids.

In molluscs, we find six supported clades (Figure 3, in indigo), which we have named Aqp1-mol1-2, Aqp1-mol4-5, Maqp and Mglp (Malacoaquaporins and Malacoglyceroporins, respectively, in accordance with previous research on the stylommatophoran snail Helix pomatia; Kosicka et al., 2016). As mentioned before, gastropods are the only molluscs that have colonized land. When looking at the mollusc-specific classical aquaporin clusters, we find three groups of special relevance. The first one is a Stylommatophora-specific copy within Aqp1-mol1. The duplication that gave rise to this copy dates back to before the origin of Panpulmonata (Tentaculata), with one copy being lost in all non-Stylommatophoran lineages (Figure 3f). Kosicka et al. (2016) identified the two duplicates of this clade in H. pomatia: HpAqp1 and HpAqp2. The second one is the Aqp1-mol2 clade, where we report an expansion in terrestrial Stylommatophora, freshwater panpulmonates, and the freshwater bivalves Margaritifera margaritifera and Cristaria plicata (Figure 3, Figure S6; MMAR1 and CPLI1, respectively). The last one is an unnamed subclade within Maqp within Maqp (Figure 3a), where an additional duplication occurred before the split of Gastropoda and Scaphopoda, with only one copy being retained in Scaphopoda, and two within panpulmonate gastropods. The other gastropods’ Maqps are not clustered within this clade.

### 3.3.2 | Aquamnioporins (Aqp8-like)

We identified several lineage-specific duplications of Aqp8-like in C. elegans, annelids and molluscs (Figure S8). In the case of annelids both terrestrial and aquatic species show duplications within this aquaporin subtype. However, not all molluscan lineages showed duplications. For instance, gastropods contain only one copy of Aqp8-like, except for the terrestrial gastropod Pomatias elegans (PELE1). In addition, while marine bivalves contain only one or two copies, freshwater bivalves contain three or even four. These two examples point to the putative role of these expansions in their transition from the marine environment.

### 3.3.3 | Aquaglycerolporins (Aqp3-like)

Compared to the phylogeny of Aqp1-like (Figure 3), where deuterostome aquaporins were nested within protostome aquaporins showing a shared origin, deuterostome and protostome Aqp3-like form two separate clades with independent evolutionary histories (Figure 4, Figure S9). Despite using a different methodology, our results for vertebrate Aqp3-like proteins are mostly consistent with previous studies (Finn et al., 2014), except for the recovery of AQP10 and AQP7 as sister clades. Within protostomes, there is a further split into two clades, which we named Aqp3-prot1 and Aqp3-prot2. Aqp3-prot1 includes representatives of all invertebrate phyla included in this study. On the contrary, Aqp3-prot2 is mostly composed of arthropod sequences but also includes some sequences from some mollusc and nematode species.

In the phylogenetic tree inferred for Aqp3-like, we recovered several ancient duplications in the nodes where a habitat transition occurred. Previous studies have reported independent duplications of Aqp3-like within chelicerates (Finn et al., 2015), myriapods and copepods (Catalán-García et al., 2021). However, their results could have been biased due to the omission of other major arthropod lineages in their inferences. Here, we overcome this limitation by having representatives from virtually all arthropod
orders analysed together. Our results pointed to an ancestral duplication that was independently retained in some groups and lost in others (Figure 4 and Figure 4c), instead of inferring independent within-lineage duplications. For instance, Finn et al. (2015) report that two clades of Aqp3-like evolved in chelicerates based on their phylogeny with only spider, scorpion and acari sequences. When including the rest of the chelicerate orders, such as xiphosurans and pycnogonids, results show that the duplication they reported is arachnid-specific (with xiphosurans nested within arachnids, Figure 4d) and derives from one of the two ancient copies of Aqp3-like that are present in arthropods, with the other one being lost. On the contrary, Pycnogonida, the earliest-splitting chelicerate lineage, has retained both copies. This is consistent with a single terrestrialization event in arachnids (considering a secondary later colonization of water by xiphosurans as discussed in Ballesteros et al., 2022), where the duplication of Aqp3-like may have helped chelicerates to transition to a different niche. These Aqp3-like copies subsequently underwent independent duplications in both clades, with two subclades in Xiphosura, Ricinulei, Opiliones and Arachnopolmonata (which we named Aqp3-arq2, not likely the result of their WGD; Figure 4d), and two subclades in Acariformes, Scorpiones, Pseudoscorpiones and Araneae (which we named Aqp3-arq1, Figure 4d). Regarding myriapods, Catalán-Garcia et al. (2021) reported that Aqp3-like genes within Pauropoda, Symphyla and Diplpodopa clustered in two separate clades, those that we inferred in this study to be ancient duplications. Symphyla and Diplodopa retain both ancient copies (named Aqp3-arh1 and Aqp3-arh2 in Figure 4c), while Chilopoda and Pauropoda only...
retain Aqp3-arth1. Last, but not least, we reproduce the Pancrustacea results obtained by Catalán-García et al. (2021) regarding the presence of two copepod Aqp3-like clusters, one of them as a sister group to Thecostraca and Arguloidea Aqp3-like (in which we also report a close relationship to collemboan Aqp3-like), and the other clustered within the rest of Pancrustacea (Figure 4c). We also obtain an early-branching clade of Aqp3-arth1 composed of freshwater copepods and terrestrial diplurans, albeit recovered with low bootstrap support. Notably, we also confirm that holometabolous insects have lost Aqp3-like sequences in favour of a potentially more efficient insect-specific Aqp1-like-derived glycerol transporter, as previously reported by Finn et al. (2015). Moreover, we observe a Branchiopoda-specific duplication in Aqp3-arth1 (Figure 4c) that was also reported in previous studies (Catalán-García et al., 2021; Finn et al., 2015; Stavang et al., 2015).

Aqp3-like sequences from the other two members of Panarthropoda (Onychophora and Tardigrada) form a separate cluster from Arthropoda, with an intertwined evolutionary history. Figure 4b shows a shared duplication between Onychophora and Tardigrada, which suggests independent gene loss of both copies in arthropods. Each of these subclades underwent subsequent duplications, with an expansion within the onychophoran family Peripatopsidea. However, whether the tardigrade copies have a species-specific or a much older origin cannot be pinpointed with the current taxonomic representation.

In the case of terrestrial annelids (Clitellata), our results show highly supported gene duplications of Aqp3-ann2 in earthworms (Crassiclitellata), but not in the terrestrial enchytraeids (Figure 4e). It is noteworthy that the exclusively terrestrial Aqp3-ann2 clade clustered together with Aqp3-mol3, which contains aquaporins from mostly terrestrial panpulmonate gastropods, Caudofoveata and Cephalopoda. Concerning molluscs, we report a duplication of Aqp3-mol2 in the freshwater-origin molluscan subclass Caenogastropoda (Figure 4a, Krug et al., 2022), that was subsequently lost in the only marine species we included in our data set (Hinea brasiliensis, HBRA1). The presence or absence of this duplication in other Caenogastropoda marine species remains to be studied, but the presence of both copies in the terrestrial P. elegans (PELE1) suggests that it was at least retained in the common ancestor of this species with other marine members of the group. Interestingly, contrary to Stylommatophora, P. elegans has not lost Aqp3-mol1.

3.4 Convergent evolution of Aqp1-like glycerol transporters in Panpulmonata and Insecta via gene duplication

In addition to Aqp3-like, previous studies have proven the ability of some terrestrial-specific Aqp1-like proteins to transport glycerol. Such is the case of entomoglyceroporins, insect-specific Aqp1-like derived glycerol transporters (Finn et al., 2015), and AQP6, a tetrapod-specific classical aquaporin that can also transport glycerol (Holm et al., 2004). Moreover, the rise of entomoglyceroporins (Eglps) was accompanied by the loss of Aqp3-like in holometabolous insects. Here, we describe a similar pattern in molluscs: the rise of Aqp1-like malacoglyceroporins (Mglp, Figure 3c), and the lack of panpulmonate representatives in Aqp3-mol1 and Aqp3-mol2 (Figure 4). Mglps are glycerol transporter aquaporins which have been previously described as molluscspecific by Kosicka et al. (2016). However, our larger taxonomic representation of molluscs restricts its specificity to gastropods, at least to Euthyneura, given our taxonomic representation. In addition, the clade contains a duplication at its root, right before the origin of Panpulmonata, gastropods that have mantle-derived lungs. The nonmarine panpulmonate Phallomedusa solida (PSOL1) is in fact the only species that retains both copies after the duplication, suggesting a potential association with Panpulmonate transition out of the sea.

Regarding Eglps, we observe that the most diverse holometabolous insects (e.g., Diptera, Coleoptera, Lepidoptera and Hymenoptera) contain two or more groups of entomoglyceroporins (Figure S10), compared to nonholometabolous insects. Further investigations including a denser taxonomic representation of insects will help shed light on this observation and pinpoint the exact origin of each duplication. Moreover, some species or groups of insects that colonized freshwater or the sea contain additional copies of Eglps (i.e., Cloeon dipterum- CDIP1- , Gerris buenoi -GBUE1- and Anopheles gambiae- AGAM2- species and the dipteran family Ephyridae). The secondarily marine Anopheles merus (AMER1) lacks Eglps, which may be the result of a secondary loss.

4 DISCUSSION

4.1 Potential implications of parallel gene duplication and loss of aquaporin coding genes for animal terrestrialization

The ecological transition of animal life from the sea to the land, either directly or through a freshwater route, has shaped Earth’s biodiversity, with 77% of extant animal species inhabiting terrestrial habitats (Román-Palacios et al., 2022). At the physiological level, this habitat shift involved the acquisition of osmoregulatory adaptations to avoid water loss and keep the osmotic homeostasis of internal fluids. Nonmarine species are osmoregulators compared to marine species, which are osmoconformers (except for teleosts, probably due to their hypothetical freshwater origin; Carrete Vega & Wiens, 2012; Rankin & Davenport, 1981). In addition, the molecular function of aquaporins as a transporter of water and other molecules has been previously associated with osmoregulation (Cabrero et al., 2020; Finn et al., 2014; Spring et al., 2009; Wittekindt & Dietl, 2019; Yi et al., 2011) and even with terrestrialization in tetrapods (Finn et al., 2014). Here, we explored Aqp-coding gene gain, duplication and loss across several animal phyla that colonized land, providing evidence of ancient parallel duplications at the base of the lineages.
that transitioned out of the sea. As aquaporin transporting properties have only been sparsely explored in these nontetrapod lineages through functional experiments, we hypothesise the underlying role aquaporin duplications may have had on the emergence of osmoregulation during animal terrestrialization based on this experimental evidence. Notwithstanding, we acknowledge that these interpretations should be understood as phylogenomically-informed hypotheses given our taxonomic representation and not as scientific claims, and that further studies are necessary to test them in depth.

The role of polyols (e.g., glycerol) and sugars (e.g., trehalose) as osmolytes in animal osmoregulation has been known for years (Christoph et al., 2007; Lamitina et al., 2004; Thorat et al., 2012), as well as the ability of Aqp1-like to transport these molecules under osmotic stress (Gil et al., 2017; Laforenza et al., 2013). Thus, during animal paths out of the sea towards land or freshwater habitats, a higher number of Aqp1-like may have provided a more efficient osmolyte flux in fluids to counteract water loss (Figure 2b, left). Further studies on the transporting properties of Aqp3-like in marine and nonmarine species will help to elucidate this.

Arthropods are the most successful animal phylum to colonize land, with three independent ancient "out-of-sea" transitions (Dunlop et al., 2013). Our comparative phylogenetic analysis reported a significantly lower number of Aqp11-like in terrestrial arthropods (Figure 2b, centre). Based on the few experimental evidences of Aqp11-like supporting the role of this intracellular aquaporin in keeping redox/pH homeostasis in the endoplasmic reticulum (ER) (Bestetti et al., 2020; Ishibashi et al., 2021; Nozaki et al., 2008), we can hypothesise a potential link of gene loss in Aqp11-like coding genes with a more efficient homeostasis in land arthropods. During hypoxia, reactive oxygen species (ROS), such as H$_2$O$_2$, are generated in the ER. As the levels of oxygen are higher on land, the transition from the sea may have removed the need for H$_2$O$_2$ transport through Aqp11-like due to a lower production of ROS, which may have triggered gene loss in Aqp11-like coding genes. This hypothesis is further reinforced by the identification of several independent duplications of Aqp11-like coding genes in aquatic arthropod species – that is, xiphosurans and mayflies (Figure S5), all of them occurring in species that had transitions back to aquatic habitats (from land to sea, and land to freshwater, respectively). In addition, we detect Aqp11-like duplications in freshwater copepods and crayfish, which may be related to their transition from sea to freshwater. Further functional studies are needed to understand the functional role of this enigmatic type of aquaporins and their role during arthropod terrestrialization.

Duplications and loss of other aquaporins may have also helped arthropods to successfully conquer land. Here, we report an ancient duplication of Aqp1-like with a secondary loss of Aqp1-arth in most Pancrustacea (especially in hexapods, Figure 3). This loss could be a consequence of the origin of the hexapod-exclusive Drip cluster, as functional studies in D. melanogaster have demonstrated the high efficiency of water transport in Malpighian tubule cells expressing Prip and Drip, as well as their importance in desiccation resistance (Cabero et al., 2020). The functional efficiency of Prip-like duplicates in the other land-dwelling arthropod lineages (arachnids and myriapods) remains to be assessed. Furthermore, our phylogenetic analysis of Aqp3-like shows an ancient gene duplication at the root of arachnids (Figure 4d). We could hypothesise that this new copy might have replaced the ancient Aqp3-like if it were neo- or subfunctionalized into a more efficient transporter, as reported with other aquaporins (Finn et al., 2015). Nevertheless, the subfunctionalization of Lepeophtheirus salmonis’ Aqp3-like in the enterocytes reported by Catalán-García et al. (2021) may suggest a similar role in intestine water absorption in other arthropods. Lastly, the presence of an Aqp3-like duplication in Branchiopoda (Figure 4c), crustaceans of marine origin that colonized brackish and freshwater environments (Oakley et al., 2013), with Thecostraca having lost one of the copies, reinforces our hypothesis of the role of aquaporins in osmoregulation.

The two phyla most closely related to arthropods, onychophorans and tardigrades, both of them having colonized land independently, also show Aqp-coding gene duplications in Aqp1- like (Figure 3) and Aqp3-like (Figure 4b). However, only the role of M. tardigradum Aqp4 (Aqp-like) in anhydrobiosis has been previously reported (Schokraie et al., 2012), hinting at the putative role onychophoran Aqp3-like may have in desiccation resistance and, therefore, in nonarthropod Panarthropoda terrestrialization. The role of Aqp1-like duplications in onychophoran physiology can only be hypothesised based on comparisons with other animals in the absence of functional studies.

The ecological transition from a marine to a freshwater environment in clitellates (Erseus et al., 2020) may have been accompanied by the reshaping of their osmoregulatory capabilities. Here we report a wide expansion of Aqp1-ann6 in terrestrial clitelates, and loss in non-Crassiclitellata annelids (Figure 3e). The class Clitellata is characterized by the presence of a clitellum that secretes a cocoon. Cocoons are egg sacs that nourish the developing embryos, whose structure and composition differ between clitellates. Previous studies have shown that earthworm but not enchytraeids cocoons accumulate polyols when exposed to cold or drought (Bauer et al., 2001; Holmstrup, 1995). Thus, successive highly supported gene duplications of putative polyol transporters Aqp3-ann2 in Crassiclitellata, but not enchytraeids (Figure 4e), suggest their role in this cocoon’s physiological response to drought and cold. Resistance to drought and cold might have facilitated their survival to the heterogeneous environmental conditions of land. Nevertheless, given the lack of functional studies of these aquaporins or the organs in which they are expressed, we can only suggest putative roles in annelid terrestrialization based on the osmoregulatory role of aquaporins in nonannelid organisms (Cabero et al., 2020; Catalán-García et al., 2021; Finn et al., 2014; Gil et al., 2017; Laforenza et al., 2013; Zieger et al., 2022).

The ecological transitions of gastropods to land, either through a freshwater path or marginal environments (Krug et al., 2022; van Straalen, 2021; Vermeij & Watson-Zink, 2022), could have been facilitated by aquaporins evolution. Our independent comparative phylogenetic analyses for each type of Aqp-coding gene revealed a
significantly higher number in nonmarine species of Aqp1-like and Aqp8-like (Figure 2b, right). As the number of aquaporins in terrestrial species compared to nonterrestrial ones is not significant, this suggests that, as in the case of aquaglyceroporins in the five-phyla analysis, aquaporin-coding gene duplication potentially enabled molluscs to cope with osmotic stress during their ecological transition to freshwater and land. In fact, this potential association is further reinforced by the identification of Aqp1-like, Aqp3-like and Aqp8-like duplications (Figures 3a, c, f and 4a, S8, respectively), and previous functional experiments. To give an example, we recovered the expansion of Aqp1-mol2 in freshwater bivalves (Figure S6) as sister group to a previously reported expansion in Dreissena rostriformis (Calcino et al., 2019), which was suggested to have a role in the transition from sea to freshwater as well as having a critical role during early cleavage in D. rostriformis (Zieger et al., 2022).

Regarding terrestrial gastropods, Stylommatophora-specific duplication within Aqp1-mol1 that contain previously described HpAqp1 and HpAqp2 (Kosicka et al., 2016), may have contributed to the success of this order in its ecological transition to land via sub- or neofunctionalization to maximize water reabsorption in the intestine. This hypothesis assumes that HpAqp2 and HpAqp1 orthologues preserve the same function in the rest of stylommatophoran species studied here, where HpAqp2 orthologues would be highly expressed in the intestine, an organ whose main function is water reabsorption (Forester, 1977), but HpAqp1 is expressed at a low level in foot, kidney and intestine (Kosicka et al., 2016). In addition, one of the Mgp (Figure 3a) and one of the Mgpl (Figure 3c) paralogues contain the aetiation-related HpAqp6 and HpAqp3, respectively, which are expressed in the foot during snail aestivation (Kosicka et al., 2020), one of the mechanisms of adaptation most relevant to humidity stress (Schweizer et al., 2019). This functional relationship to the drought-related behaviour evokes the putative role this pulmonate-specific duplication might have had on the panpulmonate terrestrialization events.

Lastly, in contrast to Aqp1-like or Aqp3-like, not much is known about Aqp8-like, except for their ability to transport water and ammonia and their expression in the inner membrane of mammalian mitochondria (Calamita et al., 2005; Lee & Thévenod, 2006; Soria et al., 2010). The relevant contribution of Aqp8-like to mitochondrial respiration (Ikaga et al., 2015) implies its importance for normal cell functioning and stress tolerance (Sokolova, 2018). Thus, the duplication of Aqp8-like in C. elegans, freshwater bivalves and the terrestrial gastropod P. elegans -PELE1- (Figure S8) suggests their importance during the “out of the sea” transition. Further studies that include more nematodes, as well as functional characterization of this aquaporin will shed light into its role during this major evolutionary transition.

5 | CONCLUSIONS

The molecular function of aquaporins as transporters of water and other molecules makes them an ideal system to interrogate the evolution of osmoregulation during animal terrestrialization, which represents one of the most dramatic ecological transitions that animals faced during their evolutionary chronicle. Here, we provide phylogenetically-informed evolutionary evidence of gene duplication and loss in Aqp-coding genes and discuss their potential key role during the evolution of terrestrial animal life. By comparing our phylogenetic results with previously described functional evidence, we hypothesise in which physiological processes aquaporins might have contributed to the ecological transition of the different animal phyla from the sea to freshwater or land, summarized in Table 1. Overall, our results suggest that ancient independent expansions of aquaporins in nonmarine animals may have played a key role in the adaption to life on land through the evolution of more sophisticated osmoregulatory mechanisms, such as a potential role of Aqp1-like in molluscan aestivation or the putative role of Aqp3-like in earthworm cocoon resistance to drought. Our results also show the complexity of aquaporin evolution in animals, highlighting the importance of taxon sampling when studying ancient gene families. All in all, our comparative phylogenomic study focuses on the emergence and reshaping of osmoregulation through aquaporin evolution and sets the road towards a deeper understanding of a fundamental transition underpinning modern biodiversity: animal terrestrialization.

AUTHOR CONTRIBUTIONS

Gemma I. Martínez-Redondo and Rosa Fernández conceived the study. Gemma I. Martínez-Redondo performed the analyses and designed the figures. Gemma I. Martínez-Redondo and Carolina Simon Guerrero annotated aquaporins. Pau Balart-García and Vanina Tonzo assisted aquaporin annotation. Leandro Aristide designed the Poisson regression analysis. Leandro Aristide and Vanina Tonzo inferred the ultrametric trees. Leandro Aristide and Gemma I. Martínez-Redondo performed the Poisson regression analysis. Gemma I. Martínez-Redondo and Rosa Fernández interpreted the results. Gemma I. Martínez-Redondo and Rosa Fernández wrote the first version of the manuscript. Rosa Fernández provided resources and supervised the study. All authors revised and approved the final version of the manuscript.

ACKNOWLEDGEMENTS

GIMR acknowledges the support of Secretaria d’Universitats i Recerca del Departament d’Empresa i Coneixement de la Generalitat de Catalunya and European Social Fund (ESF) “Investing in your future” (grant 2021 FI_B_00476) and the CSIC 2020 JAE Intro programme (JAEINT_20_02210). LA acknowledges funding from a Juan de la Cierva-Formación (grant agreement no. FJC2019-042184-I funded by MCIN/AEI/10.13039/501100011033). PBG was supported by an FPI grant (grant agreement no. BES-2017-081050) financed by MCIN/AEI/10.13039/501100011033 and by ESF “Investing in your Future”. RF acknowledges support from the following sources of funding: Ramón y Cajal fellowship (grant agreement no. RYC2017-22492 funded by MCIN/AEI/10.13039/501100011033 and ESF “Investing in your future”), Agencia Estatal de Investigación (project PID2019-108824GA-I00.
REFERENCES
Abascal, F., Irisarri, I., & Zardoya, R. (2014). Diversity and evolution of membrane intrinsic proteins. *Biochimica et Biophysica Acta*, 1840(5), 1468–1481.
Andersson, F. E., & Lindgren, A. R. (2021). Phylogenetic analyses recover a clade of large-bodied decapodiform cephalopods. *Molecular Phylogenetics and Evolution*, 156, 107038.
Aristide, L., & Fernández, R. (2022). Genomic innovation, convergence and parallelism facilitated the ecological transition to life out of the sea in molluscs. *Research Square*. https://doi.org/10.21203/rs.3.rs-2382076/v1
Baker, C. M., Buckman-Young, R. S., Costa, C. S., & Giritet, G. (2021). Phylogenetic analysis of velvet Worms (Onychophora) uncovers an evolutionary radiation in the neotoropics. *Molecular Biology and Evolution*, 38(12), 5391–5404.
Ballesteros, J. A., Santibáñez-López, C. E., Baker, C. M., Benavides, L. R., Cunha, T. J., Gainett, G., Ontano, A. Z., Setton, E. V. W., Arango, C. P., Gavish-Regev, E., Harvey, M. S., Wheeler, W. C., Homriga, G., Giritet, G., & Sharma, P. P. (2022). Comprehensive species sampling and sophisticated algorithmic approaches refute the monophyly of Arachnida. *Molecular Biology and Evolution*, 39, msac021.
Bauer, R., Worland, M. R., & Block, W. (2001). Experimental studies on cold survival of enchytraeid coocons. *Pedobiologia*, 46, 561–571.
Benga, G. (2012). On the definition, nomenclature and classification of water channel proteins (aquaporins and relatives). *Molecular Aspects of Medicine*, 33(5–6), 514–517.
Benoit, J. B., Hansen, I. A., Szuter, E. M., Drake, L. L., Burnett, D. L., & Attardo, G. M. (2014). Emerging roles of aquaporins in relation to the physiology of blood-feeding arthropods. *Journal of Comparative Physiology B*, 184, 811–825.
Bestetti, S., Galli, M., Sorrentino, I., Pinton, P., Rimessi, A., Sità, R., & Medrano-Fernandez, I. (2020). Human aquaporin-11 guarantees efficient transport of H$_2$O$_2$ across the endoplasmatic reticulum membrane. *Redox Biology*, 28, 101326.
Cabrerob, P., Terhazas, S., Dornan, A. J., Chimire, S., Holmes, H. L., Turin, D. R., Romero, M. F., Davies, S. A., & Dow, J. A. T. (2020). Specialized stellate cells offer a privileged route for rapid water flux in *drosophila* renal tubule. *Proceedings of the National Academy of Sciences*, 117, 1779–1787.
Calamita, G., Ferri, D., Gena, P., Liquori, G. E., Cavalier, A., Thomas, D., & Svelto, M. (2005). The inner mitochondrial membrane has aquaporin-8 water channels and is highly permeable to water. *The Journal of Biological Chemistry*, 280(17), 17149–17153.
Calcino, A. D., de Oliveira, A. L., Simakov, O., Schwahe, T., Zieger, E., Wollesen, T., & Wanninger, A. (2019). The quagga mussel genome and the evolution of freshwater tolerance. *DNA Research: An International Journal for Rapid Publication of Reports on Genes and Genomes*, 26(5), 411–422.
Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25(15), 1972–1973.
Carrete Vega, G., & Wiens, J. J. (2012). Why are there so few fish in the sea? *Proceedings. Biological Sciences/The Royal Society*, 279(1737), 2323–2329.
Castiglione, S., Tesone, G., Piccolo, M., Melchionna, M., Mondanaro, A., Serio, C., Di Febbraro, M., & Raia, P. (2018). A new method for testing evolutionary rate variation and shifts in phenotypic evolution. *Methods in Ecology and Evolution*, 9, 974–983.
Catalán-García, M., Chauvigné, F., Stavang, J. A., Nilsen, F., Cerdá, J., & Finn, R. N. (2021). Lineage-level divergence of copepod glycerol transporters and the emergence of isoform-specific trafficking regulation. *Communications Biology*, 4(1), 643.
Chen, H., Parry, L. A., Vinther, J., Zhai, D., Hou, X., & Ma, X. (2020). A Cambrian crownannelid reconciles phylogenomics and the fossil record. *Nature*, 583, 249–252.
Christoph, K., Beck, F.-X., & Neuhofer, W. (2007). Osmoadaptation of mammalian cells - an orchestrated network of protective genes. *Current Genomics*, 8(4), 209–218.
Colgan, D. J., Ponder, W. F., Beacham, E., & Macaranas, J. (2007). Molecular phylogenetics of Caenogastropoda (Gastropoda: Mollusca). *Molecular Phylogenetics and Evolution*, 42, 717–737.
Colgan, D. J., & Santos, R. P. (2018). A phylogenetic classification of gastropod aquaporins. *Marine Genomics*, 38, 59–65.
Cunha, T. J., & Giritet, G. (2019). A congruent topology for deep gastropod relationships. *Proceedings of the Royal Society B*, 286, 20182776.
Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Bennett, R., Berry, A., Bhai, J., Bignell, A., Billis, K., Boddu, S., Brooks, L., Charkhchi, M., Cummins, C., Fioretto, L. D. R., & Flicek, P. (2021). *Ensembl 2022. Nucleic Acids Research*, 50(D1), D988–D995.
de Marè, S. W. V., Venskutonytė, R., Eltschkner, S., de Groot, B. L., & Lindkvist-Petersson, K. (2020). Structural basis for glycerol efflux transporters and the emergence of isoform-specific trafficking regulation. *Structure*, 28(1), 215–222.
Doğan, Ö., Schrödl, M., & Chen, Z. (2020). The complete mitogencode of Moquin-Tandon, 1855 (Gastropoda: Stylommatophora): Mitochondrial genome architecture, evolution and phylogenetic considerations within Stylommatophora. *PeerJ*, 8, e8603.
Dohrmann, M., & Wörheide, G. (2017). Dating early animal evolution using phylogenomic data. *Scientific Reports*, 7, 3599.
Dunlop, J. A., Scholz, G., & Selden, P. A. (2013). Water-to-land transitions. In A. Minelli, G. Boxshall, & G. Fusco (Eds.), *Arthropod biology and evolution*. Springer.
Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Computational Biology*, 7(10), e1002195.
Zardoya, R., & Villalba, S. (2001). A phylogenetic framework for the aquaporin family in eukaryotes. *Journal of Molecular Evolution, 52*, 391–404.

Zieger, E., Schwaha, T., Burger, K., Bergheim, I., Wanninger, A., & Calcino, A. D. (2022). Midbody-localized aquaporin mediates intercellular lumen expansion during early cleavage of an invasive freshwater bivalve. *Frontiers in Cell and Developmental Biology, 10*, 894434.

**SUPPORTING INFORMATION**

Additional supporting information can be found online in the Supporting Information section at the end of this article.

---

**How to cite this article:** Martínez-Redondo, G. I., Simón Guerrero, C., Aristide, L., Balart-García, P., Tonzo, V., & Fernández, R. (2023). Parallel duplication and loss of aquaporin-coding genes during the “out of the sea” transition as potential key drivers of animal terrestrialization. *Molecular Ecology, 32*, 2022–2040. [https://doi.org/10.1111/mec.16854](https://doi.org/10.1111/mec.16854)