GBE

Genomic Exaptation and Regulatory Landscape Shifts as Key Mechanisms Enabling Flatworm Terrestrialization

Belén Carbonetto [1,*,†], Lisandra Benítez-Álvarez [1,†], Raquel García-Vernet [1], Klara Eleftheriadi [1], Carlos Vargas-Chávez [1], Nuria Escudero [1], Judit Salces-Ortiz [1], Iñaki Rojo [2], Fernando Ángel Fernández-Álvarez [3,4], Corey Alwell [5], Cristina Chiva [6,7], Eduard Sabidó [6,7], Eduardo Mateos [8,9], Fernando Carbayo [10], Rosa Fernández [1,*]

[1]Metazoa Phylogenomics & Genome Evolution Lab, Institute of Evolutionary Biology (CSIC-UPF), Barcelona 08003, Spain

[2]Citizen Scientist, Sopelana, Euskadi, Spain

[3]Institut de Ciències del Mar (ICM, CSIC), Barcelona 08003, Spain

[4]Instituto Español de Oceanografía, Centro Oceanográfico de Gijón (IEO, CSIC), Gijón 33212, Spain

[5]Department of Genetics, Trinity College Dublin, Dublin 2, Ireland

[6]Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Catalonia 08003, Spain

[7]Universitat Pompeu Fabra (UPF), Barcelona, Catalonia 08003, Spain

[8]Departament de Biologia Evolutiva, Ecologia i Ciencies Ambientals. Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain

[9]Institut de Recerca de la Biodiversitat, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain

[10]Laboratório de Ecologia e Evolução, Escola de Artes, Ciências e Humanidades, Universidade de São Paulo– USP, São Paulo, SP CEP 03828-000, Brazil

[†]These authors contributed equally.

*Corresponding authors: E-mails: belen.carbonetto@ibe.upf-csic.es; rosa.fernandez@ibe.upf-csic.es.

Accepted: March 03, 2026

Abstract

Understanding the genomic toolkit that enabled animal terrestrialization, the shift from aquatic to terrestrial habitats, is key to uncovering the evolutionary origins of land biodiversity. Yet, the genomic basis of the physiological and metabolic adaptations required for life on land remains poorly understood across most terrestrial animal phyla. Planarians (Platyhelminthes) offer a powerful model, as only one terrestrial lineage, the Geoplanidae (order Tricladida), is known. Here, we integrated genomics, transcriptomics, and proteomics to explore the genetic changes potentially supporting terrestrial adaptation. We identified a major burst of gene gain in the lineage leading to Tricladida, preceding the radiation of terrestrial planarians. Upon abiotic stress exposure, terrestrial and freshwater species exhibited distinct responses: most differentially expressed genes belonged to orthogroups gained in Tricladida, with over half under strong directional selection in terrestrial flatworms, suggesting their adaptive relevance. Transcriptomic profiles revealed divergent strategies: terrestrial species upregulated ancient genes, while freshwater species downregulated a separate set of ancestral genes. Across all datasets, the abiotic stress–response toolkit in terrestrial planarians was markedly different from freshwater relatives, with significant regulatory divergence. Our results highlight gene gain and co-option, rather than lineage-specific innovations, as key drivers of terrestrial flatworm evolution, emphasizing genomic exaptation and regulatory shifts as central to terrestrialization in Platyhelminthes. This study provides the first

genome-wide view of the genetic basis of flatworm terrestrialization and sheds light on broader patterns of animal terrestrial adaptation.

Key words: phylogenomics, genome evolution, platyhelminthes, transcriptomics.

Significance

The move from water to land reshaped animal evolution, yet the genetic changes that made terrestrial life possible remain poorly understood. This study uncovers a major burst of gene gain in the ancestor of Tricladida flatworms, an event that occurred before the rise of land-dwelling species, and shows that many of these newly acquired genes were later co-opted for surviving terrestrial stresses. By revealing that adaptation to land relied on repurposing and expanding existing genes rather than inventing entirely new ones, this work provides fresh insight into how genomic innovation and regulatory flexibility drive evolutionary transitions from aquatic to terrestrial environments.

Introduction

The colonization of land by animals is without doubt one of the most important episodes in the animal evolutionary chronicle. The ancestors of extant terrestrial arthropods, nematodes, nemerteans, velvet worms, and mollusks, among other phyla, managed to transition from marine to terrestrial environments and successfully adapt to life on land. For that, deep physiological and metabolic changes evolved in response to the markedly distinct abiotic conditions significantly differ on land and in the sea. This likely triggered profound genomic reshaping to meet new challenges such as osmoregulation, oxygen homeostasis, locomotion, reproduction, novel pathogens, and food sources. While the genomic changes facilitating this transition were studied in animal phyla, such as mollusks (Sun et al. 2019; Liu et al. 2021; Krug et al. 2022; Aristide and Fernández 2023), vertebrates (Bi et al. 2021; Meyer et al. 2021) or arthropods (Sharma 2017; Asano et al. 2019; Veldsman et al. 2021), little attention has been paid to the phylum Platyhelminthes. They constitute an excellent model to explore the genomic basis of animal terrestrialization as all described terrestrial forms are within a single clade, the family Geoplanidae. Being sister to families and superfamilies comprising freshwater forms, it is clear that the transition from marine to terrestrial environments in platyhelminthes occurred through the freshwater route (van Straalen 2021). Within the order Tricladida, we find a suborder composed mostly by marine species (suborder Maricola), a second suborder composed of freshwater and terrestrial species (suborder Continenticola) (Sluys et al. 2009) and a third clade comprising freshwater species inhabiting in epigean and hypogean habitats (suborder Cavernicola). The interrelationships between the three clades have not been properly resolved yet due to incongruences between ribosomal markers (Benítez-Álvarez et al. 2020; Stocchino et al. 2021) and transcriptomic data (Vila-Farré et al. 2023). The suborder Continenticola comprises Planarioidea (freshwater planarians) and Geoplanoidea, which are sister groups; the latter comprises Dugesiidae (freshwater planarians) and Geoplanidae (land planarians) (Sluys et al. 2009).

It is hypothesized that planarians were partially exapted for life on land even before colonization. In particular, mucus secretion may have played an important adaptative role in osmoregulation and respiration, locomotion, prey capture, and defense (Sluys 2019). In terrestrial species it has been consistently documented that the mucus-secreting epidermis mediates a tradeoff between locomotory efficiency and water conservation, ensuring survival during environmental hydric stress (McGee et al. 1997). Furthermore, internal fertilization, hard-shelled cocoons, and direct development allowed planarians to become fully terrestrial organisms, as their life cycle is not dependent on aquatic environments (Little 1983). Moreover, the presence of a protonephridial system, also present in freshwater lineages, may have played a central role in osmoregularity for adaptation in land (Silveira and Corinna 1976). However, unlike aquatic ancestors that prioritize the expulsion of excess water, terrestrial planarians exhibit a highly branched tubule network which, linked to the ability of urea excretion, is optimized for selective reabsorption and osmotic homeostasis (Campbell 1965). Additionally, the parenchymal musculature in terrestrial planarians underwent a significant development during the transition to land. This well-developed muscular network provided the hydrostatic support necessary to counteract gravity and facilitated complex locomotory patterns (Winsor 1998).

The emergence, maintenance and improvement of such specialized traits relied on a robust molecular foundation. Consequently, the evolutionary trajectory of these lineages was intimately linked to the reshaping of their underlying gene repertoire. This genomic

restructuring via gene gain, duplication and loss, is a well-known mechanism facilitating adaptation and innovation. As new coding genetic pieces arise through gene gain or duplication, new functions may arise as well associated with them (Fernández and Gabaldón 2020; Merényi et al. 2023). Furthermore, evolutionary changes in gene repertoires can lead to the diversification of gene functions. This diversification allows organisms to exploit new ecological niches, utilize different resources, or develop resistance to new pathogens or environmental stressors (Wen et al. 2023; Pal and Andersson 2024). As for gene loss, it has been proven a powerful mechanism for genomic innovation as well, as gene interactions and regulatory networks may be strongly impacted, resulting in the reshaping of the functioning gene circuits (Jiménez-Marín et al. 2023; Domazet-Lošo et al. 2024).

Gene gain and duplication has been shown to facilitate adaptation to new environments through several mechanisms, including neofunctionalization (i.e. it creates gene redundancy, allowing one gene copy to maintain its original function while the other evolves new functions), or subfunctionalization (i.e. the original function is divided between the copies) (Braasch et al. 2018; Sandve et al. 2018; Birchler and Yang 2022). This genetic redundancy increases genetic variation, providing a broader substrate for natural selection and enhancing the potential for beneficial mutations while buffering against harmful mutations, as the extra gene copy can often compensate for deleterious changes (Robinson et al. 2023; Xie et al. 2023). Gene duplication can also lead to increased gene expression levels, which may be advantageous in certain environmental conditions, and supports the evolution of complex regulatory networks and novel developmental pathways, contributing potentially to significant morphological innovations (Li et al. 2020; Van de Peer et al. 2021; Veitia and Birchler 2022). These processes can drive rapid speciation and diversification, enabling organisms to explore and thrive in new and changing environments. However, gene repertoire evolutionary dynamics within animal phyla and its role in animal terrestrialization is still largely unexplored.

Here, we used a combination of phylogenomics, transcriptomics and proteomics, with emphasis on revealing gene repertoire evolutionary dynamics to gain insight into the toolkit that may have been instrumental for adaptation of planarians to terrestrial environments. We sequenced the first comprehensive collection of high-quality transcriptomes of terrestrial flatworms, to date, comprising 16 species. We predicted the encoded genes and explored their evolutionary dynamics to understand how gene gain, duplication and loss may have contributed to terrestrialization. We further exposed both aquatic and terrestrial planarian species to abiotic stressors representing the primary environmental pressures associated with the transition from sea to land. We found differential patterns of gene expression between aquatic and terrestrial species. In addition, we used phylostratigraphy to place differentially expressed genes (DEGs) under stress conditions on the Platyhelminthes phylostrata. Our results revealed a burst of genomic innovation prior to the origin and diversification of terrestrial planarians. We also showed that most genes involved in stress–response in extant species were associated to orthologous groups that originated in the branch leading to that burst. Taken together, these results point to genomic exaptation as a key factor facilitating the colonization of land environments in platyhelminthes. Our results provide the first genome-wide evidence of the toolkit involved in flatworm terrestrialization, and more broadly contribute to delineating the big picture of the genomic basis of animal terrestrialization.

Results

A Burst of Gene Gain at the Origin of Tricladida as a Key Driver of Flatworm Terrestrialization

Gene repertoire evolution was inferred across a comprehensive Platyhelminthes dataset including 54 total species, 16 being terrestrial species. Results showed a high number of gained hierarchical orthologous groups (HOGs) in the internodes leading to Tricladida, Continenticola, and Geoplanoidea (Fig. 1a and Tables S4 and S5). In contrast, the number of gained HOGs in the internode leading to the terrestrial planarians (Geoplanidae) was lower (Fig. 1a and Table S4). The presence of the gained HOGs was confirmed experimentally at the proteomic level in extant species *O. nungara* and *S. mediterranea.* The pattern of frequency of HOGs with observed protein expression is concordant with the gained gene repertoire pattern inferred (Fig. 1b). The nodes that showed the highest number of gained HOGs also showed the highest number of HOGs with proteins present in extant species (Table S6). These observations suggest that the emergence of novel genes at the internode leading to Geoplanidae did not constitute the bulk of the genomic changes responsible of flatworm terrestrialization; instead, the exaptation of HOGs that emerged in the previous nodes, shared with aquatic species, also likely played critical role.

In order to better understand the relationship between gained HOGs and putative adaptive traits we performed a functional characterization based on Gene Ontology (GO). The analysis was based on the enrichment gained HOGs against a background of ancestral HOGs (plus gained HOGs) at each node of interest. This way we could infer enriched novel functions. First, we evaluated high-level functional signatures rather than literal mechanistic inferences.

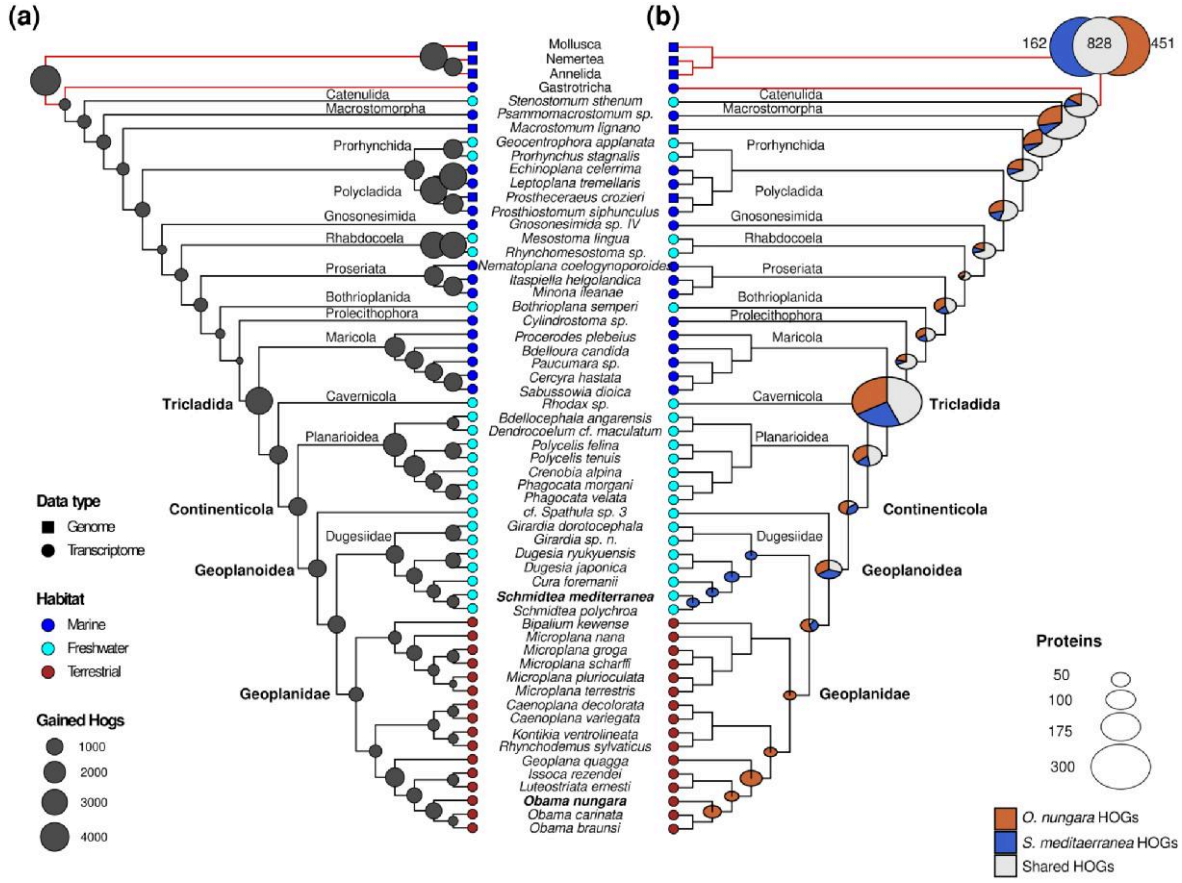

Fig. 1. Gene repertoire evolution across platyhelminthes. The outgroups are highlighted with red branches and the members of Platyhelminthes with black branches. Colors at the tips correspond to habitat, dark blue for marine species, light blue for freshwater and brown for terrestrial. a) The size of the nodes is proportional to the number of gained HOGs b) Proportion of HOGs with observed proteins in extant species: *Obama nungara* (dark brown) *Schmidtea mediterranea* (dark blue), or both (grey). The size of the pie charts is proportional to the number of HOGs except for the first node where the information is shown in a Venn diagram and the number of HOGs is indicated.

To highlight broad categories, we calculated the Wang semantic similarity between significantly enriched terms and clustered them using the K-means algorithm. Then, we obtained the representative GO term for each cluster by selecting the GO term with the farthest value in the similarity network for that cluster. The cluster representatives almost did not overlap between Tricladida, Continenticola, Geoplanoidea, Geoplanidae and Dugesiidae (Fig. 2). This high-level functional signature pattern can draft a stepwise acquisition of innovations. Our data suggests a modular evolutionary trajectory: innovation in the Tricladida and Continenticola nodes is dominated by DNA and RNA metabolic processes and their regulation and by the regulation of life cycle and cell cycle, but also by environmental response traits (*response to odorant* (GO:1990834)*, negative regulation of epithelial cell differentiation* (GO:0030857)*, adaptive immune response* (GO:0002250)). Most specialized traits for fresh water or land adaptation emerged later. Geoplanoidea gained genes showing enriched GO terms related to *osmotic stress* (GO:0006970) and a proxy of *protonephridia development (*actually *lung alveolus development,* GO:0048286, which is conventionally associated with vertebrate respiration but likely reflects a recruitment of conserved branching morphogenesis functions*)*. Finally, Geoplanidae gains, while still presenting enriched DNA and RNA metabolic related functions, showed enriched GOs like *skin development* (GO:0043588) and *creatine metabolic process* (GO:0006600), this innovation likely related to high-torque muscular contractions required for terrestrial locomotion. (Dugessidae clusters showed enriched GO terms under the umbrella of *skin development* (GO:0043588) and *positive regulation of biomineral tissue development* (GO:0070169) and *response to protozoan* (GO:0001562).

We later analyzed the list of enriched GO terms within the clusters to evaluate more specific functional signatures in Tricladida, the node where gain bust accrued, and in Geoplanidae, the node where the land transition occurred. In Tricladida, included. functions linked to osmoregulation and desiccation resistance,

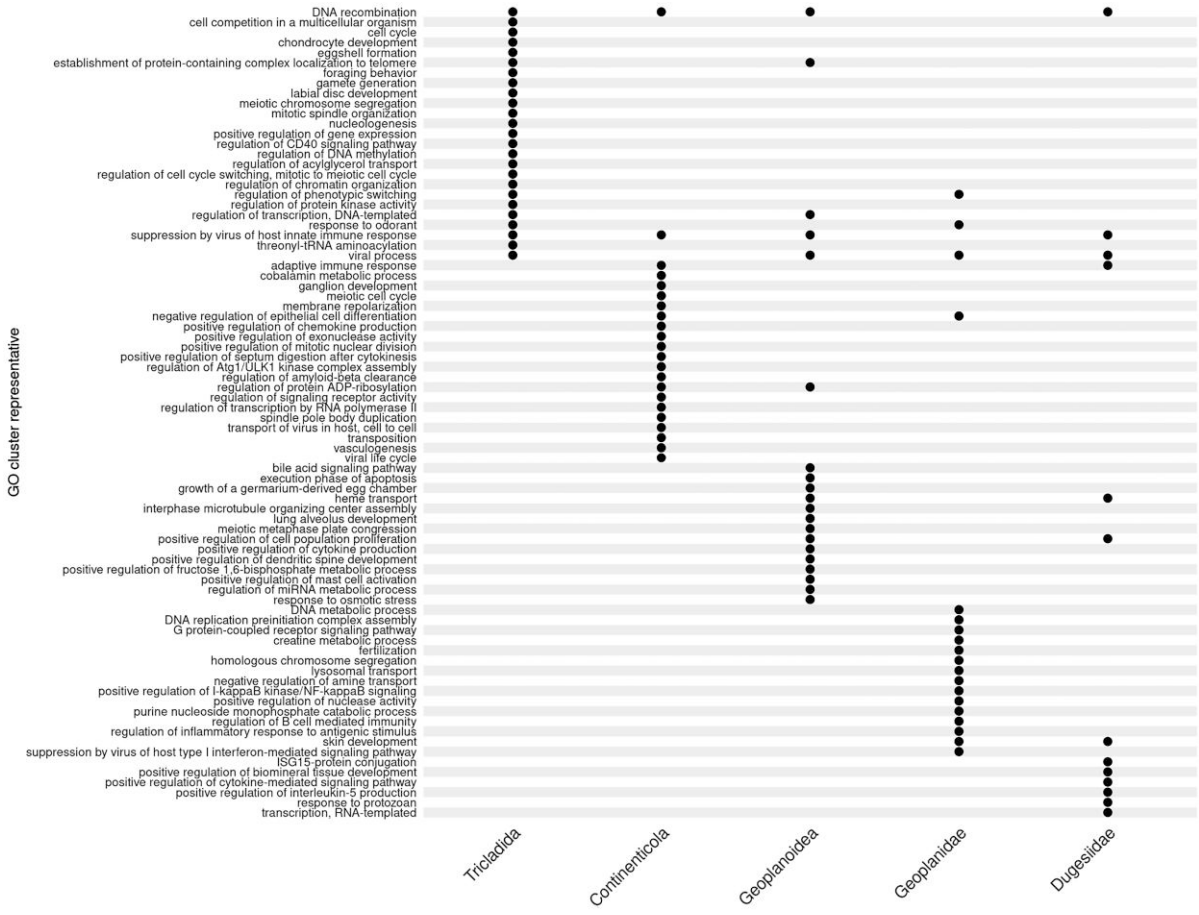

Fig. 2. Cluster farthest representatives of enriched GO terms (biological process) assigned to genes gained in Tricladida, Continenticola, Geoplanoidea, Geoplanidae and Dugesiidae.

such as *cellular water homeostasis* (GO:0009992), *cell volume homeostasis*, *cell volume negative regulation* (GO:0006884, GO:0045794) and ion transport (GO:0098719, GO:0055064), were enriched. In addition, terms linked to the regulation of protection against radiation were enriched, for example, *response to ionizing radiation* (GO:0010212, GO:0071479) and *response to gamma radiation* (GO:0010332, GO:0110011, GO:2001229). Functions associated with later water-independent reproductive strategies were also significantly enriched in Triclaida, such as *eggshell formation* (GO:0030703), *vitelline membrane formation* (GO:0007305), *eggshell chorion* gene amplification (GO:0007307), *chorion-containing eggshell pattern formation* (GO:0030381), and *regulation of timing of transition from vegetative to reproductive phase* (GO:0048510). Several genes with functions related to abiotic stress–response were gained in Tricladida, too: *cellular response to temperature stimulus* (GO:0071502), *cellular response to glucose starvation* (GO:0042149), *behavioral response to starvation* (GO:0042595), *positive phototaxis* (GO:0046956). Additionally, genes gained in the Geoplanidae node were enriched in functions probably beneficial during land conquest, too. For example terms such as *response to UV-A* (GO:0070141), *response to high light intensity* (GO:0009644), *response* to *far red light (*GO:0010218*)* and *sensory perception and detection of light stimulus* (GO:0050953, GO:0009583) reflect traits of photoprotection and sensory adaptation to direct solar radiation. Moreover, terms such as *regulation of body fluid levels* (GO:0050878), *osmosensory signaling pathway* (GO:0007231), *metanephros development* (GO:0001656), *urate transport* (GO:0015747), *glomerular visceral epithelial cell migration* (GO:0090521) describe genes that may have played an important role in osmotic homeostasis out of the water. In addition, adaptation to land must have involved the specialization of complex locomotory patterns. Concurrently, genes gained in Geoplanidae node were enriched in the following functions: *locomotory exploration behavior* (GO:003564), *negative regulation of locomotion*

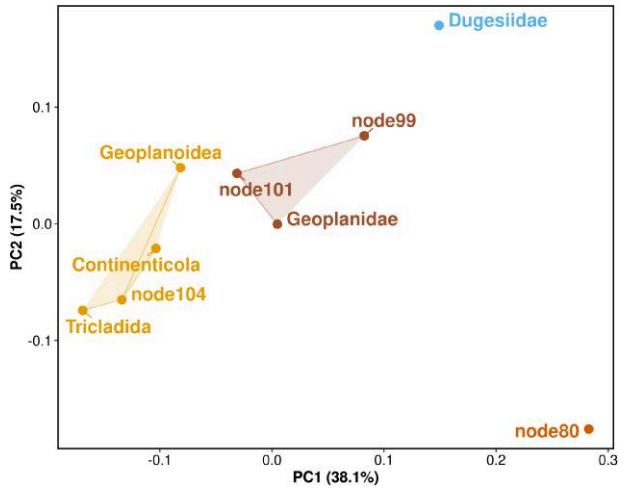

Fig. 3. Functional clustering of gained HOGs across the Tricladida phylogeny. Principal Coordinate Analysis (PCoA) based on Wang's semantic similarity of enriched GO terms associated with genes gained at each node.

involved in locomotory behavior (GO:0090327) and *O-glycan processing* (GO:0016267), needed for mucus production. These are only a sample of GO terms enriched in these nodes, and while speculative, they may represent molecular signatures of the evolutionary transition from aquatic to terrestrial life in flatworms.

To compare the acquired functional novelties across nodes, we compared the semantic similarity between enriched GO at different nodes. Results showed the presence of four clusters, the first one grouping the most ancient nodes, from Tricladida to Geoplanoidea and the second comprising Geoplanidae and the ancestor of terrestrial South American planarians (represented by Geoplaninae in this study) (Fig. 3). Dugesiidae did not cluster with the rest of the nodes and neither did the ancestor of the Eurasia group, which clustered apart. These quantitative data support our qualitative assessment, showing functional divergence in a step-wise pattern where ancestral Tricladida traits differ from those acquired during the independent transitions to freshwater and land environments. Moreover, Dugesiidae- and Geoplanidae-enriched GO terms also present functional divergence between them.

Different Genetic Toolkit Involved in the Response to Abiotic Stress in Freshwater and Terrestrial Planarians

In order to evaluate signs of adaptive genomic traits in extant species, we exposed *O. nungara* and *S. mediterranea* specimens to stressors simulating some of the major environmental challenges associated with the transition from water to land. Specifically, animals were exposed to UV radiation and visible light, hyper- and hypo-osmotic stress, hyperoxic and hypoxic conditions and chemosensory cues (see Methods section for more details).

DEGs were quantified and assigned to HOGs in order to infer their evolutionary origin (Table S15). Although a high proportion of DEGs were not assigned to HOGs (possibly due to DEGs either being species-specific or non-coding), the number of assigned genes allowed us to unveil differences in the mechanisms of response to stress between freshwater and terrestrial planarians. DEGs were assigned in 1,246 HOGs in *O. nungara* and 871 in *S. mediterranea*. Over-expressed genes were included in 865 HOGs and under- expressed genes in 381 HOGs in the terrestrial planarian. On the other hand, 138 HOGs contained over- expressed genes and 733 HOGs contained under- expressed genes, in the freshwater species.

The assignment of DEGs under stress to HOGs allowed us to infer their evolutionary origin (Fig. 4, Table S8). A high number of the HOGs including DEGs arose in the same nodes for both species, those between Tricladida and Geoplanoidea; 55% for *O. nungara* and 32.9% for *S. mediterranea*. However, the total number of shared HOGs between species was just ~ 7% ($n = 164$, Table S15, Fig. 4, Table S8). In order to evaluate the functions associated with stress–response, we compared the semantic similarity between enriched unique DEGs derived from each node by species (Fig. S1). We have observed high semantic similarity between GO terms of DEGs gained in Tricladida for both species. The shared GO terms in this node revealed cellular physiological flexibility against stress conditions, such as *muscle cell cellular homeostasis* (GO:0046716), *neuron cellular homeostasis* (GO:0070050), *unidimensional cell growth* (GO:0009826) and *canonical Wnt signaling pathway* (GO:0060070). DEGs gained in the Geoplanoidea node clustered apart for *O. nungara* and *S. mediterranea.* Intriguingly, for *O. nungara,* DEGs gained at Geoplanoidea clustered closer to *S. mediterranea* DEGs gained in more recent nodes in *S. mediterranea*, e.g. the node leading to Dugesiidae. This cluster most relevant GO terms were: *nucleic acid metabolic process* (GO:0090304), *hydrogen peroxide metabolic process* (GO:0042743), *response to cytokine* (GO:0097396) most likely in response to radiation and oxidant damage; *gravitaxis* (GO:0042332) and *response to pain* (GO:0048265) as putative behavioral sensory responses; *regulation of developmental process* (GO:0050793), *regulation of biomineral tissue development* (GO:0070167), *regulation of respiratory burst* (GO:0060263), *renal filtration* (GO:0097205), *amino sugar catabolic process* (GO:0046348), *glucosamine-containing compound metabolic process* (GO:1901073) probably against osmotic/oxic stresses. Finally, DEGs in *O. nungara* assigned to the Geoplanidae node clustered apart from every other node. The main GO terms present in this node were: *negative regulation of catalytic activity*

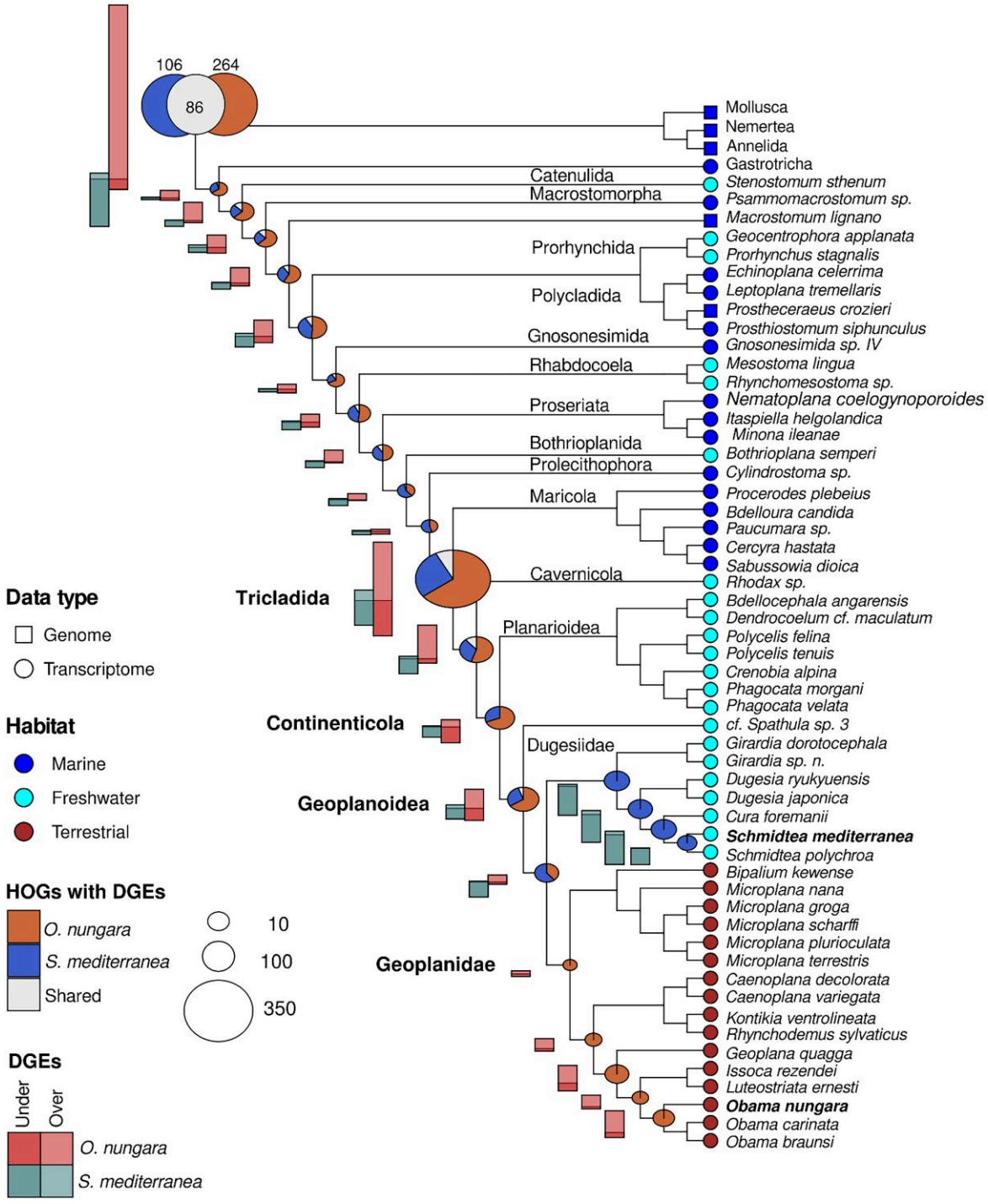

Fig. 4. Evolutionary origin of differentially expressed genes (DEGs) assigned to HOGs. Bars represent the number of over-expressed or under-expressed DEGs assigned to HOGs in *Schmidtea mediterranea* and *Obama nungara*. Pie charts represent the proportion of HOGs with assigned DEGs in each species and shared by both species. The size of the pie charts is proportional to the number of HOGs except for the most basal node where the information is shown in a Venn diagram.

(GO:0043086), *meristem development (*GO:0048507, as proxy of neoblasts development) *and anatomical structure arrangement* (GO:0048532). Suggesting, again, the gain of flexibility in the body plan to face stress conditions.

As regards the direction of change in gene expression, we observed a polarized pattern between species. While *S. mediterranea* showed that most DEGs were under-expressed, most DEGs in *O. nungara* were over-

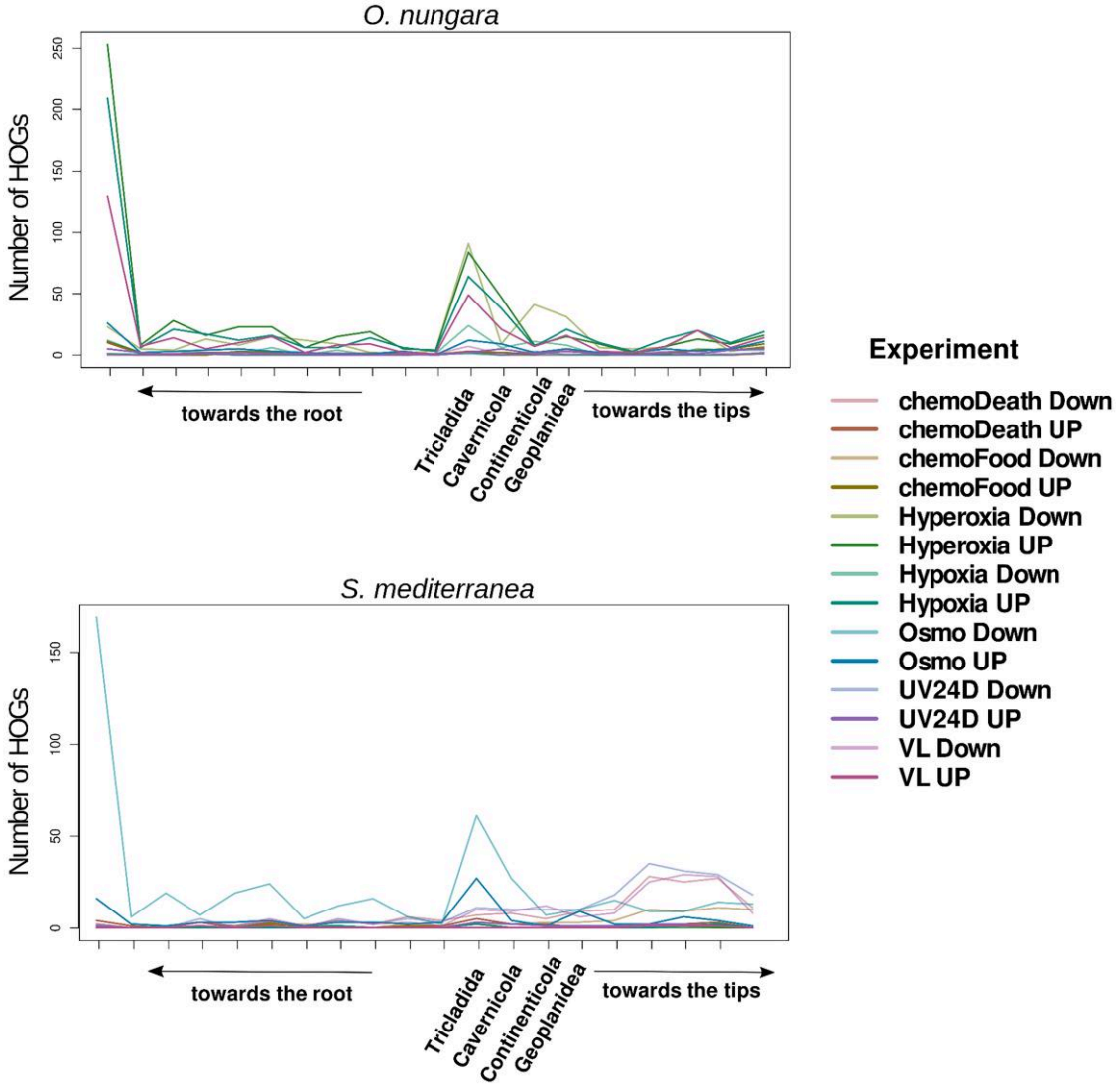

Fig. 5. Number of HOGs with DEGs assigned by node by experiment for *O. nungara* and *S. mediterranea* independently.

expressed (Fig. 4), suggesting distinct regulatory strategies in response to environmental stress. In the case of *O. nungara,* most of DEGs assigned to HOGs were detected as over-expressed in hyperoxia and hypoxia experiments. On the other hand, in the case of *S. mediterranea* most DEGs assigned to HOGs were under-expressed in the osmoregulation experiment (Fig. 4, Fig. 5).

No shared HOGs with DEGs were found in the same experiment and in the same direction of regulation between both species. However, DEGs detected in different experiments and with an opposite direction of regulation between species (i.e. over- or under-expressed) were included in the same HOGs. Most of these genes were under-expressed in *S. mediterranea* under osmotic stress, whereas they were over-expressed in *O. nungara* when exposed to hypoxia, hyperoxia, visible light or osmotic stress (Table S15; Table S9). This pattern suggests that these genes are involved in fundamental stress–response pathways but have diverged in their regulation between terrestrial and freshwater planarians, reflecting the different environmental pressures. *O. nungara* may have developed more active stress–response mechanisms, likely reflecting the greater challenges posed by terrestrial habitats, whereas S. mediterranea may rely more on down-regulation of pathways.

DEGs Present in the Non-Stressed Proteomes

We later examined whether the identified DEGs were present in the proteome datasets of *O. nungara* and *S.*

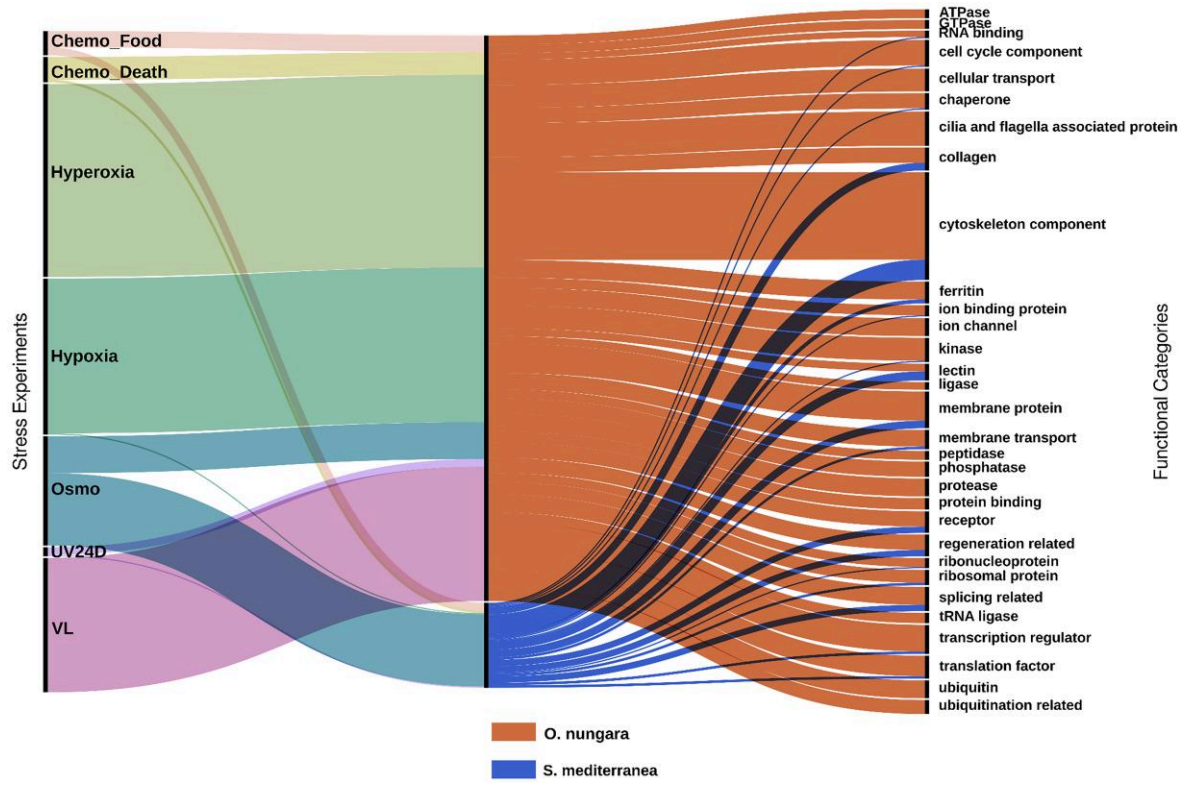

Fig. 6. Functions assigned to DEGs present *O. nungara* (brown) and *S. mediterranea* (blue) basal proteomes.

mediterranea (*DEGs-prot* hereafter). Among the DEGs assigned to HOGs, 351 out of 1358 were detected through proteomics in *O. nungara* under control conditions, and 129 out of 955 in *S. mediterranea* (Table S14). Phylostratigraphy analysis revealed that most observed DEGs-prot originated in Tricladida (Table S10).

The functional characterization of DEGs-prot showed a different functional background between the freshwater and terrestrial species, with a broader diversity of functions observed in *O. nungara* (Fig. 6, Table S11). Proteins involved in the stress–response in *O. nungara* comprised a wider range of general functions, including RNA binding, cytoskeleton components, cell cycle regulation, membrane transport, transcription regulation, ion channels, and protein ubiquitination. In contrast, the diversity of stress–response functional classes was lower in *S. mediterranea*. Most frequent protein functions in both species included cytoskeleton and membrane-associated components (Fig. 6).

A deeper look into the functional categories revealed that one of the most frequent in both species were the Ankyrin domain-containing proteins. Ankyrins are a family of adaptor proteins that link integral membrane proteins with the submembranous actin/β-spectrin cytoskeleton (Cunha and Mohler 2009). For example, the *TRPA1* ion channel, key in the transduction of nociceptive signals in *S. mediterranea*, has 14 N-terminal ankyrin repeats. *TRPA1* is known to be activated in response to stressful conditions involving also the production of H_2O_2/ROS in response to noxious heat, contributing to channel activation to mediate defensive responses (Arenas et al. 2017). Additionally, Ankyrins are related to the regeneration and cellular turnover in planarians and are part of the group of orthologous genes conserved across whole body regenerative animals (Chereddy and Makino 2024).

CEP135 was another protein detected in the proteomes and also detected as differentially expressed in both species. In agreement with enrichment results, it highlights the key role of epidermis in stimulus perception and stress protection in these organisms. *CEP135* is an essential protein for the basal body assembly in planarians (Azimzadeh and Basquin 2016). Actin components, which are also important elements of cytoskeleton, were also observed in the proteomes and DEGs patterns in both species. In planarians, actin filaments are actively involved in regeneration and cellular turnover (Pascolini et al. 1993; Liu et al. 2023). Additionally, other cytoskeleton components also observed in this analysis, such as microtubules associated with proteins are known to be fundamental in the regeneration process (Gambino et al. 2022). On the other hand, dynein, a protein also found

differentially expressed in both species, is fundamental for cilia motility, an important function related to locomotion as well as excretory and reproductive systems in planarians (Lesko and Rouhana 2020). Another protein detected here with high frequency in both species is collagen. In planarians, collagen proteins are an important component of muscular fibres and are the major structural components of the extracellular matrix (Cote et al. 2019) with an important role in regulating proliferation of neoblast during the regeneration process (Chan et al. 2021).

We also found a high number of lectins in both species. Lectins have been linked to terrestrialization in mollusks due to the expansions of lectin gene families in terrestrial mollusks and their key role in innate immunity (Aristide and Fernández 2023). Additionally, lectin domains have been predicted as toxic in *O. nungara* in a recent study, with a potential protection function in terrestrial planarians (García-Vernet et al. 2025). The existence of specific families of lectins described in free-living flatworms (Shagin et al. 2002) and our results point out the important role that these proteins can play in the evolution of Platyhelminthes, especially in complex processes of diversification such as the conquest of land.

Another protein found in the proteome and differentially expressed in both species is ferritin. This protein is the main intracellular iron storage in most eukaryotic cells (Plays et al. 2021), it is an important cytoprotective function against oxidative damage (Arosio et al. 2015). This protein also has an innate immunity role in both vertebrates and invertebrates and it is related to the regulation of apoptosis process (Ong et al. 2006). Additionally, ferritin has been described to be involved in response to anoxia in the interstitial mollusk *Littorina littorea* (Larade and Storey 2004). We identified ferritin differentially expressed in response to hyperoxia and other stressors in *O. nungara* and in osmoregulation and chemoreception experiments in *S. mediterranea* (Table S11). In planarians, ferritin is also known to be related to reproduction. In sexual freshwater planarians, ferritin in association with the Wnt pathway is fundamental for yolk-egg production (Vila-Farré et al. 2023). Given its plural role in mitigating oxidative stress and modulating innate immunity, ferritin emerges as a prime candidate for the adaptive toolkit required for life in terrestrial environments.

Several DEGs-prot were detected only in *O. nungara*. Among them, we can highlight *Cathepsin L*, a protein expressed in the intestinal system and probably related to food digestion, and also described as a biomarker in response to heavy metals contamination in *Dugesia japonica* (Ma et al. 2019). Here, we found this protein over-expressed in both chemoreception experiments as well as in the osmoregulation stress experiment. Another protein detected only in *O. nungara* was the E3 ubiquitin-protein ligase *HUWE1*. In general, the *HUWE1* ligase is a key regulator of DNA damage response, cell cycle control, apoptosis, autophagy and inflammation (Gong et al. 2020). Previously, it has been detected to be highly expressed in neoblasts with an important role in regeneration and stem cell regulation in *S. mediterranea* (Henderson et al. 2015). Here we detected this gene over-expressed in almost all experiments (chemoreception food, hyperoxia, hypoxia, osmoregulation, and visible light) only in *O. nungara*.

Regarding the expression response of the observed proteins to different abiotic stressors, the two species displayed distinct patterns. In *O. nungara*, most DEGs-prot were found in response to hyperoxia, hypoxia, and visible light treatments, while in *S. mediterranea*, the majority of DEGs-prot were associated with osmotic stress (Fig. 6). The evolutionary origin of most DEG-containing HOGs responding to hypoxia, hyperoxia, and visible light in *O. nungara*, and to osmotic stress in *S. mediterranea*, aligned with bursts of gene gain at the Tricladida node (Fig. 1, Table S10). These findings are consistent with the functional patterns observed for GO terms assigned to DEGs and suggest that stress–responses involve gene sets with potential adaptive relevance. The observation that a subset of these genes is constitutively expressed under non-stress conditions, together with the fact that most genes originated at the Tricladida node rather than at the Dugesiidae or Geoplanidae nodes, is compatible with a scenario of exaptation of ancestral molecular traits. In this framework, the "Tricladida blueprint" may have included a suite of genetic innovations that were subsequently co-opted and repurposed during the colonization of more challenging environments, including freshwater and terrestrial habitats.

Half of DEGs Gained Before the Diversification of Terrestrial Planarians are Under Directional Selection

To detect changes in selection pressures potentially related to habitat shift, we selected HOGs gained in Tricladida, Continenticola, and Geoplanoidea containing DEGs detected in *O. nungara* with a balanced representation of both terrestrial and non-terrestrial planarians (i.e. including at least 20% of the species in each habitat) resulting in a dataset comprising 444 HOGs. A statistically significant shift in directional selection in terrestrial planarians was detected in more than 50% of these HOGs (Fig. 7a, Tables S12 and S13).

In order to find putative adaptive traits associated to HOGs under directional selection we performed a functional characterization based on GO. The analysis was based on the enrichment of significant HOGs against the *O. nungara* full annotated proteome at each node of interest (Fig. 7c). Selected genes gained in

Fig. 7. Differentially expressed genes gained in Tricladida, Continenticola, and Geoplanoidea under directional selection. a) Proportion of HOGs under directional selection in the three lineages. The full bar represents the total number of DEG-containing HOGs with a balanced representation of both terrestrial and non-terrestrial planarians. b) Proportion of HOGs under directional selection by experiment and expression direction c) Top 20 enriched GO terms by *P*-value ($P < 0.01$) of DEGs under significant shifts of selection by node.

Tricladida showed GO terms related with the optimization of dormancy-related pathways such as *dauer larval development* (GO:0040024), *maintenance of dauer* (GO:0043055), *determination of adult lifespan* (GO:0008340) and *sleep* (GO:0030431). Even if these terms are originally from nematode models they act as a proxy for quiescence in planarians. Also Tricladida significant HOGs showed enriched GO terms related with

neoblast cell cycle regulation such as *regulation of mesodermal cell fate specification* (GO:0042661), *positive regulation of stem cell proliferation* (GO:2000648), *muscle cell cellular homeostasis* (GO:0046716) and *neuron cellular homeostasis* (GO:0070050). Moreover, selected functions from the Tricladida node were related with reproduction, for example GO terms: *reproductive process* (GO:0022414), *gonad development* (GO:0008406) and *regulation of vulval development* (GO:0040028). GO functions enriched in selected genes from Continenticola also showed relation with reproduction and neoblast cell cycle regulation such as *meiotic cell cycle* (GO:0051321), *regulation of mesodermal cell fate specification* (GO:0042661), *regulation of mitotic cell cycle embryonic* (GO:0009794) and *positive regulation of multicellular organism growth* (GO:0040018). In addition, Continenticola-selected HOGs showed GO terms related with environmental sensing such as *sensory processing (GO:0050893) and photoreceptor cell morphogenesis* (GO:0008594). Directional selection at the Geoplanoidea node is marked by functions involved in genomic stability and environmental sensing. Enriched genes showed GO terms for DNA repair such *positive regulation of DNA repair* (GO:0045739), *error-prone translesion synthesis* (GO:0042276), *mitotic DNA damage checkpoint signaling* (GO:0044773), *negative regulation of DNA damage checkpoint* (GO:2000002) and *positive regulation of DNA biosynthetic process* (GO:2000573). Furthermore, GO terms such as detection of *temperature stimulus involved in thermoception* (GO:0050960) and *negative gravitaxis* (GO:0048060), which can be related to terrestrial environment sensing, were also detected.

Discussion

Gene Gain and Genomic Exaptation as Key Drivers of Flatworm Terrestrialization

In this study, we unveiled a burst of gene gain detected much earlier than the origin of terrestrial planarians (i.e. in the internodes leading to Tricladida, Continenticola, and Geoplanoidea).

Gene gain is a known crucial driver of adaptation in animals, enabling them to survive and thrive in new environments by introducing novel functions and expanding genetic diversity. In various animal species, gene duplications and de novo gene gains have facilitated adaptations to environmental challenges. By generating genetic variation and providing the raw material for selection, gene gain enables organisms to develop specialized traits, such as enhanced stress tolerance, new metabolic pathways, or unique behavioral adaptations (Hull et al. 2017; Balart-García et al. 2023). Similar to our findings, Aristide and Fernández (2023) showed that multiple independent shifts from marine to freshwater and terrestrial environments in mollusks were facilitated by parallel expansions of ancient gene families. These expansions were predominantly involved in critical processes such as metabolism, osmoregulation, and defense mechanisms, showing the relevance of pre-existing genetic toolkits in adapting to new, non-marine environments. Their study also showed that while lineage-specific gene gains played a role in providing unique adaptations, much of the adaptive toolkit stemmed from pre-existing genomic features. Our results also suggest that for terrestrial planarians, older gene gains were especially important. The burst of newly gained and duplicated genes at the Tricladida node was significantly enriched in functions related to DNA and RNA metabolic processes and their regulation, as well as the regulation of the life cycle and cell cycle, and responses to environmental conditions. This situation suggests that land planarians may have relied heavily on repurposing these ancient genetic innovations to comply with the challenges of a terrestrial lifestyle. In addition, genes gained in transition nodes, i.e. those of Dugesiidae and Geoplanidae, also showed enrichment in GO terms related to adaptations to new habitats. However, semantic similarity analysis of these GO terms showed clear divergence between both lineages.

Divergent Stress–Response Mechanisms in Terrestrial and Freshwater Planarians

To assess whether the putative adaptive traits identified in the gene repertoire evolution analyses are functionally active in extant species, we exposed *O. nungara* and *S. mediterranea* to stressors designed to simulate key environmental challenges associated with the transition from aquatic to terrestrial habitats. Our comparative analysis revealed differences in how these species respond to environmental stressors. In general, *O. nungara* showed a higher number of DEGs particularly when facing changes in oxygen concentrations (hypoxia and hyperoxia) and exposure to visible light. These stressors are among the most significant challenges in the adaptation to terrestrial life, where fluctuations in oxygen availability and increased exposure to light and UV radiation required the presence of robust protective mechanisms. The extensive set of DEGs t under these conditions belonging to Tricladida-gained HOGs agrees with the idea of the repurpose of pre-existing genetic tools to develop new traits suited to comply with terrestrial challenges. On the other hand, *S. mediterranea* displayed a stress–response machinery predominantly reactive under osmoregulatory stress. Again, most of these DGEs corresponded to Tricladida-gained HOGs,

suggesting a different repurposing of a pre-existing toolkit to comply with freshwater challenges. These periods of genomic expansion may have provided a reservoir of genetic material that could be differentially exapted or co-opted in each lineage to meet specific environmental challenges. In *O. nungara*, genes initially involved in basic cellular processes may have been repurposed to develop enhanced responses to oxygen fluctuations and light exposure, conferring a selective advantage in terrestrial ecosystems. Similarly, in *S. mediterranea*, existing genes may have been co-opted to improve osmoregulatory functions, essential for maintaining homeostasis in freshwater environments. The hypothesis of genomic exaptation and/or co-option is in line with the hypothesis of the existence of exaptive traits that were present in the Tricladia lineage before habitat transitions. Those exaptations include: internal fertilization, the existence of hard shell cocoons enveloping the eggs, the presence of direct development with no larval forms and the production of mucus (Sluys 2019). In our stress experiments, the most relevant shared GO terms in DEGs gained in the Tricladida were related with cell fate and cell homeostasis (e.g. GO:0046716, GO:0070050, GO:0009826, GO:0060070). In addition, GO terms in the Geoplanoidea and Dugesiidae cluster showed strong specific signs of response to radiation and oxygen stresses (e.g. GO:0090304, GO:0042743, GO:0097396), signs of behavioral sensory responses (e.g. GO:0042332, GO:0048265), signs of osmotic and water balance response including mucus production (e.g. GO:0097205, GO:0046348, GO:1901073) and also signs of cell fate (GO:0050793, GO:0070167). Finally, DEGs in *O. nungara* assigned to the Geoplanidae transition showed, again, most relevant GO terms related with cell fate and cell homeostasis (e.g. GO:0048507, GO:0048532). These observations suggest that the adaptation to freshwater and terrestrial environments was not merely a matter of utilizing ancestral traits, but of physiologically and structurally specializing the ancestral body plan. While the tricladid common ancestor provided a "land-ready" phenotypic foundation, these traits were probably refined at transition nodes.

HOGs including DEGs showed minimal overlapping between the two species, and also showed opposite patterns of expression. While most DEGs in *S. mediterranea* were under-expressed in response to stress, *O. nungara* exhibited a predominance of over-expressed DEGs. This pattern suggests that these genes are involved in fundamental stress–response pathways but have diverged in their regulation between freshwater and terrestrial planarians, reflecting differing environmental pressures. Interestingly, shared HOGs with DEGs in both species were found to be over-expressed in *O. nungara* and under-expressed in *S. mediterranea* in response to different stressors, such as hypoxia, hyperoxia, and osmotic stress. This pattern implies that these genes are involved in core stress–response pathways but have diverged in their regulatory roles, reflecting the distinct evolutionary pressures faced by freshwater and terrestrial environments.

Key Proteins in Stress–Response and Adaptation to Terrestrial Life in Planarians

Again, following a top-down approach, we experimentally corroborated the expression at the protein level of the observed DEGs in HOGs in both species and evaluated their functions. Most DEGs with observed basal protein expression in non-stressed animals were related to cytoskeleton components, cell cycle, and regeneration processes, in agreement with previous results suggesting body plan specialization against different stressors. This links with the fact that the tissue regeneration machinery, specifically described in planarians, is key in the response to stress.

In planarians, the differentiated cells do not divide mitotically, and the neoblasts are the only dividing cells (Morita and Best 1974; Newmark and Sánchez Alvarado 2000). Thus, the cellular turnover driven by neoblasts differentiation is the only way to maintain tissue homeostasis, cell number and body size in these animals, allowing the regulation of body growth depending on environmental conditions (González-Estévez and Saló 2010; González-Estévez et al. 2012; Felix et al. 2019). The continuous cell turnover implicates the periodic elimination of selected differentiated cells and their replacement by the differentiated descendant of adult stem cells that become functional (Pellettieri and Sánchez Alvarado 2007). The differentiation process implicates a series of cellular changes under strong genetic regulation that guarantees the proper maintenance of the planarian body plan. Our findings suggest that the machinery governing cellular turnover may play a critical role in responding to environmental stress and adapting to new habitats in planarians, as mentioned in the above section.

Worth mentioning is the fact that we identified ferritin differentially expressed in response to hyperoxia and other stressors in *O. nungara* and in osmoregulation and chemoreception experiments in *S. mediterranea*. In planarians, ferritin is known to be related to reproduction. In freshwater planarians, ferritin in association with the Wnt pathway is fundamental for yolk-egg production (Vila-Farré et al. 2023). Beyond reproduction, Wnt signaling is a master regulator of axial polarity, regeneration, and neoblast-mediated cell turnover (Reddien 2018). Moreover, in our experiments the GO term *canonical Wnt signaling pathway* (GO:0060070) was within the most relevant terms in Tricladida-gained DEGs in both

species. Again, this situation points to the importance of the regulation of cell turnover and plasticity as adaptive mechanisms in response to environmental stressors.

Additionally, the expression of *HUWE1,* E3 ubiquitin ligase protein known to have an important role in regeneration and neoblast regulation in planarias (Henderson et al. 2015), specifically in *O. nungara* also connects the regenerative capabilities with stress–response mechanisms associated with life on land.

Directional Selection of Genes Gained Before the Transition of Terrestrial Planarians

In a last top-down step we evaluated if DEGs under stress conditions in *O. nungara* were directionally selected. Results showed that ca. half of DEGs that arose prior to the origin and diversification of terrestrial planarians (i.e. genes gained in Tricladida, Continenticola, and Geoplanoidea) were under directional selection. The selected genes in Tricladida were still enriched in functions related with cell fate, neoblast cell cycle, and environmental sensing.

The identified directional selection patterns provide a robust validation of our multi-omics findings, converging on a single evolutionary narrative. The functions observed in the selected HOGs are consistent with those enriched in stress–response DEGs, which were predominantly acquired at the Tricladida node, suggesting that the genetic "raw material" for land colonization was already present in the common ancestor. Finally, the fact that most of these adaptive traits correspond to proteins constitutively detected in our proteomic profile confirms that this "Tricladida blueprint" is not merely a genomic potential but a functional reality. Collectively, this evidence supports a model of evolutionary exaptation, where an ancestral suite of tools for DNA repair, metabolic control, and structural integrity was later refined by selection to master the transition from aquatic to terrestrial environments.

Final Remarks

This study unveils the genomic mechanisms underlying planarian terrestrialization, offering significant insights into the evolutionary processes that enabled this transition. We identified a burst of gene gain occurring earlier in flatworm evolutionary history—prior to the emergence of land planarians. This genomic innovation, enriched in regulatory and stress–response functions, likely provided a foundational genetic toolkit that was later co-opted through genomic exaptation to facilitate adaptation to terrestrial environments. This illustrates how genomic exaptation can be more important than gene gains occurring precisely at the node where phenotypic changes are observed. The ability to utilize and readapt existing genes underscores the significance of gene regulation and functional divergence in evolutionary adaptation. Overall, this study provides the first comprehensive genome-wide evidence elucidating the mechanisms driving flatworm terrestrialization. Our findings not only expand our understanding of land colonization in flatworms but also align with broader evolutionary patterns observed across other taxa. Specifically, we reveal that multiple evolutionary strategies, including early gene gains, expansion of conserved gene families, directional selection, and the repurposing of existing genes, have been instrumental in overcoming the environmental challenges of transitioning from aquatic to terrestrial habitats. The fact that significant genomic innovations occurred prior to the phenotypic transition highlights the importance of pre-existing genetic diversity and the role of exaptation in evolutionary innovation. Collectively, this research contributes to a deeper understanding of the genomic factors underlying large-scale evolutionary convergence during habitat transitions, elucidating how animals have repeatedly adapted to life on land throughout evolution.

Materials and Methods

An overall representation of the evolutionary context of planarian evolution and analytical workflow used in this study is shown in Fig. 8.

Sampling and Sequencing of Terrestrial Planarians

We hand-collected 15 species of terrestrial planarians from Spain and Brazil. Sampled species, coordinates and localities are available at Table S1. Samples were collected and stored in RNALater at <20 °C until processed. RNA extraction was performed using the TRIzol® reagent (Invitrogen, USA) method following the manufacturer's instructions and adding 5 µg of RNase-free glycogen as a carrier to the aqueous phase for low input samples. Concentration of all samples was assessed by Qubit RNA BR Assay kit (Thermo Fisher Scientific). Samples were subjected to Illumina's TruSeq Stranded mRNA library preparation kit and sequenced on a NovaSeq 6000 (Illumina, 2 × 150 bp) for a 6 Gb coverage.

Gene Repertoire Evolutionary Dynamics

In order to infer patterns of gene gain, duplication, and loss potentially associated with flatworm terrestrialisation, we assembled a dataset of proteomes representing the main lineages of Platyhelminthes. In addition to publicly available datasets from MATEdb2 (Martínez-Redondo et al. 2024a), we included 15 terrestrial planarian species newly collected for this study (Table S1; Fig. 8a). RNA from these species was sequenced and assembled

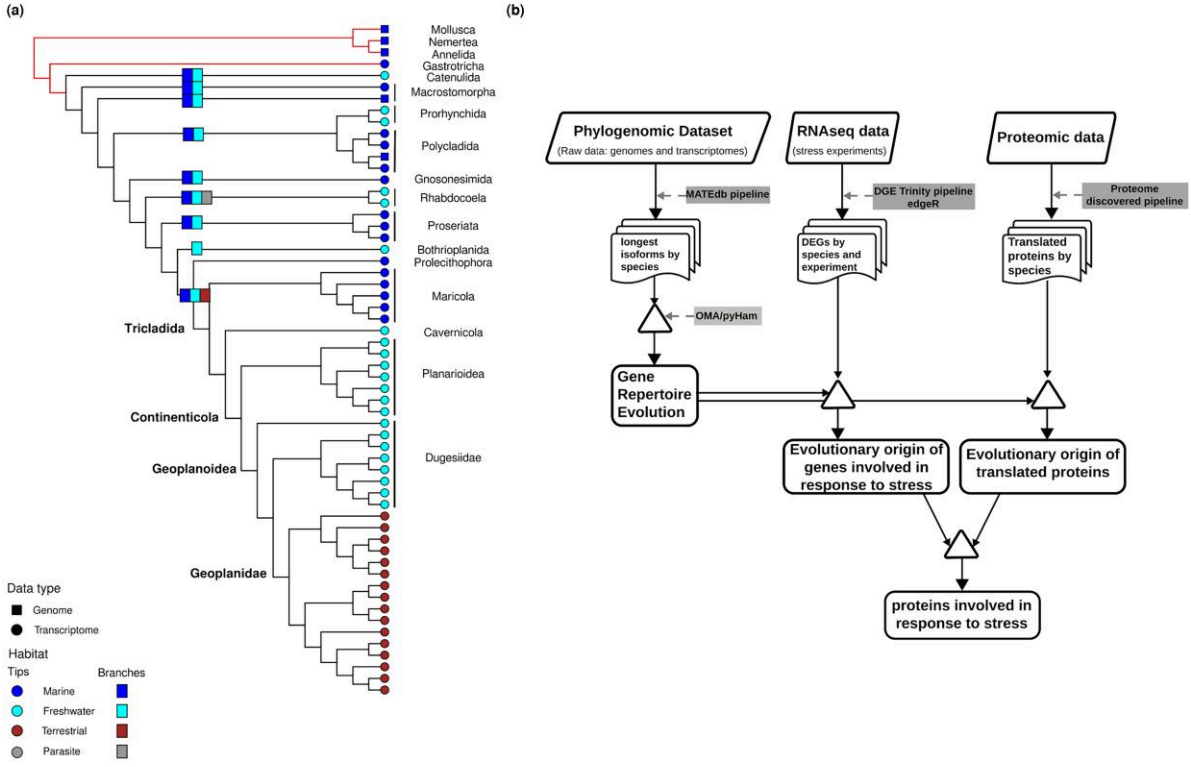

Fig. 8. Multi-omics analytical workflow of this study. a) Schematic species tree showing the phylogenomic dataset. Outgroups are represented with red branches and Platyhelminthes with black ones. Habitats reported for each flatworm lineage are represented by squares in the branches, and the habitat of each species is indicated by circles in the tips (dark blue: marine, light blue: freshwater, dark brown: terrestrial,). Tree topology and habitats extracted from Laumer et al. (2015). The type of data is represented by squares in the terminals (transcriptome, grey, or genome, yellow). b) Workflow followed in this study integrating genomic, transcriptomic and proteomics data.

following the MATEdb2 pipeline, and proteomes were generated by extracting the longest isoform per gene, consistent with the MATEdb2 standard. These newly generated datasets were then combined with additional platyhelminth proteomes retrieved from MATEdb2, with particular emphasis on dense sampling within Tricladida, the only order comprising both freshwater and terrestrial representatives.To polarize the comparative genomics results, we used a topology reflecting our current understanding of platyhelminth interrelationships (Laumer et al. 2015; Vila-Farré et al. 2023; Benítez-Álvarez et al. 2025). Per-gene protein sequences were obtained from MATEdb database or were obtained as previously explained in Martínez-Redondo et al. (2025). For each species, the longest isoform per gene was parsed for orthology inference. The predicted proteomes were subject to orthology inference using OMA v2.6 (Altenhoff et al. 2019) to infer HOGs, defined as set of genes that have descended from a single common ancestor within a taxonomic range of interest (Altenhoff et al. 2013). pyHam (Train et al. 2019) was used to infer gene repertoire evolutionary dynamics across the Platyhelminthes phylogeny based on the previously-inferred HOGs following the analytical strategy described in (Vargas-Chavez et al. 2024). Gene gains were defined as HOGs whose taxonomic origin (root) was precisely assigned to the ancestral node of interest by pyHam, indicating an absence in parental lineages. To ensure the robustness of these inferred gains, we applied a completeness filter: only HOGs represented in at least 20% of the extant species descending from the node were retained for downstream analysis. Subsequent curation was performed to differentiate between true gains, duplications, and losses, ensuring that genes previously present in parental nodes were not misclassified as gains due to incomplete lineage sampling. Specifically, to account for potential annotation gaps in the outgroups, we implemented a recovery step: HOGs initially flagged as "lost" by the automated pyHam reconstruction were re-examined and retained if they maintained a significant representation in the descending lineages, defined as being present in at least 20% of the species within the node with at least one sequence detected. This rigorous filtering ensured that our final set of gained HOGs excluded ancestral genes that might have been erroneously omitted from parental nodes due to technical artifacts.

Proteomics Data and Protein Identification

Proteomic raw data of *O. nungara* and *S. mediterranea* was obtained from García-Vernet et al. (2025). Briefly, proteins were extracted by using TRIzol® reagent (Invitrogen, USA) following (Simões et al. 2013). The final protein pellets were resuspended in 25 µl of resuspension buffer (1% SDS and 8 M Urea in Tris-HCl 1 M, pH 8). Samples were digested and desalted for the LC-MS/MS analysis and analyzed using a Orbitrap Eclipse mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) coupled to an EASY-nLC 1200 (Thermo Fisher Scientific [Proxeon], Odense, Denmark). The Proteome Discoverer software suite (v.2.5, Thermo Fisher Scientific) and the Mascot search engine (v2.6.2) were used for preprocessing the data and identifying the peptides. To perform the peptide identification, we used as reference databases the longest isoforms of *O. nungara* and *S. mediterranea*, as previously described for the gene repertoire dynamics analysis analyses. For constructing the final proteomic dataset, we only included proteins identified with high confidence levels and designated as the master protein within a protein group. We defined a protein group as a set of proteins containing at least two unique peptides. Peptide-spectrum matches were filtered at a false discovery rate (FDR) of 1% using a target–decoy approach. For protein-level inference, only high-confidence master proteins supported by at least two unique peptides were retained. For quantitative comparisons, normalized protein abundances were statistically evaluated using the built-in statistical framework in Proteome Discoverer, applying multiple-testing correction (Benjamini–Hochberg) and considering proteins significant at an adjusted P-value < 0.05. Detailed experimental and statistical methods can be found in García-Vernet et al. (2025).

Differential Gene Expression Involved in Response to Abiotic Stress

With the goal of pinpointing the gene repertoire used in response to abiotic stress and explore its evolutionary origin, specimens of *Obama nungara* (terrestrial planarian, family Geoplanidae, $n = 30$) and *Schmidtea mediterranea* (freshwater planarian, family Dugesiidae, $n = 36$) were subjected to several stressors.

Natural light—Terrestrial species are regularly exposed to visible light as part of their daily environmental conditions, while many aquatic invertebrates typically experience limited or no direct light exposure. This ecological contrast suggests that light may act as a more frequent and biologically relevant stressor in terrestrial environments. To assess differential responses to visible light, we exposed individuals to ambient natural light for 15 min prior to dissection and flash freezing (experiment code: *VL*).

UV radiation—Ultraviolet (UV-B) radiation (280 to 315 nm) is a potent genotoxic stressor that induces DNA damage. To investigate UV-induced stress–responses specimens were exposed to UV-B radiation (302 nm) at 8.48 W/m^2 using a UVM-28 EL Series UV lamp (Analytik Jena US) for 2 min. Following exposure, animals were allowed to recover during 24 h in darkness (experiment code: *UV24D*).

Osmoregulation experiments—Successful transition from marine to terrestrial environments required major physiological adaptations in water balance regulation. Among the most critical challenges for terrestrial animals is the risk of dehydration due to increased evaporative water loss in air. To assess desiccation tolerance in *O. nungara*, we reduced ambient relative humidity to 29–30% RH and placed individuals in petri dishes where surface moisture was removed using filter paper. Behavioral and morphological indicators of dehydration, such as reduced activity or loss of cuticular shine, were carefully monitored, and the experiment was terminated once early signs of desiccation became apparent/15 to 50 min). Osmoregulatory capacity was evaluated in *S. mediterranea* by exposing specimens to seawater for 30 s. (experiment code: *Osmo*).

Oxygen pressure experiments—One of the major physiological barriers to terrestrialisation is the shift from extracting oxygen in water to breathing air, which involves coping with different oxygen solubility and availability. To simulate this transition, we exposed animals to controlled oxygen extremes that reflect the lowest and highest oxygen concentrations known to support life in aquatic and terrestrial environments. Hypoxic conditions were generated using a HypoxyLab™ Hypoxia chamber (Oxford Optronix). For hyperoxia, pure oxygen (ClearO2, UK) was sprayed into airtight glass containers until the target oxygen concentration was reached. Oxygen levels were continuously monitored using an OXY-1 SMA Single Channel Fiber Optic Oxygen Transmitter (PreSens Precision Sensing GmbH) and Planar Oxygen-Sensitive Spots affixed inside the experimental chambers. In hypoxia experiments, *O. nungara* individuals were placed directly into airtight glass containers, which were then inserted into the HypoxyLab to achieve oxygen concentrations between 13% to 15% O_2. For *S. mediterranea* the containers were pre-filled with freshwater and the oxygen concentration was reduced to 8% O_2 before animals were introduced. For hyperoxia, *O. nungara* individuals were transferred into pre-equilibrated airtight chambers after spraying with pure oxygen to achieve 40% to 42% O_2. Oxygen levels were maintained by additional injections as needed. For *S. mediterranea* water was pre-oxygenated to reach 42% to 45% O_2 before animals were added. All exposures were conducted for

20 min (experiment codes: *Hypoxia* or *Hypo* and *Hyperoxia* or *Hyper*).

Exposure to chemical cues—The transition from aquatic to terrestrial environments necessitated profound changes in sensory systems, particularly in the detection and processing of chemical cues. While chemosensation is a universal mechanism across animals, the nature of chemical stimuli differs markedly between water and air, requiring lineage-specific adaptations in receptor types and signaling pathways. To investigate how planarians respond to ecologically relevant chemosensory cues, we designed experiments using both attractive and aversive stimuli. For attractive cues, individuals were exposed to beef liver as a food source (experiment code: *CF, chemo_food*). To simulate warning signals, we used decomposition-associated fluids. In brief, one conspecific individual per species was euthanised and allowed to decompose in phosphate-buffered saline (PBS) or freshwater depending on the organism's native environment. The resulting fluid (dead conspecific) was used as an aversive chemosensory stimulus (experiment code: *CD, chemo_death*). During trials, live specimens were placed individually in petri dishes—either dry or containing a small volume of water—while the stimulus (food or decomposed fluid) was introduced at the center. Behavioral responses, such as approach or avoidance, were monitored in real time.

Control samples were directly processed without any additional treatment. All samples were left in starvation for at least 24 h before any experiment to reduce gut content. Three to five biological replicates per species and experiment were included as: controls $n = 4$, VL $n = 4$, UV24D $n = 3$, Osmo $n = 3$, CF $n = 3$, CD $n = 3$, Hyper $n = 5$, Hypo $n = 5$ for *O. nungara* and controls $n = 5$, VL $n = 4$, UV24D $n = 3$, Osmo $n = 4$, CF $n = 5$, CD $n = 5$, Hyper $n = 5$, Hypo $n = 5$ for *S. mediterranea*. After the experiments, the specimens were flash frozen in liquid nitrogen and kept <−70 °C until further processing. mRNA libraries were built and sequenced as indicated above for the terrestrial planarians. Quality assessment of the raw Illumina RNA-seq reads of *O. nungara* and *S. mediterranea* of all the experiments was performed using FastQC v0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc). Adapters and low-quality bases were trimmed using Trimmomatic v0.39 (MINLEN: 75, SLIDINGWINDOW: 4:15, LEADING: 10, TRAILING: 10, AVGQUAL: 30) (Bolger et al. 2014) and the trimmed RNA-seq reads were also assessed with FastQC before further analysis. For both species, a de novo reference transcriptome assembly generated with Trinity v2.11.0 (Haas et al. 2013) using as input a combination of the RNA-seq reads of the stress experiments and used further for differential gene expression analysis. The indexing and quantification of the transcripts was performed using Salmon v1.5.2 (Patro et al. 2017) for each sample (Table S7). For each gene, the transcripts per million (TPM) value was calculated and the perl script *align_and_estimate_abundance.pl* included in the Trinity package was used to generate the counts and trimmed mean of M-values (TMM) scaling normalized expression matrices. Differential gene expression analysis was conducted using the perl script run_DE_analysis.pl also included in Trinity toolkit and the edgeR (Robinson et al. 2010) package. Finally, the Benjamini-Hochberg correction method (Benjamini and Hochberg 1995) was applied with the script analyze_diff_expr.pl with cutoffs of adjusted *P*-value at 0.001 (FDR) and a four-fold change.

Evolutionary Origin of the Gene Repertoire Involved in Response to Abiotic Stress

We assigned the DEGs in the stress experiments and the proteins detected by proteomics to HOGs. We then reconciled the HOGs with the Platyhelminthes species tree to infer the evolutionary origin of these genes following a phylostratigraphic reconstruction (i.e. in which node each HOG originated) based on pyHAM (Train et al. 2019) results.

Detection of Shifts in Selection Regimes in Genes Involved in Response to Abiotic Stress

We have chosen the HOGs which included DEGs in *O. nungara under stress* experiments and that arose earlier to the diversification of terrestrial planarians (i.e. at Tricladida, Continenticola, and Geoplanoidea nodes), to detect selective pressure changes during transition to terrestrial habitat used *Pelican* v.1.0.8 (Duchemin et al. 2023). Pelican follows the approach of Parto and Lartillot (2018) to pinpoint sites associated with trait changes by analyzing shifts in amino acid selection. The software evaluates the hypothesis that protein substitution processes are driven by environmental conditions. Using a Likelihood Ratio Test (LRT), it contrasts a uniform model of evolution across a phylogeny with a flexible model where terrestrial and non-terrestrial branches are governed by different substitution rules. We selected HOGs with a balanced representation of both terrestrial and non-terrestrial species (i.e. HOGs to be present in at least 20% of the species in each group). We coded the non-terrestrial species as background character and terrestrial species as foreground character. Initial quality control was performed with PREQUAL to mask non-homologous regions, followed by alignment with MAFFT and curation with trimAl (-automated1). Phylogenetic trees were generated using IQ-TREE (LG4M model) and reconciled with the species tree using Treerecs. These reconciled trees were then annotated via parsimony to distinguish

between terrestrial and non-terrestrial lineages. Pelican was subsequently run in scan discrete mode, contrasting a null model of uniform evolution against a trait-dependent model to pinpoint sites associated with the transition between environments. Because Pelican's output includes a set of *P*-values for each site in the alignment, we used these site-level *P*-values to obtain gene-level prediction using the Gene-wise Truncated Fisher's method (Duchemin et al. 2023) considering the best k = 10 *P*-values in each alignment. Additionally, we applied the adjustment method Benjamini and Hochberg (1995) (Benjamini and Hochberg 1995) (BH) based on the FDR.

Functional Annotation and Enrichment Analyses

The putative function of genes identified as relevant in each of the analyses (i.e. gene repertoire evolution, differential gene expression and directional shift selection) was explored through functional annotation and enrichment. For functional annotation, we used FANTASIA (Martínez-Redondo et al. 2024), since homology-based methods did not recover GO terms for a high percentage of the genes of interest (Table S2). On average, 98.96% of sequences were annotated with FANTASIA at the GO level, in contrast to a mean number of 67.16% sequences annotated with eggNOG-mapper at GO level and 48.68% at gene name level. In order to extract functional insight from these annotations, enrichment analyses for GO terms were performed In brief, the *Fisher*'s exact test was used along with the the *elim* algorithm as implemented in the topGO package with default parameters (i.e. *P*-value cutoff = 0.01 for *elim* method) (Rahnenfuhrer 2025). topGo takes into account the DAG (Directed Acyclic Graph) of the GO to account for redundancy. Specifically, the *elim* algorithm computes the significance of a node conditional on the significance of its children and removes all genes that are annotated to a significantly enriched node from all its ancestors (Alexa et al. 2006). To visualize the significantly enriched GO terms, we first calculated the Wang semantic similarity matrix and clustered significant terms using the K-means algorithm as implemented in the simplifyEnrichment R package. Then, to summarize the function of each cluster, we obtained their representative GO term by selecting the GO term with the lowest closeness value (the "farthest") in the subset semantic similarity network for that cluster. The background list of Go terms used for the Fisher tests is described in Table S3.

Moreover, to evaluate functional divergence between nodes we computed "Biological Process" GO terms semantic similarity between nodes. We created a custom script that used *pygosemsim* (https://github.com/mojaie/pygosemsim) to estimate the Wang GO semantic similarity matrix between groups of GO terms using the Best-Match Average (BMA) method. Later, we used the R package constellatoR v0.1.0 (https://github.com/MetazoaPhylogenomicsLab/constellatoR) to cluster nodes based on the computed similarity matrices. The affinity propagation cluster algorithm was used to identify groups exhibiting high functional similarity. To visualize these results, we used PCoA plots where the functional clusters are represented as polygons. Additionally, we used constellatoR v0.1.0 to identify the most relevant GO terms associated with each node and obtained the three GO terms with the lowest level belonging to each GO and retrieved their description.

Finally, we characterized the DEGs detected by proteomics in both species using the best hits obtained with BLAST (Camacho et al. 2009) these sequences against the no-redundant database (Available from: https://www.ncbi.nlm.nih.gov/gene/last available update on May 1st 2024).

Data Availability

The raw data for the transcriptomes generated in this study can be found at the European Nucleotide Archive with Project IDs: PRJEB81501, PRJEB108593 and PRJEB108587. Raw data for the proteomics data set can be found at the Proteomics Identifications Database with Project ID: PXD074775. Both at https://www.ebi.ac.uk.

A repository containing additional input, intermediate and supporting files can de found here: https://github.com/MetazoaPhylogenomicsLab/Benitez-Alvarez_et_al_2025_Gene_repertoire_evolution_land_planarians.

Acknowledgments

We thank Adrian Altenhoff for guidance and discussions about pyHAM. Gemma I. Martínez-Redondo kindly helped us collect *Microplana* specimens, and shared scripts to improve the GO enrichment treeplots. RG-V acknowledges the support of the Margarita Salas grant funded by the Spanish Ministry of Universities and the European Union Next Generation EU/PRTR. 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"), the European Research Council (this project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 948281)), the OSCARS project (funding from the European Commission's Horizon Europe Research and Innovation programme under grant agreement no.

101129751) and the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement de la Generalitat de Catalunya (AGAUR 2021-SGR00420). FÁF-Á was supported by a Beatriu de Pinós fellowship from Secretaria d'Universitats i Recerca del Departament de Recerca i Universitats of the Generalitat de Catalunya (Ref. BP 2021 00035) and a Ramón y Cajal fellowship (Ref. RYC2023-043494-I) funded by MCIN/AEI/10.13039/501100011033 and FSE+. This research was supported by the Spanish government through the "Severo Ochoa Centre of Excellence" accreditation (CEX2019-000928-S). We also thank Centro de Supercomputación de Galicia and CSIC for access to computer resources (CESGA and DRAGO, respectively).

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Literature Cited

Alexa A, Rahnenführer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics. 2006:22:1600–1607. https://doi.org/10.1093/bioinformatics/btl140.

Altenhoff AM, et al. OMA standalone: orthology inference among public and custom genomes and transcriptomes. Genome Res. 2019:29:1152–1163. https://doi.org/10.1101/gr.243212.118.

Altenhoff AM, Gil M, Gonnet GH, Dessimoz C. Inferring hierarchical orthologous groups from orthologous gene pairs. PLoS One. 2013:8:e53786. https://doi.org/10.1371/journal.pone.0053786.

Arenas OM, et al. Activation of planarian TRPA1 by reactive oxygen species reveals a conserved mechanism for animal nociception. Nat Neurosci. 2017:20:1686–1693. https://doi.org/10.1038/s41593-017-0005-0.

Aristide L, Fernández R. Genomic insights into mollusk terrestrialization: parallel and convergent gene family expansions as key facilitators in out-of-the-sea transitions. Genome Biol Evol. 2023:15:evad176. https://doi.org/10.1093/gbe/evad176.

Arosio P, Carmona F, Gozzelino R, Maccarinelli F, Poli M. The importance of eukaryotic ferritins in iron handling and cytoprotection. Biochem J. 2015:472:1–15. https://doi.org/10.1042/BJ20150787.

Asano T, Seto Y, Hashimoto K, Kurushima H. Mini-review: an insect-specific system for terrestrialization: laccase-mediated cuticle formation. Insect Biochem Mol Biol. 2019:108:61–70. https://doi.org/10.1016/j.ibmb.2019.03.007.

Azimzadeh J, Basquin C. Basal bodies across eukaryotes series: basal bodies in the freshwater planarian *Schmidtea mediterranea*. Cilia. 2016:5:15. https://doi.org/10.1186/s13630-016-0037-1.

Balart-García P, et al. Parallel and convergent genomic changes underlie independent subterranean colonization across beetles. Nat Commun. 2023:14:3842. https://doi.org/10.1038/s41467-023-39603-1.

Benítez-Álvarez L, et al. Phylogeny and biogeography of the Cavernicola (platyhelminthes: tricladida)): relicts of an epigean group sheltering in caves? Mol Phylogenet Evol. 2020:145:106709. https://doi.org/10.1016/j.ympev.2019.106709.

Benítez-Álvarez L, et al. Disentangling the evolutionary history of terrestrial planarians through phylogenomics. Mol Phylogenet Evol. 2025:210:107381. https://doi.org/10.1016/j.ympev.2025.108371.

Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol. 1995:57:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Bi X, et al. Tracing the genetic footprints of vertebrate landing in non-teleost ray-finned fishes. Cell. 2021:184:1377–1391. https://doi.org/10.1016/j.cell.2021.01.046.

Birchler JA, Yang H. The multiple fates of gene duplications: deletion, hypofunctionalization, subfunctionalization, neofunctionalization, dosage balance constraints, and neutral variation. Plant Cell. 2022:34:2466–2474. https://doi.org/10.1093/plcell/koac076.

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics. 2014:30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Braasch I, Bobe J, Guiguen Y, Postlethwait JH. Reply to: subfunctionalization versus neofunctionalization after whole-genome duplication. Nat Genet. 2018:50:910–911. https://doi.org/10.1038/s41588-018-0163-3.

Camacho C, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009:10:421. https://doi.org/10.1186/1471-2105-10-421.

Campbell JW. Arginine and urea biosynthesis in the land planarian: its significance in biochemical evolution. Nature. 1965:208:1299–1301. https://doi.org/10.1038/2081299a0.

Chan A, Ma S, Pearson BJ, Chan D. Collagen IV differentially regulates planarian stem cell potency and lineage progression. Proc Natl Acad Sci U S A. 2021:118:e2021251118. https://doi.org/10.1073/pnas.2021251118.

Chereddy SCRR, Makino T. Conserved genes in highly regenerative metazoans are associated with planarian regeneration. Genome Biol Evol. 2024:16:evae082. https://doi.org/10.1093/gbe/evae082.

Cote LE, Simental E, Reddien PW. Muscle functions as a connective tissue and source of extracellular matrix in planarians. Nat Commun. 2019:10:1592. https://doi.org/10.1038/s41467-019-09539-6.

Cunha SR, Mohler PJ. Ankyrin protein networks in membrane formation and stabilization. J Cell Mol Med. 2009:13:4364–4376. https://doi.org/10.1111/j.1582-4934.2009.00943.x.

Domazet-Lošo M, Široki T, Šimičević K, Domazet-Lošo T. Macroevolutionary dynamics of gene family gain and loss along multicellular eukaryotic lineages. Nat Commun. 2024:15:2663. https://doi.org/10.1038/s41467-024-47017-w.

Duchemin L, Lanore V, Veber P, Boussau B. Evaluation of methods to detect shifts in directional selection at the genome scale. Mol Biol Evol. 2023:40:msac247. https://doi.org/10.1093/molbev/msac247.

Felix DA, Gutiérrez-Gutiérrez Ó, Espada L, Thems A, González-Estévez C. It is not all about regeneration: planarians' striking power to stand starvation. Semin Cell Dev Biol. 2019:87:169–181. https://doi.org/10.1016/j.semcdb.2018.04.010.

Fernández R, Gabaldón T. Gene gain and loss across the metazoan tree of life. Nat Ecol Evol. 2020:4:524–533. https://doi.org/10.1038/s41559-019-1069-x.

Gambino G, et al. Microtubule-associated protein 1B is implicated in stem cell commitment and nervous system regeneration in planarians. PLoS One. 2022:17:e0278966. https://doi.org/10.1371/journal.pone.0278966.

García-Vernet R, et al. 2025. A proteo-transcriptomic investigation of toxin evolution in planarians and their role in flatworm

terrestrialization [preprint]. bioRxiv. 632098. https://doi.org/10.1101/2025.01.09.632098.
Gong X, et al. The structure and regulation of the E3 ubiquitin ligase HUWE1 and its biological functions in cancer. Invest New Drugs. 2020:38:515–524. https://doi.org/10.1007/s10637-020-00894-6.
González-Estévez C, Felix DA, Rodríguez-Esteban G, Aboobaker AA. Decreased neoblast progeny and increased cell death during starvation-induced degrowth. Int J Dev Biol. 2012:56:83–91. https://doi.org/10.1387/ijdb.113452cg.
González-Estévez C, Saló E. Autophagy and apoptosis in planarians. Apoptosis. 2010:15:279–292. https://doi.org/10.1007/s10495-009-0445-4.
Haas BJ, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013:8:1494–1512. https://doi.org/10.1038/nprot.2013.084.
Henderson JM, et al. Identification of HECT E3 ubiquitin ligase family genes involved in stem cell regulation and regeneration in planarians. Dev Biol. 2015:404:21–34. https://doi.org/10.1016/j.ydbio.2015.04.021.
Hull RM, Cruz C, Jack CV, Houseley J. Environmental change drives accelerated adaptation through stimulated copy number variation. PLoS Biol. 2017:15:e2001333. https://doi.org/10.1371/journal.pbio.2001333.
Jiménez-Marín B, et al. Gene loss during a transition to multicellularity. Sci Rep. 2023:13:5268. https://doi.org/10.1038/s41598-023-29742-2.
Krug PJ, et al. 2022. Phylogenomic resolution of the root of panpulmonata, a hyperdiverse radiation of gastropods: new insight into the evolution of air breathing. Proc Biol Sci. 289:20211855. https://doi.org/10.1098/rspb.2021.1855.
Larade K, Storey KB. Accumulation and translation of ferritin heavy chain transcripts following anoxia exposure in a marine invertebrate. J Exp Biol. 2004:207:1353–1360. https://doi.org/10.1242/jeb.00872.
Laumer CE, Hejnol A, Giribet G. Nuclear genomic signals of the 'microturbellarian' roots of platyhelminth evolutionary innovation. eLife. 2015:4:e05503. https://doi.org/10.7554/eLife.05503.
Lesko SL, Rouhana L. Dynein assembly factor with WD repeat domains 1 (DAW1) is required for the function of motile cilia in the planarian *Schmidtea mediterranea*. Dev Growth Differ. 2020:62:423–437. https://doi.org/10.1111/dgd.12669.
Li F-W, et al. Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. Nat Plants. 2020:6:259–272. https://doi.org/10.1038/s41477-020-0618-2.
Little C. The colonisation of land: origins and adaptations of terrestrial animals. Cambridge University Press; 1983.
Liu C, et al. Giant African snail genomes provide insights into molluscan whole-genome duplication and aquatic–terrestrial transition. Mol Ecol Resour. 2021:21:478–494. https://doi.org/10.1111/1755-0998.13261.
Liu X, Sun Y, Wang S, Zhang S, Tian Q. Actin restricts cell proliferation and promotes differentiation during planarian regeneration. Biochem Biophys Res Commun. 2023:640:150–156. https://doi.org/10.1016/j.bbrc.2022.12.008.
Ma K-X, et al. Identification of a potential tissue-specific biomarker cathepsin L-like gene from the planarian *Dugesia japonica*: molecular cloning, characterization, and expression in response to heavy metal exposure Ecotoxicol Environ Saf. 2019:180:73–79. https://doi.org/10.1016/j.ecoenv.2019.04.079.
Martínez-Redondo GI, et al. MATEdb2: a collection of high-quality metazoan proteomes across the animal tree of life to speed up phylogenomic studies. Genome Biol Evol. 2024:16:evae235. https://doi.org/10.1093/gbe/evae235.
Martínez-Redondo GI, et al. FANTASIA leverages language models to decode the functional dark proteome across the animal tree of life. Commun Biol. 2025:8:1227. https://doi.org/10.1038/s42003-025-08651-2.
McGee C, Fairweather I, Blackshaw RP. Ultrastructural features of the epidermis of the planarian *Artioposthia triangulata* (Dendy). Hydrobiologia. 1997:347:15–24. https://doi.org/10.1023/A:1002934504296.
Merényi Z, et al. Genomes of fungi and relatives reveal delayed loss of ancestral gene families and evolution of key fungal traits. Nat Ecol Evol. 2023:7:1221–1231. https://doi.org/10.1038/s41559-023-02095-9.
Meyer A, et al. Giant lungfish genome elucidates the conquest of land by vertebrates. Nature. 2021:590:284–289. https://doi.org/10.1038/s41586-021-03198-8.
Morita M, Best JB. Electron microscopic studies of planarian regeneration . II. Changes in epidermis during regeneration. J Exp Zool. 1974:187:345–373. https://doi.org/10.1002/jez.1401870305.
Newmark PA, Sánchez Alvarado A. Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. Dev Biol. 2000:220:142–153. https://doi.org/10.1006/dbio.2000.9645.
Ong ST, Shan Ho JZ, Ho B, Ding JL. Iron-withholding strategy in innate immunity. Immunobiology. 2006:211:295–314. https://doi.org/10.1016/j.imbio.2006.02.004.
Pal A, Andersson DI. Bacteria can compensate the fitness costs of amplified resistance genes via a bypass mechanism. Nat Commun. 2024:15:2333. https://doi.org/10.1038/s41467-024-46571-7.
Parto S, Lartillot N. Molecular adaptation in rubisco: discriminating between convergent evolution and positive selection using mechanistic and classical codon models. PLoS One. 2018:13:e0192697. https://doi.org/10.1371/journal.pone.0192697.
Pascolini R, et al. Cytoskeletal protein expression in planarians. Boll Zool. 1993:60:403–406. https://doi.org/10.1080/11250009309355848.
Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017:14:417–419. https://doi.org/10.1038/nmeth.4197.
Pellettieri J, Sánchez Alvarado A. Cell turnover and adult tissue homeostasis: from humans to planarians. Annu Rev Genet. 2007:41:83–105. https://doi.org/10.1146/annurev.genet.41.110306.130244.
Plays M, Müller S, Rodriguez R. Chemistry and biology of ferritin. Metallomics. 2021:13:mfab021. https://doi.org/10.1093/mtomcs/mfab021.
Rahnenfuhrer AA. topGO: enrichment analysis for Gene Ontology. 2025. https://doi.org/10.18129/B9.bioc.topGO.
Reddien PW. The cellular and molecular basis of planarian regeneration. Cell. 2018:175:327–345. https://doi.org/10.1016/j.cell.2018.09.021.
Robinson D, et al. Gene-by-environment interactions influence the fitness cost of gene copy-number variation in yeast. G3 (Bethesda). 2023:3:13. https://doi.org/10.1093/g3journal/jkad159.
Robinson MD, McCarthy DJ, Smyth GK. Edger: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010:26:139–140. https://doi.org/10.1093/bioinformatics/btp616.
Sandve SR, Rohlfs RV, Hvidsten TR. Subfunctionalization versus neofunctionalization after whole-genome duplication. Nat Genet. 2018:50:908–909. https://doi.org/10.1038/s41588-018-0162-4.

Shagin DA, et al. Identification and characterization of a new family of C-type lectin-like genes from planaria girardia tigrina. Glycobiology. 2002:12:463–472. https://doi.org/10.1093/glycob/cwf056.

Sharma PP. Chelicerates and the conquest of land: a view of arachnid origins through an evo-devo spyglass. Integr Comp Biol. 2017:57:510–522. https://doi.org/10.1093/icb/icx078.

Silveira M, Corinna A. Fine structural observations on the protonephridium of the terrestrial triclad *Geoplana pasipha*. Cell Tissue Res. 1976:168:455–463. https://doi.org/10.1007/BF00215996.

Simões AES, et al. Efficient recovery of proteins from multiple source samples after trizol® or trizol®LS RNA extraction and long-term storage. BMC Genomics. 2013:14:181. https://doi.org/10.1186/1471-2164-14-181.

Sluys R. The evolutionary terrestrialization of planarian flatworms (platyhelminthes, tricladida, geoplanidae): a review and research programme. Zoosyst Evol. 2019:95:543–556. https://doi.org/10.3897/zse.95.38727.

Sluys R, Kawakatsu M, Riutort M, Baguñà J. A new higher classification of planarian flatworms (platyhelminthes, tricladida). J Nat Hist. 2009:43:1763–1777. https://doi.org/10.1080/00222930902741669.

Stocchino GA, et al. Amphibioplanidae: a new branch and family on the phylogenetic tree of the triclad flatworms (platyhelminthes: tricladida), represented by a species from sardinian caves with a remarkable lifestyle. Zool J Linn Soc. 2021:193:1364–1391. https://doi.org/10.1093/zoolinnean/zlaa183.

Sun J, et al. Signatures of divergence, invasiveness, and terrestrialization revealed by four apple snail genomes. Mol Biol Evol. 2019:36:1507–1520. https://doi.org/10.1093/molbev/msz084.

Train C-M, Pignatelli M, Altenhoff A, Dessimoz C. Iham and pyHam: visualizing and processing hierarchical orthologous groups. Bioinformatics. 2019:35:2504–2506. https://doi.org/10.1093/bioinformatics/bty994.

Van de Peer Y, Ashman T-L, Soltis PS, Soltis DE. Erratum to: polyploidy: an evolutionary and ecological force in stressful times. Plant Cell. 2021:33:2899. https://doi.org/10.1093/plcell/koab149.

Van Straalen NM. Evolutionary terrestrialization scenarios for soil invertebrates. Pedobiologia (Jena). 2021:87–88:150753. https://doi.org/10.1016/j.pedobi.2021.150753.

Vargas-Chavez C, et al. An episodic burst of massive genomic rearrangements and the origin of non-marine annelids. Nat Ecol Evol. 2024:9:1263–1279. https://doi.org/10.1038/s41559-025-02728-1.

Veitia RA, Birchler JA. Gene-dosage issues: a recurrent theme in whole genome duplication events. Trends Genet. 2022:38:1–3. https://doi.org/10.1016/j.tig.2021.06.006.

Veldsman WP, et al. Comparative genomics of the coconut crab and other decapod crustaceans: exploring the molecular basis of terrestrial adaptation. BMC Genomics. 2021:22:313. https://doi.org/10.1186/s12864-021-07636-9.

Vila-Farré M, et al. Evolutionary dynamics of whole-body regeneration in planarians flatworms. Nat Ecol Evol. 2023:7:2108–2124. https://doi.org/10.1038/s41559-023-02221-7.

Wen Z-Y, Kang Y-J, Ke L, Yang D-C, Gao G., Genome-wide identification of gene loss events suggests loss relics as a potential source of functional lncRNAs in humans. Mol Biol Evol. 2023:40:msad103. https://doi.org/10.1093/molbev/msad103.

Winsor L. Aspects of taxonomy and functional histology in terrestrial flatworms (tricladida: terricola). Pedobiologia (Jena). 1998:42:412–432. https://doi.org/10.1016/S0031-4056(24)00463-3.

Xie A, et al. Upf3a but not Upf1 mediates the genetic compensation response induced by leg1 deleterious mutations in an H3K4me3-independent manner. Cell Discov. 2023:9:63. https://doi.org/10.1038/s41421-023-00550-2.

Associate editor: Christopher Wheat