

# The monoaminergic system is a bilaterian innovation

Received: 15 August 2022

Accepted: 25 May 2023

Published online: 06 June 2023

 Check for updates

Matthew Goulty<sup>1</sup>, Gaele Botton-Amiot<sup>2</sup>, Ezio Rosato<sup>1</sup>,  
Simon G. Sprecher<sup>2</sup> & Roberto Feuda<sup>1</sup> ✉

Monoamines like serotonin, dopamine, and adrenaline/noradrenaline (epinephrine/norepinephrine) act as neuromodulators in the nervous system. They play a role in complex behaviours, cognitive functions such as learning and memory formation, as well as fundamental homeostatic processes such as sleep and feeding. However, the evolutionary origin of the genes required for monoaminergic modulation is uncertain. Using a phylogenomic approach, in this study, we show that most of the genes involved in monoamine production, modulation, and reception originated in the bilaterian stem group. This suggests that the monoaminergic system is a bilaterian novelty and that its evolution may have contributed to the Cambrian diversification.

Monoamines have different biological roles; in the nervous system, they are fundamental neuromodulators<sup>1–5</sup> required for neuronal plasticity<sup>6,7</sup>. They regulate behaviour and contribute to cognitive functions, including learning and memory formation, emotional states and homeostatic processes such as feeding and sleep<sup>1,3,4,8</sup>. In animals, monoamines are synthesised from aromatic amino acids—primarily tyrosine and tryptophan—using multiple enzymes and co-factors<sup>1</sup> (see Fig. 1A and Table 1 for definitions). Typically, the first reaction adds a hydroxyl group to the amino acid. Members of the aromatic amino acid hydroxylase (AAAH, Table 1 for definitions) family, phenylalanine hydroxylases (PAHs), tyrosine hydroxylases (THs), and tryptophan hydroxylases (TPHs), mediate this chemical reaction<sup>1,8–11</sup>. Then, enzymes of the amino acid decarboxylase family (AADC), which includes dopa decarboxylases (DDCs), histidine decarboxylases (HDCs), and tyrosine decarboxylases (TDCs), remove a carboxyl group<sup>1,8,12–14</sup>. Some monoamines, such as octopamine, epinephrine and norepinephrine, are modified further by the addition of a second hydroxyl group in a reaction catalysed by Cu(II) monooxygenases, such as dopamine beta-hydroxylases (DBHs), tyramine beta-hydroxylases (TBHs) and monooxygenase DBH-like (MOXD)s<sup>1,13,15,16</sup>. Other monoamines are methylated by phenylethanolamine-*N*-methyltransferases (PNMTs)<sup>1</sup>. Different combinations of these enzymes produce all the major monoamines described in animals (Fig. 1A).

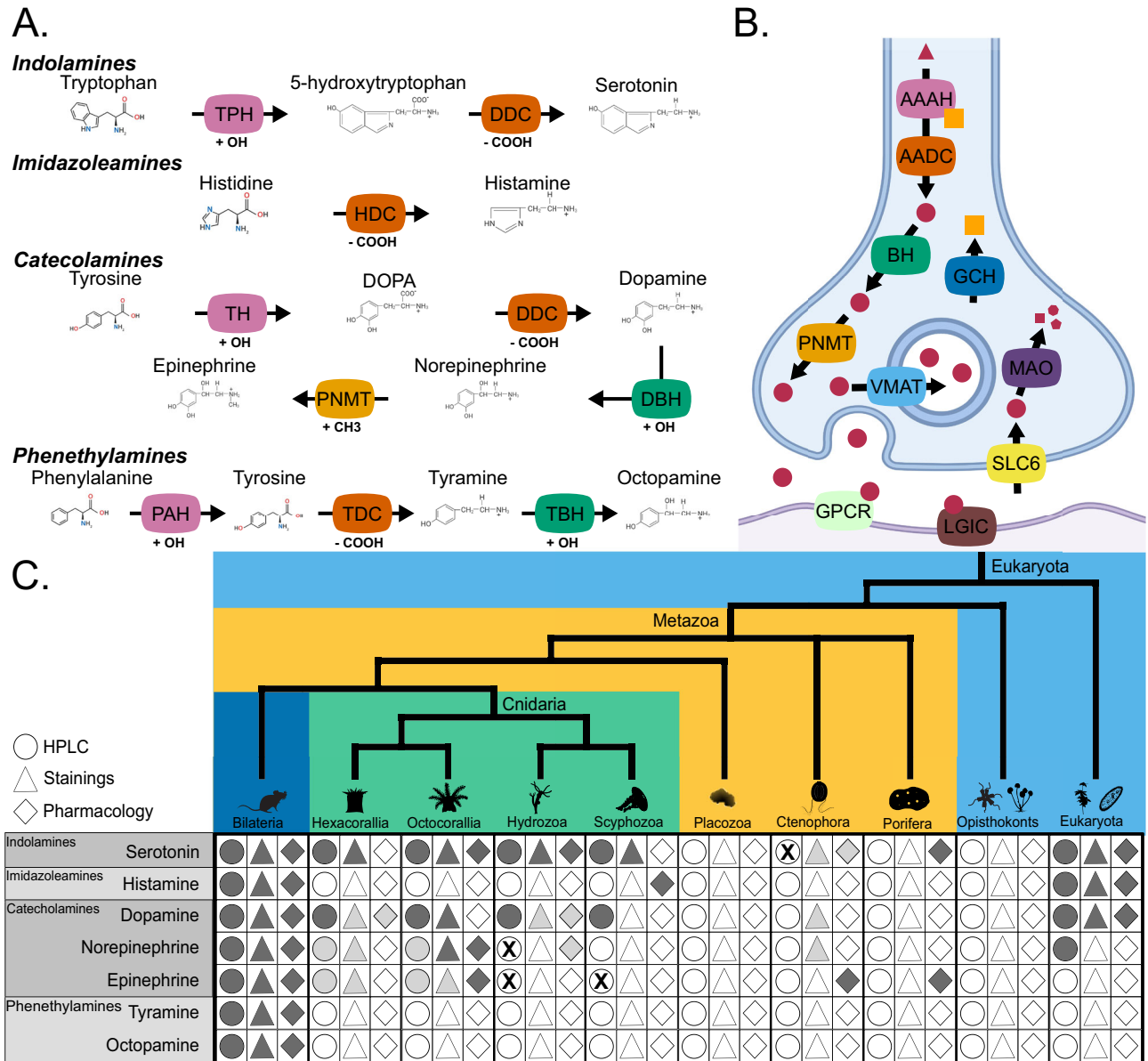
However, to function in the nervous system, the production of monoamines is necessary but not sufficient, and additional

elements are required. Vesicular monoamine transporters (VMATs) concentrate the monoamines in vesicles before secretion into the synaptic cleft (Fig. 1B)<sup>1,17–19</sup>. Several types of G-protein coupled receptors (GPCRs, e.g., dopaminergic receptors, serotonergic receptors, etc.) detect monoamines on either side of the synapse, triggering signalling cascades<sup>20–25</sup>. Finally, different proteins control the level of monoamines in the synaptic cleft by reuptake (such as transporters of the SLC6 family, like SERTs and DATs)<sup>26–30</sup> and by degradation (by catabolic enzymes, like monoamine oxidases, MAOs)<sup>31–36</sup>.

Different lines of evidence suggest that monoamines are present in some non-bilaterian animals (Fig. 1C). However, in these animals, the presence of the genes required for the production, modulation, and reception of monoamines remains unclear<sup>20,37–40</sup>. Several studies<sup>11,26,37,39,41–45</sup> have addressed this problem before, but they were unable to reach a definitive conclusion (summarised in Table 2). A main cause of uncertainty is the limited number of non-bilaterian genomes, such as those of sponges, ctenophores, placozoans, and cnidarians, that have been investigated so far<sup>11,26,37,39,41–45</sup>. Additionally, these studies have focussed on subsets of genes rather than on the monoaminergic system as a whole<sup>11,26,37,39,41–45</sup>. Furthermore, single-gene families have limited phylogenetic signal, making it difficult to resolve their relationships in ‘deep time’<sup>46,47</sup>.

Here, we investigate the coordinated evolution of the genes involved in the formation of the monoaminergic system in animals.

<sup>1</sup>Department of Genetics and Genome Biology, University of Leicester, Leicestershire, UK. <sup>2</sup>Department of Biology, Institute of Zoology, University of Fribourg, CH-1700 Fribourg, Switzerland. ✉e-mail: [rf190@le.ac.uk](mailto:rf190@le.ac.uk)



**Fig. 1 | An overview of the monoamine system.** **A** Synthesis pathways for key monoamines including the substrate, chemical modification, enzymes, and products. Arrows indicate reactions with the facilitating enzyme overlain. For each enzyme, the label shows the name while the colour indicates the gene family. Chemical modifications are shown next to the enzymes responsible. **B** Cartoon of a synapse with the different enzymes required for the production and detection of monoamines. The red circles represent monoamines and precursor molecules. Yellow squares represent tetrahydrobiopterin co-factor. **C** Current molecular evidence from the literature supporting the presence of monoamines outside Bilaterians in the literature. Dark grey indicates positive results, light grey displays uncertain or partial evidence (e.g., precursors, related compounds) and an X indicates negative results. Blank shapes indicate a lack of evidence. Staining refers to any chemical or immuno-staining experiments; pharmacology refers to evidence-based drug perturbations, adding inhibitors or other chemical interference experiments; HPLC High-Pressure Liquid Chromatography (see Supplementary Data 1 for references and details). PAH phenylalanine hydroxylase, TPH

tryptophan hydroxylase, TH tyrosine hydroxylase, DDC dopa decarboxylase, TDC tyrosine decarboxylase, HDC histidine decarboxylase, DBH dopamine beta hydroxylase, TBH tyramine beta hydroxylase, PNMT phenylethanolamine-N-methyltransferase, AAAH aromatic amino acid decarboxylase, AADC aromatic amine decarboxylase, BH beta hydroxylase, VMAT vesicular monoamine transporter, GCH GTP cyclo-hydrolase, GPCR g-protein coupled receptor, LGIC ligand gated ion channel, SLC solute ligand carrier, MAO monoamine oxidase. (A) and (B) were made with Biorender. Silhouettes obtained from Phylopic.org. Silhouette images are by Christoph Schomburg (Dendronephthya gigantea); Daniel Jaron (Mus musculus); Emily Jane McTavish, from <http://chestofbooks.com/animals/Manual-Of-Zoology/images/I-Order-Ciliata-41.jpg> (Ciliophora); Konsta Happonen, from a CC-BY-NC image by sokolkov 2002 on iNaturalist (Geranium sylvaticum); Mali'o Kodis, photograph by Ching (<http://www.flickr.com/photos/36302473@N03/>) (Chrysaora fuscescens); Noah Schlottman (Pleurobrachia); Oliver Voigt (Trichoplax adhaerens); Steven Traver (Hydra); and Tess Linden (Salpingoeca rosetta).

We study the pattern of duplication of 18 genes that encode the proteins involved in the synthesis, the turnover and the detection of monoamines. We take advantage of modern phylogenomic data covering a wide range of animals, especially non-bilaterians. Moreover, we used methods to minimise the effect of unstable sequences and reconstructed the pattern of gene duplication using

recently developed maximum likelihood reconciliation techniques. Our results strongly indicate that many orthologs of the genes involved in the monoaminergic system originated in the bilaterian stem-group. This suggests that the monoaminergic system is a bilaterian innovation and may have contributed to their evolutionary success.

**Table 1 | Definition of key abbreviations used in the paper**

Abbreviations	Definition
AAAH	Aromatic Amino Acid Hydroxylase
AADC	Aromatic Amine Decarboxylase
ACM	Acetylcholine Muscarinic Receptor
BUSCO	Benchmarking Universal Single Copy Orthologs
COMT	Catechol-O-Methyltransferase
DAT	Dopamine Transporter
DBH	Dopamine Beta Hydroxylase
DDC	Dopa Decarboxylase
GCH	GTP Cyclo-hydrolase
GPCR	G-protein Coupled Receptor
HDC	Histidine Decarboxylase
HNMT	Histamine-N-Methyltransferase
HRH	Histamine Receptor
IDAT	Invertebrate Dopamine Transporter
INE	Drosophila melanogaster transporter Ine
INMT	Indolethylamine-N-Methyltransferase
MAO	Monoamine Oxidase
MOXD	Monoxygenase DBH-like
NET	Nor-epinephrine Transporter
NNMT	Nicotinamide-N-Methyltransferase
PAH	Phenylalanine Hydroxylase
PNMT	Phenylethanolamine-N-Methyltransferase
SAM-MT	S-Adenosyl Methionine Methyltransferase
SERT	Serotonin Transporter
SLC	Solute Ligand Carrier
TBE	Transfer Bootstrap Expectation
TBH	Tyramine Beta Hydroxylase
TDC	Tyrosine Decarboxylase
TH	Tyrosine Hydroxylase
TPH	Tryptophan Hydroxylase
UFB	Ultrafast Bootstrap
VACHT	Vesicular Acetylcholine Transporter
VMAT	Vesicular Monoamine Transporter

## Results

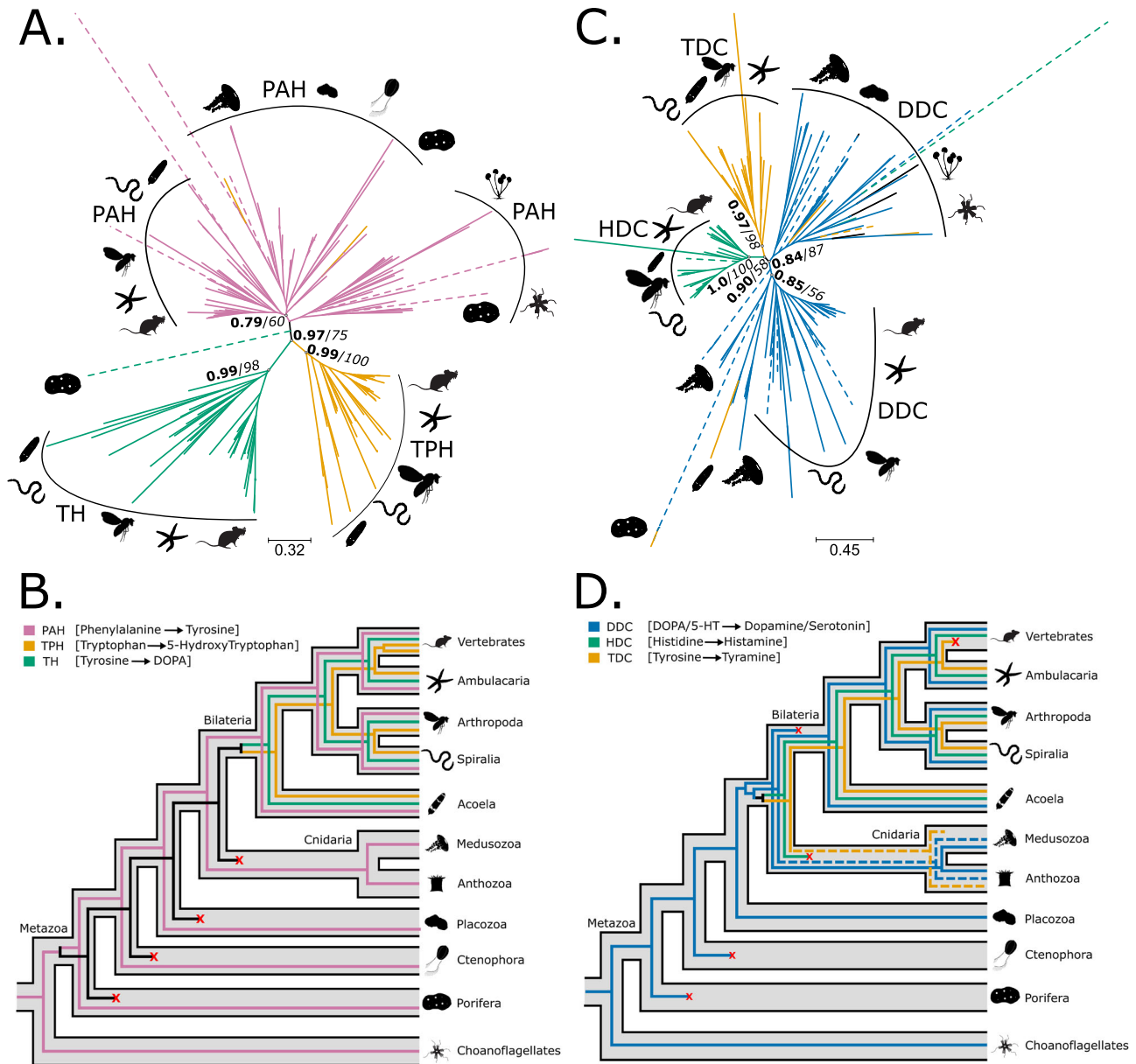
### The key enzymatic machinery for monoamine synthesis is a bilaterian novelty

Common problems with previous studies were the sparse taxonomical sampling and/or the use of limited phylogenetic methods (see Table 2<sup>11,26,37,39,41-45</sup>). To overcome these issues, we sampled 47 animal species, including 21 non-bilaterians and 18 opisthokonts (Supplementary Data 2 and Methods for further details). To minimise bias associated with poor-quality genomes, we selected species with a high BUSCO completeness score<sup>48</sup> (see Supplementary Data 2) while maintaining large taxonomical diversity. After homology and orthogroup identification (see Material and Methods and Supplementary Data 3 and 4), we computed gene trees using maximum likelihood (ML) inference. In addition to ultrafast bootstrap<sup>49</sup> (UFB), nodal support was estimated using the transferable bootstrap expectation score<sup>50</sup> (TBE). This method has been designed to identify and account for short and problematic sequences that have limited or conflicting phylogenetic signals<sup>50</sup>. Importantly, the presence of rogue/unstable sequences can affect phylogenetic relationships and the inferred duplication pattern. To identify unstable sequences, we used the Leaf Stability index<sup>51</sup> (LSI, as implemented in RogueNaRok<sup>52</sup>) and the t-index<sup>50</sup> (from the TBE analysis). Finally, we used GeneRax<sup>53</sup> to reconstruct the gene duplication events. In brief, given a gene and

**Table 2 | Summary of previous work and main conclusions**

Paper	Genes	Non-Bilaterians	Methods	Conclusions
Iyer et al. <sup>42</sup>	AAAH, AADC, BH, PNMT, COMT, HNMT, MAO	0	Similarity; NJ Tree; ML Tree; Bayesian Tree	DDC and HDC diverge in Metazoa TH sister to TPH and PAH
Caveney et al. <sup>26</sup>	SLC6	0	PCR; Similarity; NJ Tree	3 transporter Clades Predate Bilateria
Anciti <sup>41</sup>	AAAH, AADC, BH, PNMT, MAO, GPCRs, SLC6, VMAT	Nematostella vectensis ( <i>Renilla</i> sp.)	Similarity; Motif; NJ Tree; ML Tree	Homologues of most genes found in <i>N. vectensis</i>
Cao et al. <sup>11</sup>	AAAH	2 ( <i>N. vectensis</i> ; <i>Trichoplax adhaerens</i> )	Similarity; NJ Tree; ML Tree	TPH and PAH sisters TH and TPH Bilaterian
Sittberg-Liberles et al. <sup>43</sup>	AAAH	0	Bayesian Tree	TPH and TH are sisters Duplications in Metazoa
Kutcho and Sittberg-Liberles <sup>44</sup>	AADC, BH, MAO	0	Similarity; ML Tree	DDC Homologue to TDC and HDC; Metazoan unique TBH homologue of DBH and MOXD PNMT Vertebrate Specific
Krishnan and Schiöth <sup>45</sup>	GPCRs, AAAH, AADC, BH, PNMT	4 ( <i>N. vectensis</i> ; <i>T. adhaerens</i> ; <i>Mnemiopsis leidyi</i> ; <i>Amphimedon queenslandica</i> )	Similarity	Whole Pathway Present in Cnidaria DDC, DBH and GPCRs in other non-bilaterians
Francis et al. <sup>39</sup>	AADC, BH, MAO	2 Placozoans; 36 Cnidarians; 13 Ctenophores	Similarity; ML Tree	DBH Homologues in non-bilaterians DDC Ortholog in Hydrozoa DDC Homologues in non-bilaterians
Moroz et al. <sup>37</sup>	AAAH, AADC, SLC6, BH, PNMT, VMAT, 5HT3R	<i>T. adhaerens</i> ; 6 Cnidarians; <i>M. leidyi</i> ; <i>Pleurobrachia bachei</i> ; <i>A. queenslandica</i>	Uncertain	Only DDC and DBH have Orthologs in non-Bilateria

AAAH aromatic amino acid hydroxylase, AADC aromatic amine decarboxylase, BH beta hydroxylase, PNMT phenylethanolamine-N-methyltransferase, VMAT vesicular monoamine transporter, SLC solute ligand carrier, MAO monoamine oxidase, COMT catechol-O-methyltransferase, HNMT histamine-N-methyltransferase, GPCR g-protein coupled receptor, 5HT3R 5-hydroxytryptamine receptor, 3, PAH phenylalanine hydroxylase, TPH tryptophan hydroxylase, TH tyrosine hydroxylase, DDC dopa decarboxylase, TDC tyrosine decarboxylase, HDC histidine decarboxylase, DBH dopamine beta hydroxylase, MOXD monoxygenase DBH-like, NJ neighbour hood joining, ML maximum likelihood.



**Fig. 2 | Phylogeny and reconciliation for aromatic amino acid hydroxylases (AAAHs) and amino acid decarboxylases (AADCs).** **A** Transfer bootstrap expectation tree and **(B)** simplified illustration of reconciliation calculated using Generax for AAAH sequences. **C** Transfer bootstrap expectation tree and **(D)** simplified illustration of reconciliation calculated using Generax for AADC sequences. The nodal supports shown are transfer bootstrap expectation (TBE) scores (in bold), and ultrafast bootstrap proportion supports (in italic) for key nodes. Dashed lines indicate sequences identified as unstable in the t-index and leaf stability index (LSI) analysis (see Supplementary Data 6 for details). PAH phenylalanine hydroxylase, TPH tryptophan hydroxylase, TH tyrosine hydroxylase, DDC dopa decarboxylase,

TDC tyrosine decarboxylase, HDC histidine decarboxylase. Silhouettes obtained from Phylopic.org. Silhouette images are by Andrew R. Gehrke (*Hofstenia miamia*); Daniel Jaron (*Mus musculus*); Mali'o Kodis, photograph by Ching (<http://www.flickr.com/photos/36302473@N03/>) (*Chrysaora fuscescens*); Mario Quevedo (Asteriidae); Oliver Voigt (*Trichoplax adhaerens*); Ramiro Morales-Hojas (*Drosophila americana*); Steven Haddock, Jellywatch.org (*Hormiphora californensis*); and Tess Linden (*Salpingoeca rosetta*). Source Data are available at [[https://figshare.le.ac.uk/articles/dataset/Monoamine\\_Neuromodulation\\_is\\_a\\_Bilaterian\\_Innovation\\_Results/20391477](https://figshare.le.ac.uk/articles/dataset/Monoamine_Neuromodulation_is_a_Bilaterian_Innovation_Results/20391477)] in Results/GeneTrees labelled OG\_AAAH\_TBE.tbe.tree, OG\_AAAH\_UFB-treefile, OG\_AADC\_TBE.tbe.tree and OG\_AADC\_UFB.treefile.

species tree, GeneRax uses an ML approach to optimise the duplication and loss events (see<sup>53</sup> and<sup>54</sup> for details).

First, we investigated all the sequences encoding the aromatic amino acid hydroxylase (AAAH) family, enzymes that add a hydroxyl group to the aromatic ring of amino acid. This includes three key enzyme types: phenylalanine hydroxylases (PAHs) that synthesise tyrosine from phenylalanine; tyrosine hydroxylases (THs) that mediate the initial step in the synthesis of dopamine, epinephrine, and nor-epinephrine; and tryptophan hydroxylases (TPHs), which start the synthesis of serotonin (see Fig. 1A). The phylogenetic trees (Fig. 2A, Supplementary 1A–D) recovered the monophyly of TPHs (TBE = 0.99,

UFB = 100), THs (TBE = 0.99 and UFB = 98) and 'TPHs plus THs' (TBE = 0.97 and UFB = 75). Unlike PAHs, TPHs and THs were unique to Bilateria, apart from a single sequence from the sponge *Amphimedon queenslandica*. However, the results of the LSI and the t-index (see Methods and Supplementary Data 6) suggested that it most likely corresponds to a rogue taxa<sup>55,56</sup>. The absence of TPHs or THs in other sponges supported this conclusion, and we removed the *A. queenslandica* sequence from the dataset. The gene tree to species tree reconciliation corroborated the phylogenetic analyses indicating that THs and TPHs originated in the bilaterian stem group (Fig. 2B, Supplementary Fig. 1F–I).

Amino acid hydroxylases require the co-factor tetrahydrobiopterin to function<sup>1</sup>, which is synthesised by GTP cyclohydrolases (GCHs), among other enzymes<sup>1</sup>. Unlike AAAH, our phylogenetic tree for GCHs (Supplementary Fig. 3A, B) indicated that these enzymes are found across all animal species and opisthokonts. Phylogeny and reconciliation analyses showed a clear resemblance between gene tree and species tree, suggesting that no major duplication or loss occurred in this gene family (Supplementary Figs. 2, 3D, E). GCHs are not specific to monoaminergic pathways<sup>1</sup>, which may explain their wider distribution.

Next, we investigated the evolution of aromatic amino acid decarboxylase (AADC) enzymes, which remove a carboxyl group as part of monoamine synthesis<sup>1</sup> (Fig. 1A). Enzymes of the AADC family include dopa decarboxylases (DDCs), histidine decarboxylases (HDCs) and tyrosine decarboxylases (TDCs). The phylogenetic trees (Fig. 2C, Supplementary Fig. 4A–D) recovered the monophyly of HDCs and TDCs (TBE = 0.9 and UFB = 58). Both groups included only bilaterian sequences, except putative TDCs from the cnidarian *Paramuricea biscaya*. In contrast, we identified two clades of DDC encoding genes. The first comprised mainly bilaterians (TBE = 0.86, UFB = 69), while the second included most non-bilaterian and non-metazoan sequences (TBE = 0.84, UFB = 87). In the first DDC clade, the sequences from the sponge *Sycon ciliatum* and the cnidarians *Hydra magnipapillata* and *Clytia hemisphaerica* (TBE = 0.57) were positioned as a sister group to the bilaterian DDCs (TBE = 0.85). LSI and the t-index (see Supplementary Data 6) identified the sequences from *S. ciliatum* as unstable (supported by the alternate positioning in the UFB tree, Supplementary Fig. 4B) but not the sequences from *H. magnipapillata* and *C. hemisphaerica* (see Supplementary Data 6). This suggests that the latter sequences likely represent putative DDC orthologs, consistent with previous observations<sup>39</sup>. The reconciliation analysis (Fig. 2D, Supplementary Fig. 4F–I) performed after excluding the unstable sequences (see above) suggested that HDCs and TDCs originated in the stem group of Cnidaria and Bilateria. However, such a scenario relies merely on sequences from a single species (the coral *P. biscaya*). For DDCs, the reconciliation suggested a similar pattern: an origin in the stem group of Cnidaria and Bilateria. However, such a hypothesis would require the loss of DDC orthologs in all cnidarians except hydrozoans (Fig. 2D, Supplementary Fig. 4F–I). In summary, our results implied that DDCs, HDCs and TDCs are almost exclusive to Bilateria, with the reconciliation analysis placing their origin in the stem lineage of Cnidaria and Bilateria. However, this ‘earlier’ (compared to an origin in the stem lineage of Bilateria) scenario depends upon a few sparsely distributed cnidarian sequences.

Dopamine beta hydroxylases (DBHs), tyramine beta hydroxylases (TBHs), and monooxygenase dopamine beta hydroxylase-like (MOXD) are involved in the synthesis of norepinephrine and octopamine<sup>1</sup> (Fig. 1A). In vertebrates, DBHs are used to synthesise norepinephrine<sup>1</sup>, while in arthropods TBHs are involved in the production of octopamine<sup>1</sup>. MOXD is not functionally characterised. The phylogenetic trees (Fig. 3A, Supplementary Fig. 5A, B) supported the monophyly of MOXD (TBE = 0.95, UFB = 87) and of DBHs and TBHs (TBE = 1.0, UFB = 100), which were almost exclusively limited to bilaterians except for a single sequence from *P. biscaya* in both clades. While the TBE tree placed the non-bilaterian sequences as the sister group of the DBHs/TBHs clade (TBE = 0.64), the UFB tree assigned them as the sister group of the bilaterian MOXD (UFB = 48) (Supplementary Fig. 5A, B, respectively). To consider these differences, we performed the reconciliation analysis on both trees (Supplementary Fig. 5D–G). When reconciled, both analyses supported the classification of cnidarian sequences as orthologs of DBHs/TBHs (as in the TBE tree, Supplementary Fig. 5A). Furthermore, most reconciliations supported the split between MOXD and DBHs/TBHs occurring in the ancestor of Cnidaria and Bilateria (Fig. 3B and Supplementary Fig. 5D–F). Interestingly and contrary to previous observations<sup>44</sup>, both

trees (Supplementary Fig. 5A, B) indicated that the fruit fly TBH is orthologous (1:1) to human DBH, suggesting that the name distinction is purely semantic (see Supplementary Fig. 5A, B).

Phenylethanolamine-*N*-methyltransferases (PNMTs) catalyse the last step in the production of epinephrine from norepinephrine and belong to the SAM-binding methyltransferase superfamily (SAM-MT)<sup>57</sup>. Our analysis indicated that the enzymes of the SAM-MT family are widely distributed across Metazoa, but PNMTs are uniquely present in vertebrates (TBE = 1, UFB = 100 Fig. 3C, Supplementary 6A, B). Additionally, we showed that these sequences form a monophyletic group with nicotinamide-*N*-methyltransferases (NNMTs) and indolethylamine-*N*-methyltransferases (INMTs), both also unique to vertebrates (Fig. 3C). The reconciliation (Fig. 3D, Supplementary 6D, E) suggested that PNMTs and INMTs/NNMTs are the product of a gene duplication in the ancestor of jawed vertebrates (gnathostomes). A subsequent duplication in mammals gave rise to additional INMT and NNMT paralogs.

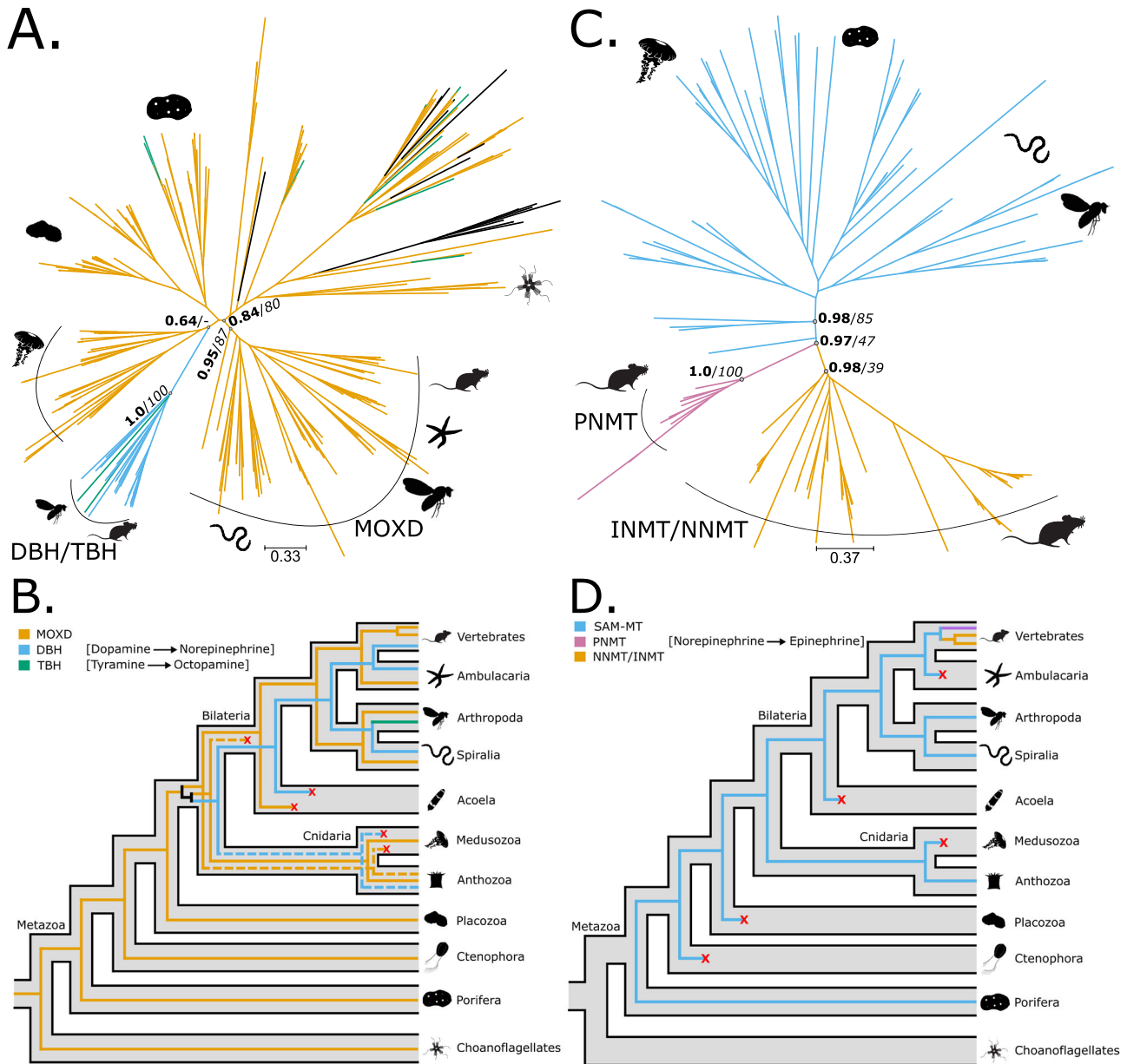
### Monoamine transporters orthologs are unique to Bilateria

Transporters concentrate monoamines inside vesicles for secretion into the synaptic cleft<sup>1</sup>. In Bilateria, vesicular monoamine transporters (VMATs) perform this role. These are part of the solute ligand carrier family 18 (SLC18) and vesicular acetylcholine transporters (VACHTs)<sup>17</sup>. The phylogeny of the SLC18 orthogroup (Fig. 4A, B) supported the monophyly of bilaterian VMATs (TBE = 0.86, UFB = 44) and VACHTs (TBE = 1.0, UFB = 100) (Fig. 4A, Supplementary Fig. 7A, B). Three cnidarian sequences were placed within the VMATs clade; however, LSI and the t-index identified them as unstable (see Supplementary Data 6). Furthermore, our phylogenetic results supported the monophyly of VMATs and VACHTs (TBE = 0.99, UFB = 100, Fig. 4A, Supplementary Fig. 7A, B). Interestingly, the analysis placed some sequences from sponges, choanoflagellates and other holozoans as outgroups of the bilaterian VMAT/VACHT clade (Fig. 4A and Supplementary Fig. 7A, B). After removing the unstable cnidarian sequences, the reconciliation confirmed these findings, namely that VMATs and VACHTs descend from a duplication in the bilaterian stem group (Fig. 4B, Supplementary Fig. 7G, I) and that all non-bilaterians except sponges lack VMAT/VACHT orthologs.

SLC6 is a large family of transporters, including those for serotonin (SERTs/SLC6a4), dopamine in invertebrates (iDATs) and dopamine/epinephrine/norepinephrine in vertebrates (SLC6a2/3s). These proteins regulate the concentration of monoamines in the synaptic cleft<sup>1,26</sup>. All monoamine transporters formed a clade almost exclusive to bilaterians (TBE = 0.99, UFB = 100), apart from a single sequence from the cnidarian *P. biscaya* (Fig. 4C, Supplementary Fig. 8A, B). Our analyses supported the monophyly of SLC6a4 (TBE = 0.94, UFB = 100), iDATs (TBE = 0.98, UFB = 100), and SLC6a2/3s (TBE = 0.96, UFB = 47). Furthermore, we identified a sister group relationship between the transporters for monoamines and a cluster that included transporters for GABA, taurine and creatine (TBE = 0.79, UFB = 50). Such grouping mainly consisted of bilaterian sequences (i.e., SLC6a1s, SLC6a6s and SLC6a8s in Fig. 4C). Both categories of transporters were recovered as a sister group to a large cnidarian-specific clade (TBE = 0.83, UFB = 61). Thus, reconciliation analyses indicated that the transporters for monoamine originated in the stem of Bilateria (Fig. 4D, Supplementary Fig. 8D, E). In addition, based on a single sequence from *P. biscaya* we identified an ortholog to all monoamine transporters in corals. However, evidence from more species is needed to corroborate this finding.

### Evolutionary history of monoamine catabolic enzymes

The correct functioning of neural circuits requires tight control of monoamine levels. Several catabolic enzymes are used in animals for breaking down/inactivating monoamines, which provide a key mechanism of regulation<sup>58</sup>. In bilaterians, this is mediated by



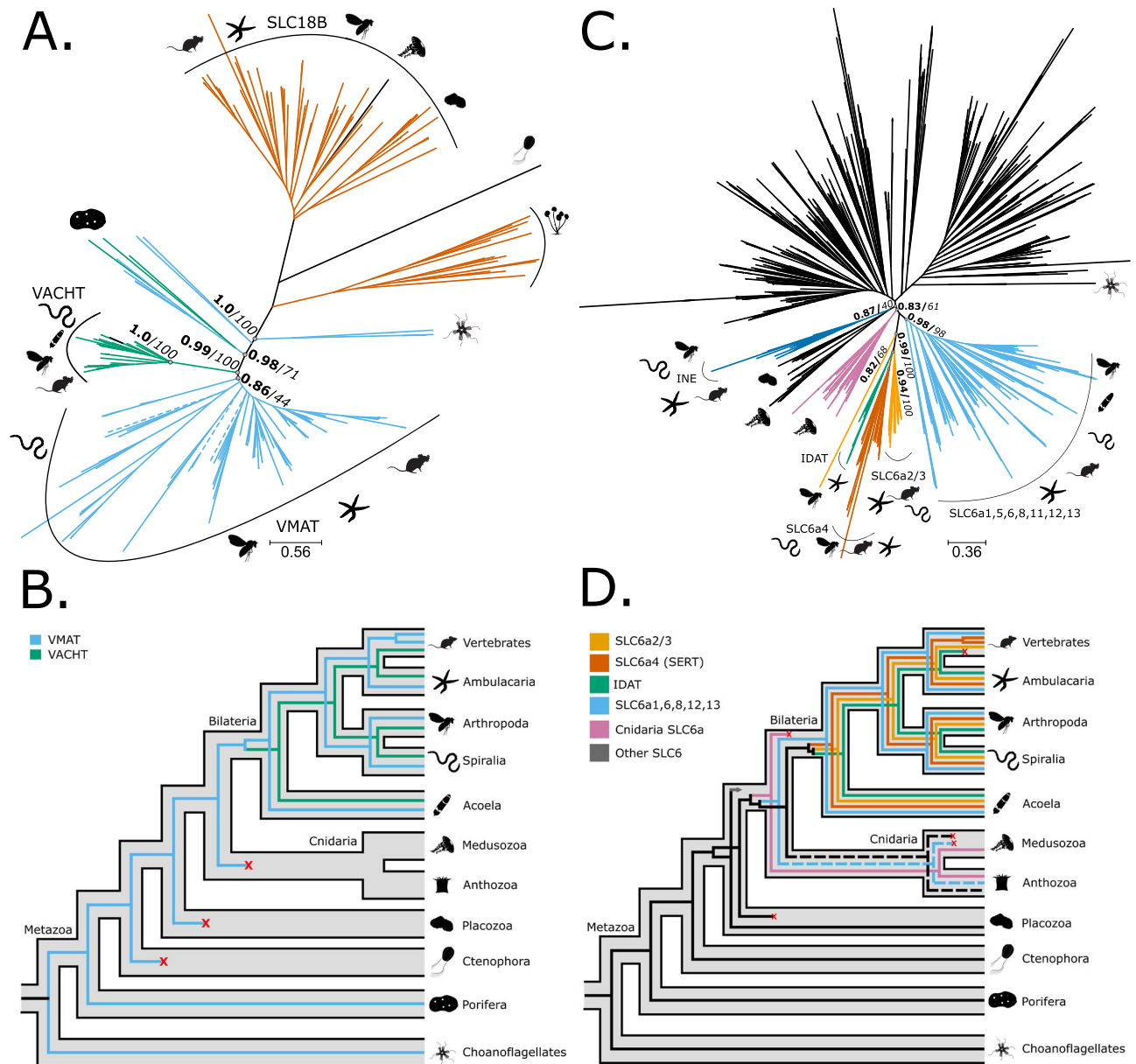
**Fig. 3 | Phylogeny and reconciliation for beta-hydroxylases (BHs) and phenylethanolamine-*N*-methyltransferases (PNMTs).** **A** Transfer bootstrap expectation tree and **(B)** simplified illustration of reconciliation calculated using Generax for BH sequences. **C** Transfer bootstrap expectation tree and **(D)** simplified illustration of reconciliation calculated using Generax for PNMT sequences. The nodal supports shown are transfer bootstrap expectation (TBE) scores (in bold), and ultrafast bootstrap proportion supports (in italic) for key nodes. Dashed lines indicate sequences identified as unstable in the t-index and leaf stability index (LSI) analysis (see Supplementary Data 6 for details). DBH dopamine beta hydroxylase, TBH tyramine beta hydroxylase, MOXD monooxygenase DBH-like, INMT indolethylamine-*N*-methyltransferase, NNMT nicotinamide-*N*-methyltransferase, SAM-MT

S-adenosyl methionine methyltransferase. Silhouettes obtained from Phylopic.org. Silhouette images are by Andrew R. Gehrke (*Hofstenia miamia*); Daniel Jaron (*Mus musculus*); Mali'o Kodis, photograph by Ching (<http://www.flickr.com/photos/36302473@N03/>) (*Chrysaora fuscescens*); Mario Quevedo (Asteriidae); Oliver Voigt (*Trichoplax adhaerens*); Ramiro Morales-Hojas (*Drosophila americana*); Steven Haddock, Jellywatch.org (*Hormiphora californensis*); and Tess Linden (Salpingoeca rosetta). Source Data are available at [[https://figshare.le.ac.uk/articles/dataset/Monoamine\\_Neuromodulation\\_is\\_a\\_Bilaterian\\_Innovation\\_Results/20391477](https://figshare.le.ac.uk/articles/dataset/Monoamine_Neuromodulation_is_a_Bilaterian_Innovation_Results/20391477)] in Results/GeneTrees labelled OG\_BH\_TBE.tbe.tree, OG\_BH\_UFB.treefile, OG\_PNMT\_TBE.tbe.tree and OG\_PNMT\_UFB.treefile.

catechol-*o*-methyltransferases (COMTs), histamine-*N*-methyltransferases (HNMTs) and monoamine oxidases (MAOs) (see Fig. 1B).

COMTs add a methyl group and consequently inactivate dopamine, epinephrine, and norepinephrine<sup>32,59</sup>. Phylogenetic trees and reconciliation analyses (Supplementary Figs. 9, 10) identified COMTs as primarily present in deuterostomes. The phylogenetic analyses placed echinoderm COMTs as a sister group to the vertebrate transmembrane-*O*-methyltransferase clade (TBE = 0.65 and UFB = 60), which also may be able to inactivate dopamine,

epinephrine, and norepinephrine<sup>60</sup>. Additionally, this orthogroup included sequences from two corals (*P. biscaya* and *Heliopora coerulea*), two sponges (*S. ciliatum* and *Oscarella pearsei*), a rotifer (*Adineta vaga*) and several non-metazoans (Supplementary Data 3). The reconciliation study inferred many secondary losses across Metazoa (Supplementary Figs. 9, 10D, E). Thus, the phylogenetic distribution of COMTs seem to contrast with the other monoaminergic components that are usually highly conserved across Bilateria (see Supplementary Data 3).



**Fig. 4 | Phylogeny and reconciliation for vesicular monoamine transporters (VMATs) and Solute ligand carrier 6 (SLC6) family members.** **A** Transfer bootstrap expectation tree and **(B)** simplified illustration of reconciliation calculated using Generax for VMAT sequences. **C** Transfer bootstrap expectation tree and **(D)** simplified illustration of reconciliation calculated using Generax for SLC6 sequences. The nodal supports shown are transfer bootstrap expectation (TBE) scores (in bold), and ultrafast bootstrap proportion supports (in italic) for key nodes. Dashed lines indicate sequences identified as unstable in the t-index and leaf stability index (LSI) analysis (see Supplementary Data 6 for details). VACHT vesicular acetylcholine transporter, INE *Drosophila melanogaster* transporter Ine, IDAT invertebrate dopamine transporter, SERT serotonin transporter, SLC solute ligand

carrier. Silhouettes obtained from Phylopic.org. Silhouette images are by Andrew R. Gehrke (*Hofstenia miamia*); Daniel Jaron (*Mus musculus*); Mali'o Kodis, photograph by Ching (<http://www.flickr.com/photos/36302473@N03/>) (*Chrysaora fuscescens*); Mario Quevedo (*Asteriidae*); Oliver Voigt (*Trichoplax adhaerens*); Ramiro Morales-Hojas (*Drosophila americana*); Steven Haddock, Jellywatch.org (*Hormiphora californensis*); and Tess Linden (*Salpingoeca rosetta*). Source Data are available at [[https://figshare.le.ac.uk/articles/dataset/Monoamine\\_Neuromodulation\\_is\\_a\\_Bilaterian\\_Innovation\\_Results/20391477](https://figshare.le.ac.uk/articles/dataset/Monoamine_Neuromodulation_is_a_Bilaterian_Innovation_Results/20391477)] in Results/GeneTrees labelled OG\_VMAT\_TBE.tbe.tree, OG\_VMAT\_UFB.treefile, OG\_SLC6\_TBE.tbe.tree and OG\_SLC6\_UFB.treefile.

The enzymes of the Histamine-*N*-methyltransferase (HNMTs) family inactivate histamine by adding a methyl group<sup>33,61</sup>. We identified HNMTs in Deuterostomia, Acoela, and Anthozoa (Cnidaria), with sequences present in almost every species from these groups (Supplementary Data 3). Phylogenetic trees and reconciliation analyses (Supplementary Figs. 11, 12A, B) suggested that HNMTs originated in the ancestral group of Cnidaria/Bilateria followed by many secondary losses (Supplementary Figs. 11, 12D, E).

Monoamine oxidases (MAOs) inactivate all monoamines and other amino acid derivatives such as tryptamine, benzylamine and

kynuramine<sup>31</sup>. Both the TBE and the UFB trees (Supplementary Figs. 14A, B) supported the monophyly of bilaterian MAOs, including both vertebrate paralogs (TBE = 0.92, UFB = 43). In parallel, we identified a monophyletic clade of cnidarian sequences (TBE = 0.93, UFB = 94) as sister to the bilaterian MAOs. Several choanoflagellate and other non-metazoans contributed to a larger MAO clade that included the metazoan sequences (TBE = 0.99, UFB = 86) (Supplementary Fig. 14A, B). Furthermore, we recognised a MAO-like group that included animal and choanoflagellate sequences related to the large MAO clade (TBE = 0.97, UFB = 91). Interestingly, Medusozoa,

Porifera and *D. melanogaster* retain MAO-like sequences only (Supplementary Fig. 13, 14A, B, D, E). We did not identify any MAO in ctenophores or placozoans, while at least one homologue from each clade (MAO and MAO-like) is present in all other animal groups (Supplementary Fig. 13, 14A, B, D, E). The reconciliation study suggested that the two major MAO clades have an ancient origin, significantly predating the origin of animals and neurons (Supplementary Fig. 13, 14D, E).

### Most monoaminergic receptors are specific to bilaterians

Once released in the synaptic cleft, monoamines are detected by specific transmembrane proteins expressed on both sides of the synapse. These receptors are the effectors that trigger signalling cascades inducing a physiological response. Most monoamine receptors are G-protein-coupled receptors (GPCRs) of class A<sup>20,45</sup>. However, serotonin can also be detected by ligand-gated ion channel receptors from the cys-loop repeated gene family<sup>1</sup> (5HT3Rs).

Our studies revealed that all putative 5HT3R sequences form an orthogroup with the ligand-gated zinc-activated ion channels. This group comprised vertebrates and some urochordates and acodels sequences (see Supplementary Data 3). The phylogenetic tree indicated that most vertebrate 5HT3Rs are monophyletic (TBE = 0.96, UFB = 97; Supplementary Fig. 16A, B). Some fish 5HT3R-annotated sequences clustered with zinc-activated ion channels. Accordingly, the reconciliation study suggested that while several losses occurred in bilaterians, 5HT3Rs underwent an expansion in vertebrate lineages (Supplementary Fig. 15, and 16D, E).

Then, we turned our attention to GPCRs. Our analysis divided the monoaminergic GPCRs into two orthogroups (see Supplementary Data 3). The first, hereafter OG30, contained histamine receptors 1, 3 and 4 (HRH1, HRH3, HRH4) and acetylcholine muscarinic receptors (ACMs). It included bilaterian and non-bilaterian sequences. The second, hereafter OG1, comprised the remaining known monoaminergic GPCRs, including the receptors for serotonin, dopamine, epinephrine/norepinephrine, octopamine/tyramine, trace amines and histamine receptor 2 (HRH2). OG1 included sequences from bilaterians and non-bilaterians (Supplementary Data 3). However, given the taxonomic composition and the lack of outgroups, it was impossible to understand the duplication pattern and the following orthology relationships.

To overcome this limitation, we expanded our study to include all GPCR orthogroups. To reduce the computational burden, we considered non-redundant, seven transmembrane domains GPCRs only (see methods). Thus, we selected 2,837 GPCRs and used CLANs<sup>62</sup> (a similarity-based method) to identify clusters of related sequences. In brief, CLANs use an all-*vs.*-all BLAST to compute similarity and to cluster sequences at different similarity thresholds<sup>62</sup> (i.e. different P-values). Applying a threshold including all connections with P-value < 1e<sup>-60</sup>, we observed that most bilaterian sequences from OG1 and OG30 clustered together (Supplementary Fig. 19A). To identify potential outgroups, we relaxed the threshold to include all connections with P-value < 1e<sup>-40</sup> (Supplementary Fig. 19C). At this threshold, we observed that OG1, OG30, adenosine receptors, and melatonin receptors formed a cluster with many interconnected sequences and that such a cluster had sporadic connections with opsins and other GPCRs. To better focus on monoaminergic GPCRs, we increased the P-value threshold to 1e<sup>-42</sup> (Supplementary Fig. 19B) and identified 1,277 sequences connected to OG1 and OG30.

While CLANs is a powerful tool to investigate the relationships between a large number of sequences, it cannot clarify patterns of gene duplication and specific evolutionary histories. Thus, we performed phylogenetic studies using opsin sequences as the outgroup. Both the UFB and the TBE trees (Fig. 5A, Supplementary Fig. 20 A, B) provided support for the monophyly of OG30 (TBE = 0.97 and UFB = 74) and of OG1+OG30 (TBE = 0.96 and UFB = 25). Mainly, these clades are composed by bilaterian sequences, except for a few from

cnidarians that are a sister group to the ACM receptors within OG30. Additionally, both trees supported the existence of a large cnidarian-specific monophyletic group (TBE = 0.65, UFB = 66). Additionally, we identified a clade of adenosine receptors (TBE = 0.73, UFB = 63), which included sequences from bilaterians, cnidarians and placozoans. The adenosine receptors clade was placed as sister to OG1 and OG30 and the cnidarian-specific clade. The main disagreement between the UFB and the TBE trees was the position of OG1. The UFB tree recovered OG1 as monophyletic, albeit with low support (UFB = 37). The TBE tree divided OG1 in two clades with OG30 nested within (TBE = 0.70) (Fig. 5A, Supplementary Fig. 20 A, B). Another difference between the trees was the phylogenetic position of two placozoan sequences. These were placed within OG1 by UFB but as a sister group to OG1 and OG30 by TBE. However, the LSI analysis identified these sequences as unstable (Supplementary Data 6).

The reconciliation performed using the TBE and UFB trees supported that OG1, OG30 and the cnidarian-specific clade originated from a duplication that gave rise to the adenosine receptors and to a “proto-monoamine” receptor in the ancestor of placozoans, cnidarians and bilaterian. The “proto-monoamine” receptor subsequently gave rise, in the cnidarian/bilaterian stem group, to the canonical monoamine receptors (OG30 and OG1) and to the cnidarian-specific clade (Fig. 5B, Supplementary Fig. 20C, D–F). Both trees supported OG1 and OG30, originating from a single gene in the cnidarian/bilaterian stem group. Subsequently, OG1 underwent a significant expansion in the bilaterian stem group giving rise to the modern bilaterian monoaminergic receptors. However, the two trees disagree on the precise number of copies of OG1 in the common ancestor of bilaterians, with UFB indicating a single copy (Supplementary Fig. 20D, F) and TBE suggesting multiple copies (Supplementary Fig. 20C, E).

In summary, our results suggest that monoaminergic GPCRs originated from a gene duplication event that also gave rise to the adenosine receptors in the eumetazoan stem-group. Modern monoaminergic receptors can be divided into two subclasses. The first is bilaterian specific and contains receptors for serotonin, dopamine, epinephrine/norepinephrine, octopamine/tyramine, trace amines and HRH2. The second includes HRH1, HRH3, HRH4 and ACM receptors comprising both bilaterian and cnidarian sequences. Our reconciliation analyses indicate that the two subclasses diverged in the Bilateria/Cnidaria stem group.

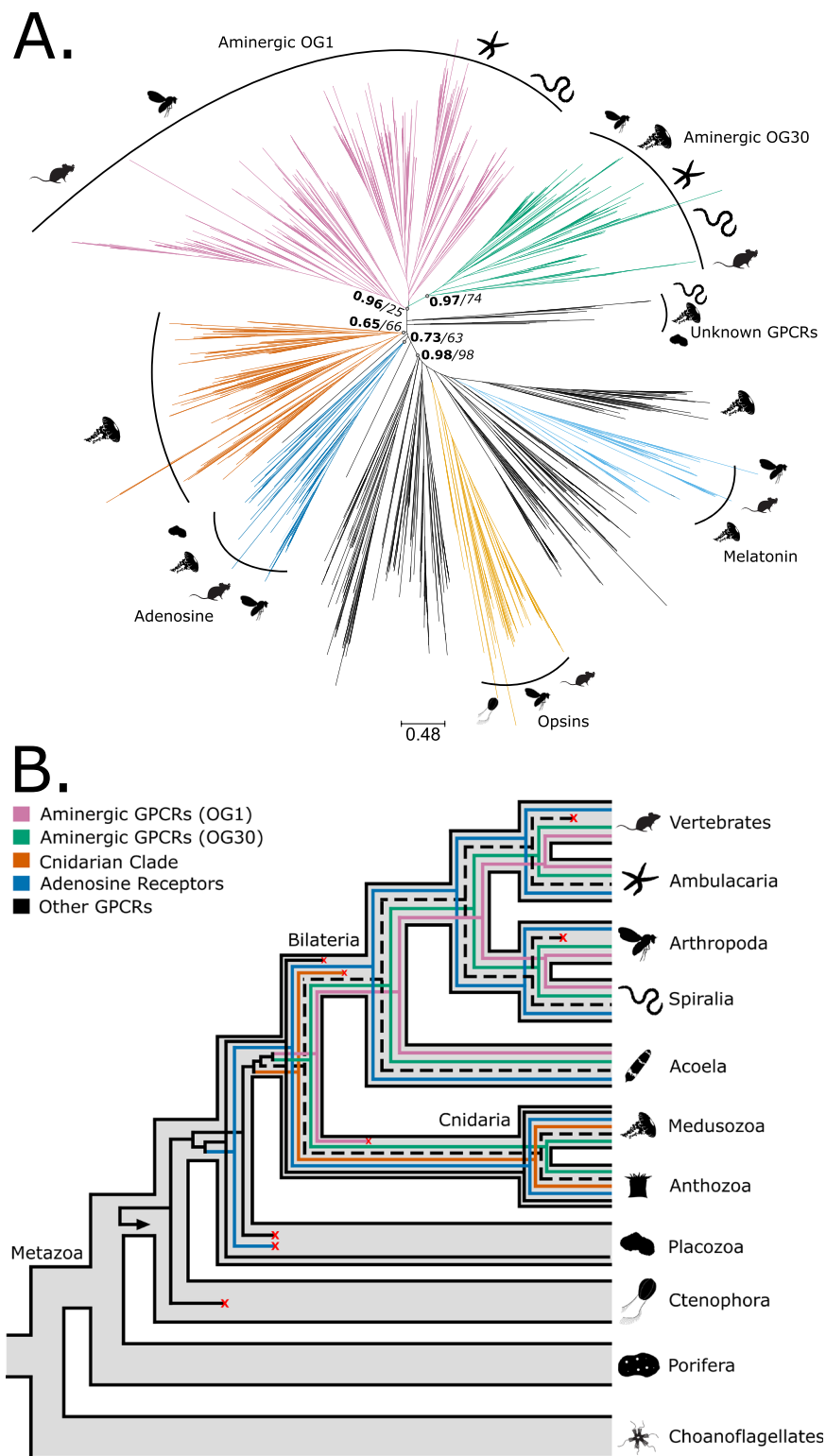
## Discussion

In this study, we have provided strong phylogenetic evidence that the genes that are essential for building a monoaminergic system (i.e., encoding the proteins required for the production, detection, and degradation of monoamines in the nervous system) appear as a whole in bilaterians. Our results (summarised in Fig. 6) allow for a substantial clarification of the evolution of the monoaminergic system.

Compared to previous work<sup>11,26,37,39,41–45</sup> we have included a larger taxonomic sample, which has provided greater resolution in clarifying the evolution of the different components of the monoaminergic system. We have used statistical methods to pinpoint unstable taxa and to reconcile the gene trees with the species tree, which has improved the robustness of our phylogenetic inferences. We have confirmed previous observations, for instance, that THs and TPHs are present only in Bilateria<sup>11,37,43</sup>. However, in contrast to previous work, we show that most monoaminergic genes evolved through gene duplication in the Bilateria stem group and that others, such as 5HT3Rs and PNMTs, originated in vertebrates. The distribution of orthologs summarised in Fig. 6, suggests a scenario where the monoaminergic system was assembled by combining ancient enzymes with broad functions (such as GCH and MAO) with newly evolved bilaterian-specific genes (such as TPH and TH) into a new functional unit.

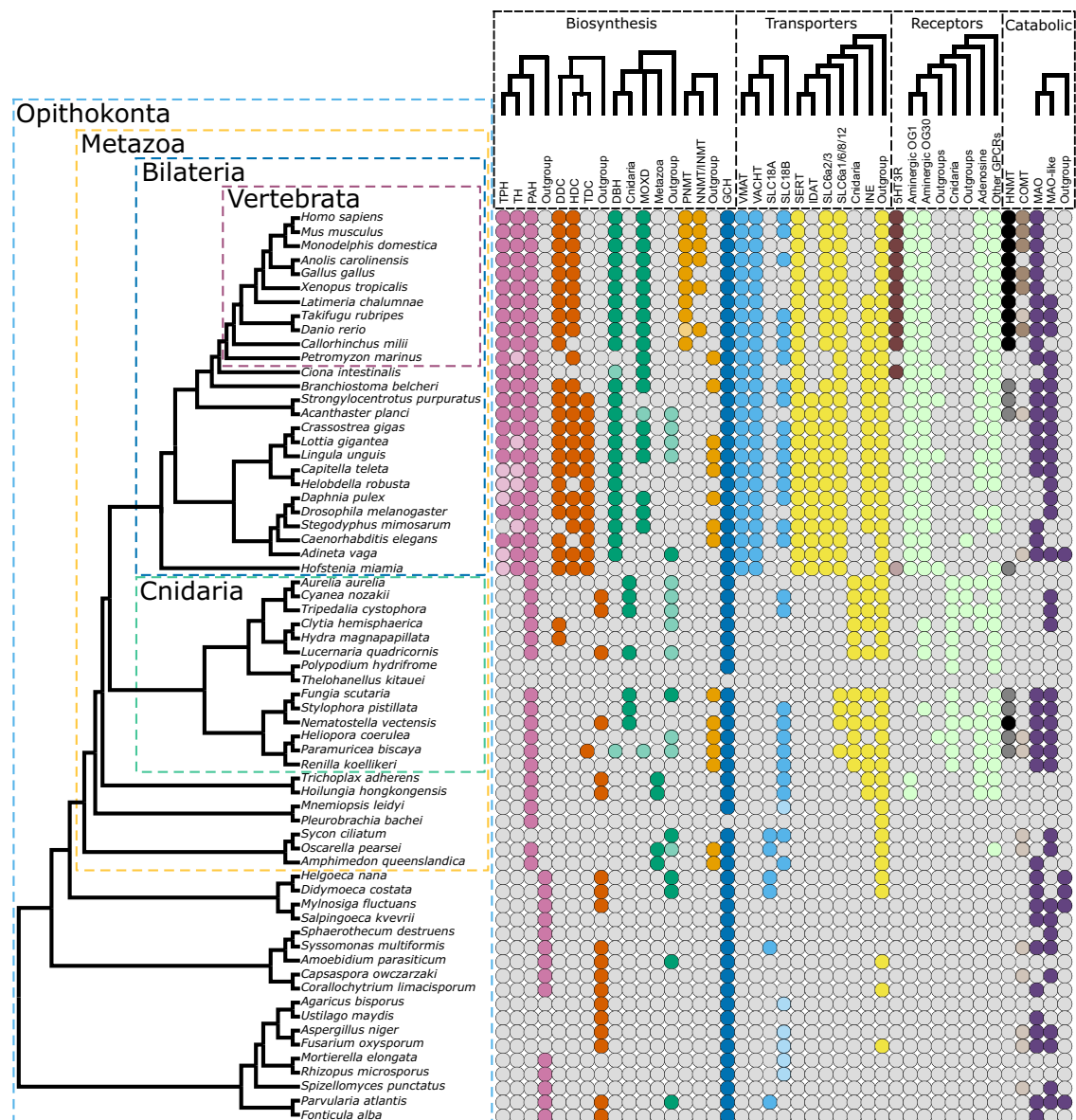
The conclusion that the monoaminergic system is a bilaterian innovation may seem in conflict with existing experimental





**Fig. 5 | Phylogeny and reconciliation of the monoaminergic GPCRs.** **A** Transfer bootstrap expectation tree and **B** simplified illustration of reconciliation calculated using Generax for monoaminergic g-protein coupled receptors (GPCRs). The nodal supports shown are transfer bootstrap expectation (TBE) scores (in bold), and ultrafast bootstrap proportion supports (in italic) for key nodes. Dashed lines indicate sequences identified as unstable in the t-index and leaf stability index (LSI) analysis (see Supplementary Data 6 for details). OG orthogroup. Silhouettes obtained from Phylopic.org. Silhouette images are by Andrew R. Gehrke (*Hofstena*

*miamia*); Daniel Jaron (*Mus musculus*); Mali'o Kodis, photograph by Ching (<http://www.flickr.com/photos/36302473@N03/>) (*Chrysaora fuscescens*); Mario Quevedo (Asteroidea); Oliver Voigt (*Trichoplax adhaerens*); Ramiro Morales-Hojas (*Drosophila americana*); Steven Haddock, Jellywatch.org (*Hormiphora californensis*); and Tess Linden (*Salpingoeca rosetta*). Source Data are available at [[https://figshare.le.ac.uk/articles/dataset/Monoamine\\_Neuromodulation\\_is\\_a\\_Bilaterian\\_Innovation\\_Results/20391477](https://figshare.le.ac.uk/articles/dataset/Monoamine_Neuromodulation_is_a_Bilaterian_Innovation_Results/20391477)] in Results/GPCRAnalysis/GeneTrees labelled GPCR\_CLANS\_TBE.tbe.tree and GPCR\_CLANS\_UFB.treefile.



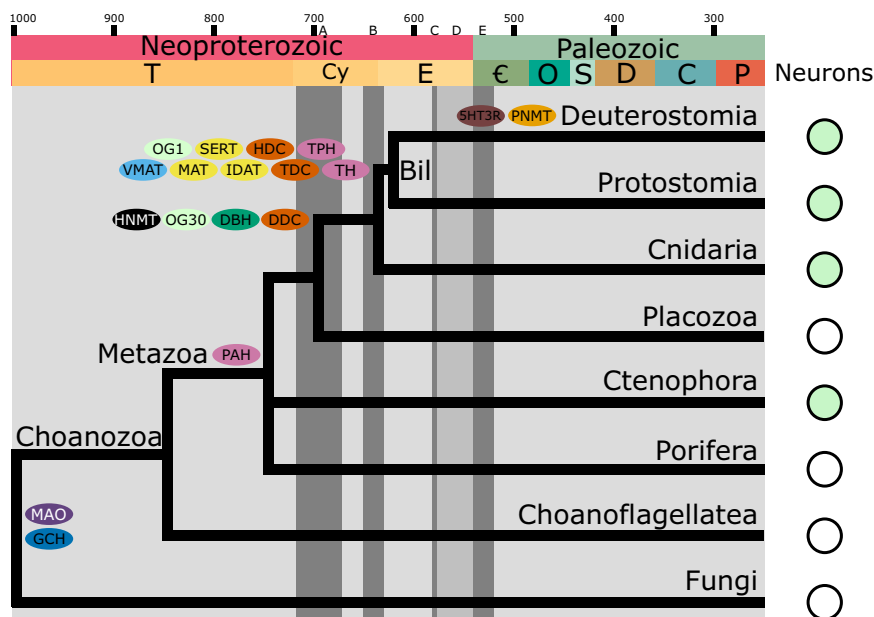
**Fig. 6 | Distribution of monoaminergic system genes. Presence/absence of monoamine pathway genes inferred using the reconciliation analysis.** Colours correspond to orthogroups, with darker shades indicating matches to specific InterProScan profiles (Supplementary Data 4). OG orthogroup, PAH phenylalanine hydroxylase, TPH tryptophan hydroxylase, TH tyrosine hydroxylase, DDC dopa decarboxylase, TDC tyrosine decarboxylase, HDC histidine decarboxylase, DBH dopamine beta hydroxylase, TBH tyramine beta hydroxylase, MOXD mono-oxygenase DBH-like, PNMT phenylethanolamine-*N*-methyltransferase, INMT indolethylamine-*N*-methyltransferase, NNMT nicotinamide-*N*-methyltransferase, VMAT

vesicular monoamine transporter, VACHT vesicular acetylcholine transporter, GCH GTP cyclo-hydroxylase, SLC solute ligand carrier, MAO monoamine oxidase, HNMT histamine-*N*-methyltransferase, SERT serotonin transporter, IDAT invertebrate dopamine transporter, SHT3R 5-hydroxytryptamine receptor 3, INE *Drosophila melanogaster* transporter Ine, COMT catechol-*O*-methyltransferase, GPCR g-protein coupled receptor. Source Data are available at [[https://figshare.le.ac.uk/articles/dataset/Monoamine\\_Neuromodulation\\_is\\_a\\_Bilaterian\\_Innovation\\_Results/20391477](https://figshare.le.ac.uk/articles/dataset/Monoamine_Neuromodulation_is_a_Bilaterian_Innovation_Results/20391477)] in Results/Broccoli labelled table\_OGs\_protein\_names.txt and in Results/InterProScan labelled \*\_IPS.tsv.

observations reporting the presence of monoamines in non-bilaterian metazoans<sup>38</sup> (Fig. 1C). How could non-bilaterians produce monoamines? One possible solution is through the functional flexibility of synthetic enzymes. Recently, studies have identified that PAH orthologs produce serotonin in *Drosophila melanogaster*, *Caenorhabditis elegans*, and mice<sup>9,10,63,64</sup>. Additionally, mutational analyses of mammalian PAH have shown that a few simple mutations can change their substrate specificity from phenylalanine to tryptophan, the key amino acid required for serotonin production<sup>65,66</sup>. Furthermore, alternative pathways could be responsible for the production of monoamines. For example, it has been proposed that the production of L-DOPA (a precursor of dopamine—see Fig. 1A) could be mediated by a tyrosinase in *Hydra*<sup>67</sup> and in the sea anemone *Metridium senile*<sup>68,69</sup>. There is evidence of a similar

pathway functioning in young mice<sup>70,71</sup>. Similarly, *Caenorhabditis elegans* TH-mutants maintain a relatively high dopamine level, suggesting that alternative synthesis pathways exist<sup>5</sup>. Outside animals, other pathways are known, such as cytochrome T5H and other monooxygenases that can produce serotonin and dopamine in plants<sup>72–75</sup>.

Although production is a necessary premise, it is not sufficient to imply function, which depends on the nature (and location) of the receptors involved. Our phylogenetic analysis suggests that the receptors for serotonin, dopamine, epinephrine/norepinephrine, octopamine/tyramine, trace amines and the HRH2s are uniquely present in Bilateria. Thus, they originated from a single gene duplication event in the cnidarian/bilaterian stem group. While the orthologs of histamine receptors 1, 3 and 4 are limited to Bilateria, they appear to be



**Fig. 7 | Synopsis of the evolution of the monoaminergic system.** A simplified time-calibrated species tree illustrating the origin of the key monoaminergic genes and the presence of neurons. A Sturtian glaciation, B Marinoan glaciation, C Gaskiers glaciation, D Occurrence of Ediacaran Biota/early animal fossils, E the Cambrian explosion. Species tree dated following Dos reis et al.<sup>105</sup>, Geological column is shown in accordance with the ICS International Chronostratigraphic Chart (updated 2017)<sup>106</sup>. Bil Bilateria, OG orthogroup, PAH phenylalanine

hydroxylase, TPH tryptophan hydroxylase, DDC dopa decarboxylase, TDC tyrosine decarboxylase, HDC histidine decarboxylase, DBH dopamine beta hydroxylase, PNMT phenylethanolamine-*N*-methyltransferase, VMAT vesicular monoamine transporter, GCH GTP cyclo-hydrolase, MAO monoamine oxidase, HNMT histamine-*N*-methyltransferase, SERT serotonin transporter, MAT monoamine transporter, IDAT invertebrate dopamine transporter, SHT3R 5-hydroxytryptamine receptor 3.

secondarily absent in Cnidaria, with just the ACM receptors conserved across both Cnidaria and Bilateria. We also identified an independent expansion of GPCRs that are unique to Cnidaria but closely related to the canonical monoaminergic receptors. The functional role of these receptors is yet to be determined, but we speculate that they might drive the response to monoamines observed in cnidarians<sup>38</sup>. Despite the considerable progress presented here, we are aware that more remains to be done to unravel the function of monoamine-related enzymes and GPCRs in non-bilaterian animals. Increasing the number of high-quality genomes (e.g., chromosomal level assembly) for many more organisms should enable more precise characterisation of orthology relationships.

In summary, our data indicate that the key genes required for monoamine production, modulation, and reception emerged in the stem-bilaterians. This discovery suggests that the monoaminergic system evolved to the Cryogenian/Ediacaran boundary, about 650–600 Mya (Fig. 7), pre-dating the Cambrian (~540 Mya). In the transition between Ediacaran and Cambrian, the oceans experienced a progressive, likely non-uniform, increase in oxygen concentration<sup>76–80</sup>. The accepted view is that the increased oxidative capacity of the environment favoured the emergence of more energetically intensive and complex modes of life<sup>71</sup>. Compared to previous epochs, the fossil record shows a shift from filter-feeding and low-motility forms to more elaborated body plans compatible with locomotion and predatory behaviour<sup>81–83</sup>. The origin of the monoaminergic system in the Cryogenian/Ediacaran (Fig. 7), the established roles of monoamines in bilaterians<sup>4–7</sup>, and the morphological changes evident in the fossil record<sup>81–83</sup>, lead us to speculate that it could have played a role in the Cambrian diversification for example by providing flexibility of the neural circuits to facilitate the interaction with the environment.

## Methods

This study utilised wholly computational methods using data sourced from publicly accessible databases, as such ethics approval was not necessary for the methods used.

## Species selection

Whole proteomes were downloaded from NCBI and ENSEMBL (see Supplementary Data 2). Additionally, to improve our diversity for Cnidaria, we assembled transcriptomic data for seven species (see Supplementary Data 2) using Trinity<sup>84</sup> (with-trimmomatic argument) and used Transdecoder<sup>85</sup> (LongOrfs tool with default options) to infer the amino acid sequences. We used the BUSCO<sup>48</sup> eukaryota\_odb10 database (default options) of 255 single-copy genes to measure the completeness of the proteomes. We selected 101 species (including 47 animals and 54 other eukaryotes) considering both taxonomical diversity and BUSCO completeness score (see Supplementary Data 2).

## Homologues and orthogroup identification

The KEGG pathways<sup>86</sup> for serotonergic and dopaminergic synapses as well as published literature, were used to identify monoamine pathway genes (See Supplementary Data 4). To identify homologous sequences for each gene of interest, we used BLASTP<sup>87</sup> (default options with -evalue 1e-10, -outfmt 6) with known sequences from SwissProt<sup>88</sup>, EggNOG<sup>89</sup> or KEGG<sup>86</sup> as seeds (detailed in Supplementary Data 4). All BLAST hits with an e-value < 1e<sup>-25</sup> were extracted and annotated using BLASTP against the SwissProt database. BLAST hits for each species were filtered to remove redundant sequences using CD-HIT (-c 1.0 with default options)<sup>90</sup>.

Starting from the putative homologues to identify the orthogroups, we used Broccoli using default parameters and the maximum likelihood tree reconstruction method<sup>91</sup>. Orthogroups were then annotated based on sequences from *Homo sapiens*, *Mus musculus* and *Drosophila melanogaster* (Supplementary Data 3). Furthermore, orthogroups of interest were analysed using InterProScan<sup>92</sup> (-f TSV-goterm-pathways options used) and manually inspected for matches to profiles associated with known monoamine genes (details in Supplementary Data 4).

Initially, we performed the BLAST searches and Broccoli orthology inference using all 101 species and constructed phylogenetic trees on orthogroups of interest. However, phylogenetic analysis (see

Supplementary Figs. 1E 3C, 4E, 5C, 6C, 7E, 8C, 10C, 12C, 14C, 16C, 17C and 18C) indicates that the vast majority of monoaminergic genes are metazoans-specific (Supplementary Data 3). To reduce the computational burden, we focused on 47 animal and 18 opisthokont species.

### GPCR analysis

We assembled a dataset of monoamine receptors by combing the different Broccoli orthogroups with GPCR annotations (see main text for justification). Phobius<sup>93</sup> (-s output) was used to estimate the number of transmembrane domains (TMDs) and sequences with 7TMDs were kept, and CD-hit<sup>90</sup> was used to eliminate homologues with >80% similarity (-c 0.8 with default options). We used CLANS<sup>62</sup> (all-vs-all BLAST calculated using the online tool: <https://toolkit.tuebingen.mpg.de/tools/clans>) to cluster sequences based on their similarity across different stringency thresholds/p-values (from  $1e^{-100}$  to  $1e^{-30}$ ). The clusters were annotated using *Homo sapiens*, *Mus musculus* and *Drosophila melanogaster* sequences. All sequences connecting to the cluster which contained the known aminergic GPCRs at p-value= $1e^{-42}$  were extracted, and opsin sequences were used as the outgroup.

### Phylogenetic analyses

**Gene trees.** Each gene family of interest was aligned using MAFFT<sup>94</sup> using (-auto option and 1,000 max iterations), and we removed sites with >70% gaps using trimAl<sup>95</sup> (-gt 0.3 with default options, see Supplementary Data 5 for alignments statistics calculated using PhyKIT<sup>96</sup>). IQ-TREE2<sup>97</sup> was used to reconstruct the gene trees under the best-fitting model selected with BIC (arguments used: -m MFP -mset Blosum62,cpREV,Dayhoff,DCMut,FLU,HIVb,HIVw,JTT,JTTDCMut,LG,mtART,mtMAM,mtREV,mtZOA,mtMet,mtVer,mtInV,Poisson,PMB,rtREV,VT,WAG,GTR20 -mrate E,I,G,I+G,R -B 1000-wbtl-bnni-alrt 1000-abayes,-alrt 1,000 and -abayes not used in GPCR analyses, Supplementary Data 4). Node support was calculated using 1,000 Ultrafast Bootstrap<sup>49</sup> (UFB) repeats. We also estimated nodal support using transferable bootstrap expectation (TBE)<sup>50</sup> scores from 100 non-parametric bootstrap repeats (-b 100 -tbe, see main text for justification).

**Reconciliation analyses.** Generax<sup>53</sup> was used to reconcile the gene tree with the species tree using the undatedDL model (-r UndatedDL and -unrooted-gene-tree arguments) using the closest proxies to the best-fit models, selected by IQ-TREE2. Specific rate parameters were estimated for each orthogroup (-family-specific-rates option). Reconciliations were visualised using recphyloXML<sup>98</sup>.

To account for the uncertainty in non-bilaterian animal relationships, we performed the reconciliation using the ctenophore-first<sup>99,100</sup> and sponge-first hypothesis<sup>101-103</sup>. Similarly, we performed a reconciliation including the non-opisthokont species. However, given the lack (or instability - see above) of monoaminergic genes in ctenophores and sponges and in general outside metazoans (see Supplementary Data 3), the differences between the scenarios were marginal (Supplementary Material).

### Rogue taxa analysis

To evaluate the presence of problematic sequences<sup>46,52,55,56</sup>, we used two orthogonal methods the t-index<sup>50</sup> and the Leaf stability index<sup>51,104</sup>. The t-index evaluates how often the taxa change position across the bootstrap repeats (i.e., how unstable the sequence is). The Leaf stability index (LSI) uses quartet frequencies to estimate sequence stability, and it was calculated using RogueNaRok<sup>52</sup> (rnr-lsi default options). The LSI was independently estimated using the TBE bootstrap trees and the UFB trees. To identify unstable taxa, we mainly used the t-index and considered sequences with a score > 2 as unstable. We further validated this instability with LSI where, for each sequence, the lower the value, the more unstable it is. Where unstable sequences may have an

influence on the topology or reconciliation, we removed them from the alignment and re-ran phylogenetic analyses as described above.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

All data and results underlying this study have been deposited in the figshare repository under accession codes: 20391462; 20391477. Source data trees for Figs. 2A, 2C, 3A, 3C, 4A and 4C can be found in Figshare repository 20391477 in Results/GeneTrees and are labelled \*\_TBE.tbe.tree and \*\_UFB.treefile. Source data tree for Fig. 5A can be found in Figshare repository 20391477 in Results/GPCRAnalysis/GeneTrees labelled GPCR\_CLANS\_TBE.tbe.tree and GPCR\_CLANS\_UFB.treefile. Source data for Fig. 6 can be found in Figshare repository 20391477 in Results/Broccoli labelled table\_OGs\_protein\_counts.txt and in Results/InterProScan labelled \*\_IPS.tsv. Source data trees for Supplementary Figures can be found in Figshare repository 20391477 in Results/GeneTrees, Results/GPCRAnalysis/GeneTrees, Results/Reconciliations and Results/RogueTaxaAnalysis/GeneTreesReanalysis. Publicly available data/databases used in this study can be accessed at: NCBI (<https://www.ncbi.nlm.nih.gov/>); ENSEMBL (<https://ensemblgenomes.org/>); EukProt (<https://evocellbio.com/eukprot/>); KEGG pathways (<https://www.kegg.jp/kegg/pathway.html>); EggNOG (<http://eggno5.embl.de/#/app/home>); SwissProt (<https://www.uniprot.org/>). Accession codes/sources for publicly available original sequence data used in this study are outlined in Supplementary Data 2.

### Code availability

Computation used in this study used published software, references are provided, and necessary options/parameters are outlined in the Methods section.

### References

1. Flames, N. & Hobert, O. Transcriptional control of the terminal fate of monoaminergic neurons. *Annu. Rev. Neurosci.* **34**, 153–184 (2011).
2. Nagaya, Y., Kutsukake, M., Chigusa, S. I. & Komatsu, A. A trace amine, tyramine, functions as a neuromodulator in *Drosophila melanogaster*. *Neurosci. Lett.* **329**, 324–328 (2002).
3. Libersat, F. & Pflueger, H.-J. Monoamines and the orchestration of behavior. *BioScience* **54**, 17–25 (2004).
4. Wise, R. A. Dopamine, learning and motivation. *Nat. Rev. Neurosci.* **5**, 483–494 (2004).
5. Sanyal, S. et al. Dopamine modulates the plasticity of mechanosensory responses in *Caenorhabditis elegans*. *EMBO J.* **23**, 473–482 (2004).
6. Di Giovanni, G., Chagraoui, A., Bharatiya, R. & De Deurwaerdère, P. in *Handbook of Behavioral Neuroscience* (eds. Müller, C. P. & Cunningham, K. A.) Vol. 31, 197–215 (Elsevier, 2020).
7. Quraishi, S. A. & Paladini, C. A. in *Handbook of Behavioral Neuroscience* (eds. Steiner, H. & Tseng, K. Y.) Vol. 24, 361–372 (Elsevier, 2016).
8. Azmitia, E. C. in *Handbook of Behavioral Neuroscience* (eds. Müller, C. P. & Jacobs, B. L.) vol. 21 3–22 (Elsevier, 2010).
9. Coleman, C. M. & Neckameyer, W. S. Serotonin synthesis by two distinct enzymes in *Drosophila melanogaster*. *Arch. Insect Biochem. Physiol.* **59**, 12–31 (2005).
10. Mordhorst, A. et al. Phenylalanine hydroxylase contributes to serotonin synthesis in mice. *FASEB J.* **35**, e21648 (2021).
11. Cao, J., Shi, F., Liu, X., Huang, G. & Zhou, M. Phylogenetic analysis and evolution of aromatic amino acid hydroxylase. *FEBS Lett.* **584**, 4775–4782 (2010).
12. Huang, H., Li, Y., Liang, J. & Finkelman, F. D. Molecular regulation of histamine synthesis. *Front. Immunol.* **9** (2018).

13. Cole, S. H. et al. Two functional but noncomplementing drosophila tyrosine decarboxylase genes: distinct roles for neural tyramine and octopamine in female fertility \*. *J. Biol. Chem.* **280**, 14948–14955 (2005).
14. Livingstone, M. S. & Tempel, B. L. Genetic dissection of monoamine neurotransmitter synthesis in Drosophila. *Nature* **303**, 67–70 (1983).
15. Roeder, T. Octopamine in invertebrates. *Prog. Neurobiol.* **59**, 533–561 (1999).
16. Joh, T. H. & Hwang, O. Dopamine beta-hydroxylase: biochemistry and molecular biology. *Ann. N. Y. Acad. Sci.* **493**, 342–350 (1987).
17. Eiden, L. E., Schäfer, M. K.-H., Weihe, E. & Schütz, B. The vesicular amine transporter family (SLC18): amine/proton antiporters required for vesicular accumulation and regulated exocytotic secretion of monoamines and acetylcholine. *Pflüg. Arch.* **447**, 636–640 (2004).
18. Erickson, J. D., Eiden, L. E. & Hoffman, B. J. Expression cloning of a reserpine-sensitive vesicular monoamine transporter. *Proc. Natl. Acad. Sci.* **89**, 10993–10997 (1992).
19. Peter, D. et al. Chromosomal localization of the human vesicular amine transporter genes. *Genomics* **18**, 720–723 (1993).
20. Bauknecht, P. & Jékely, G. Ancient coexistence of norepinephrine, tyramine, and octopamine signaling in bilaterians. *BMC Biol.* **15**, 6 (2017).
21. Gainetdinov, R. R., Hoener, M. C. & Berry, M. D. Trace amines and their receptors. *Pharmacol. Rev.* **70**, 549–620 (2018).
22. Missale, C., Nash, S. R., Robinson, S. W., Jaber, M. & Caron, M. G. Dopamine receptors: from structure to function. *Physiol. Rev.* **78**, 189–225 (1998).
23. Nichols, D. E. & Nichols, C. D. Serotonin receptors. *Chem. Rev.* **108**, 1614–1641 (2008).
24. Pucadyil, T. J., Kalipatnapu, S. & Chattopadhyay, A. The Serotonin1A AReceptor: a representative member of the serotonin receptor family. *Cell. Mol. Neurobiol.* **25**, 553–580 (2005).
25. Neve, K. A., Seamans, J. K. & Trantham-Davidson, H. Dopamine receptor signaling. *J. Recept. Signal Transduct.* **24**, 165–205 (2004).
26. Caveney, S., Cladman, W., Verellen, L. & Donly, C. Ancestry of neuronal monoamine transporters in the Metazoa. *J. Exp. Biol.* **209**, 4858–4868 (2006).
27. Blakely, R. D. et al. Cloning and expression of a functional serotonin transporter from rat brain. *Nature* **354**, 66–70 (1991).
28. Kilty, J. E., Lorang, D. & Amara, S. G. Cloning and expression of a cocaine-sensitive rat dopamine transporter. *Science* **254**, 578–579 (1991).
29. Pacholczyk, T., Blakely, R. D. & Amara, S. G. Expression cloning of a cocaine- and antidepressant-sensitive human noradrenaline transporter. *Nature* **350**, 350–354 (1991).
30. Chen, N. & Reith, M. E. A. Structure and function of the dopamine transporter. *Eur. J. Pharmacol.* **405**, 329–339 (2000).
31. Geha, R. M., Rebrin, I., Chen, K. & Shih, J. C. Substrate and inhibitor specificities for human monoamine oxidase A and B are influenced by a single amino acid\*. *J. Biol. Chem.* **276**, 9877–9882 (2001).
32. AXELROD, J. & TOMCHICK, R. Enzymatic O-methylation of epinephrine and other catechols. *J. Biol. Chem.* **233**, 702–705 (1958).
33. Brown, D. D., Tomchick, R. & Axelrod, J. The distribution and properties of a histamine-methylating enzyme. *J. Biol. Chem.* **234**, 2948–2950 (1959).
34. Jeffery, D. R. & Roth, J. A. Characterization of membrane-bound and soluble catechol-O-methyltransferase from human frontal cortex. *J. Neurochem.* **42**, 826–832 (1984).
35. Girard, B. et al. Human histamine N-methyltransferase pharmacogenetics: cloning and expression of kidney cDNA. *Mol. Pharmacol.* **45**, 461–468 (1994).
36. Ma, J. et al. Structure of rat monoamine oxidase A and its specific recognitions for substrates and inhibitors. *J. Mol. Biol.* **338**, 103–114 (2004).
37. Moroz, L. L., Romanova, D. Y. & Kohn, A. B. Neural versus alternative integrative systems: molecular insights into origins of neurotransmitters. *Philos. Trans. R. Soc. B: Biol. Sci.* **376**, 20190762 (2021).
38. Kanaya, H. J. et al. A sleep-like state in Hydra unravels conserved sleep mechanisms during the evolutionary development of the central nervous system. *Sci. Adv.* **6**, eabb9415 (2020).
39. Francis, W. R. et al. The genome of the contractile demosponge *Tethya wilhelma* and the evolution of metazoan neural signalling pathways. *bioRxiv* <https://doi.org/10.1101/120998> (2017).
40. Moroz, L. L. & Kohn, A. B. Unbiased view of synaptic and neuronal gene complement in ctenophores: are there pan-neuronal and pan-synaptic genes across metazoa? *Integr. Comp. Biol.* **55**, 1028–1049 (2015).
41. Anctil, M. Chemical transmission in the sea anemone *Nematostella vectensis*: a genomic perspective. *Comp. Biochem. Physiol. D: Genomics Proteomics* **4**, 268–289 (2009).
42. Iyer, L. M., Aravind, L., Coon, S. L., Klein, D. C. & Koonin, E. V. Evolution of cell–cell signaling in animals: did late horizontal gene transfer from bacteria have a role? *Trends Genet.* **20**, 292–299 (2004).
43. Siltberg-Liberles, J., Steen, I. H., Svebak, R. M. & Martinez, A. The phylogeny of the aromatic amino acid hydroxylases revisited by characterizing phenylalanine hydroxylase from *Dictyostelium discoideum*. *Gene* **427**, 86–92 (2008).
44. Kutchko, K. M. & Siltberg-Liberles, J. Metazoan innovation: from aromatic amino acids to extracellular signaling. *Amino Acids* **45**, 359–367 (2013).
45. Krishnan, A. & Schiöth, H. B. The role of G protein-coupled receptors in the early evolution of neurotransmission and the nervous system. *J. Exp. Biol.* **218**, 562–571 (2015).
46. Felsenstein, J. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* **27**, 401–410 (1978).
47. Fleming, J. F., Feuda, R., Roberts, N. W. & Pisani, D. A novel approach to investigate the effect of tree reconstruction artifacts in single-gene analysis clarifies opsin evolution in nonbilaterian metazoans. *Genome Biol. Evol.* **12**, 3906–3916 (2020).
48. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
49. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
50. Lemoine, F. et al. Renewing Felsenstein’s phylogenetic bootstrap in the era of big data. *Nature* **556**, 452–456 (2018).
51. Thorley, J. L. & Wilkinson, M. Testing the phylogenetic stability of early tetrapods. *J. Theor. Biol.* **200**, 343–344 (1999).
52. Aberer, A. J., Krompass, D. & Stamatakis, A. Pruning rogue taxa improves phylogenetic accuracy: an efficient algorithm and webservice. *Syst. Biol.* **62**, 162–166 (2013).
53. Morel, B., Kozlov, A. M., Stamatakis, A. & Szöllösi, G. J. GeneRax: a tool for species-tree-aware maximum likelihood-based gene family tree inference under gene duplication, transfer, and loss. *Mol. Biol. Evol.* **37**, 2763–2774 (2020).
54. Williams, T. A. et al. The power and limitations of species tree-aware phylogenetics. <https://doi.org/10.1101/2023.03.17.533068> (2023).
55. Wilkinson, M. Majority-rule reduced consensus trees and their use in bootstrapping. *Mol. Biol. Evol.* **13**, 437–444 (1996).
56. Sanderson, M. J. & Shaffer, H. B. Troubleshooting molecular phylogenetic analyses. *Annu. Rev. Ecol. Syst.* **33**, 49–72 (2002).

57. Martin, J. L., Begun, J., McLeish, M. J., Caine, J. M. & Grunewald, G. L. Getting the adrenaline going: crystal structure of the adrenaline-synthesizing enzyme PNMT. *Structure* **9**, 977–985 (2001).
58. Bortolato, M., Chen, K. & Shih, J. C. Monoamine oxidase inactivation: from pathophysiology to therapeutics. *Adv. Drug Deliv. Rev.* **60**, 1527–1533 (2008).
59. Tunbridge, E. M., Harrison, P. J. & Weinberger, D. R. Catechol-o-methyltransferase, cognition, and psychosis: Val158Met and beyond. *Biol. Psychiatry* **60**, 141–151 (2006).
60. Cunningham, C. L. et al. The murine catecholamine methyltransferase mTOMT is essential for mechanotransduction by cochlear hair cells. *eLife* **6**, e24318 (2017).
61. Heidari, A. et al. Mutations in the histamine N-methyltransferase gene, HNMT, are associated with nonsyndromic autosomal recessive intellectual disability. *Hum. Mol. Genet.* **24**, 5697–5710 (2015).
62. Frickey, T. & Lupas, A. CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics* **20**, 3702–3704 (2004).
63. Alcañiz, S. & Silva, F. J. Phenylalanine hydroxylase participation in the synthesis of serotonin and pteridines in *Drosophila melanogaster*. *Comp. Biochem. Physiol. C. Pharmacol. Toxicol. Endocrinol.* **116**, 205–212 (1997).
64. Yu, J. et al. Parallel pathways for serotonin biosynthesis and metabolism in *C. elegans*. *Nat. Chem. Biol.* 1–10 (2022) <https://doi.org/10.1038/s41589-022-01148-7> (2022).
65. Teigen, K., Alan McKinney, J., Haavik, J. & Martinez, A. Selectivity and affinity determinants for ligand binding to the aromatic amino acid hydroxylases. *Curr. Med. Chem.* **14**, 455–467 (2007).
66. Hufton, S. E., Jennings, I. G. & Cotton, R. G. Structure and function of the aromatic amino acid hydroxylases. *Biochem. J.* **311**, 353–366 (1995).
67. Carlberg, M. Localization of dopamine in the freshwater hydrozoan *Hydra attenuata*. *Cell Tissue Res.* **270**, 601–607 (1992).
68. Carlberg, M. & Elofsson, R. Presence of 3,4-dihydroxyphenylalanine and 3,4,5-trihydroxyphenylalanine in a coelenterate nervous system: Possible tyrosinase-mediated accumulation. *Neurochem. Int.* **11**, 161–167 (1987).
69. Carlberg, M., Jergil, B., Lindblad, C. & Rosengren, E. Enzymatic 5-hydroxylation of l-dopa by a tyrosinase isolated from the sea anemone *Metridium senile*. *Gen. Pharmacol. Vasc. Syst.* **15**, 301–307 (1984).
70. Rios, M. et al. Catecholamine synthesis is mediated by tyrosinase in the absence of tyrosine hydroxylase. *J. Neurosci.* **19**, 3519–3526 (1999).
71. Eisenhofer, G. et al. Tyrosinase: a developmentally specific major determinant of peripheral dopamine. *FASEB J.* **17**, 1248–1255 (2003).
72. Erland, L. A. E., Turi, C. E. & Saxena, P. K. in *Serotonin* (ed. Pilowsky, P. M.) 23–46 (Academic Press, 2019). <https://doi.org/10.1016/B978-0-12-800050-2.00002-4>.
73. Kang, K., Kang, S., Lee, K., Park, M. & Back, K. Enzymatic features of serotonin biosynthetic enzymes and serotonin biosynthesis in plants. *Plant Signal. Behav.* **3**, 389–390 (2008).
74. Kang, S., Kang, K., Lee, K. & Back, K. Characterization of tryptamine 5-hydroxylase and serotonin synthesis in rice plants. *Plant Cell Rep.* **26**, 2009–2015 (2007).
75. Liu, Q. et al. Functions of dopamine in plants: a review. *Plant Signal. Behav.* **15**, 1827782 (2020).
76. Erwin, D. H. & Valentine, J. W. *The Cambrian Explosion: the Construction of Animal biodiversity*. (Roberts, 2012).
77. Darroch, S. A. F., Smith, E. F., Laflamme, M. & Erwin, D. H. Ediacaran extinction and cambrian explosion. *Trends Ecol. Evol.* **33**, 653–663 (2018).
78. Erwin, D. H. Early metazoan life: divergence, environment and ecology. *Philos. Trans. R. Soc. B: Biol. Sci.* **370**, 20150036 (2015).
79. Wood, R. et al. Integrated records of environmental change and evolution challenge the Cambrian Explosion. *Nat. Ecol. Evol.* **3**, 528–538 (2019).
80. Droser, M. L., Tarhan, L. G. & Gehling, J. G. The rise of animals in a changing environment: global ecological innovation in the Late Ediacaran. *Annu. Rev. Earth Planet. Sci.* **45**, 593–617 (2017).
81. Plotnick, R. E., Dornbos, S. Q. & Chen, J. Information landscapes and sensory ecology of the Cambrian Radiation. *Paleobiology* **36**, 303–317 (2010).
82. Hsieh, S., Plotnick, R. E. & Bush, A. M. The Phanerozoic aftermath of the Cambrian information revolution: sensory and cognitive complexity in marine faunas. *Paleobiology* **48**, 397–419 (2022).
83. Carbone, C. & Narbonne, G. M. When life got smart: the evolution of behavioral complexity through the Ediacaran and Early Cambrian of NW Canada. *J. Paleontol.* **88**, 309–330 (2014).
84. Grabherr, M. G. et al. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* **29**, 644–652 (2011).
85. Transdecoder:: Anaconda Cloud. <https://anaconda.org/bioconda/transdecoder> (2020).
86. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361 (2017).
87. Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
88. Gasteiger, E., Jung, E. & Bairoch, A. SWISS-PROT: Connecting Biomolecular Knowledge Via a Protein Database. *Curr. Issues Mol. Biol.* **3**, 47–55 (2001).
89. Huerta-Cepas, J. et al. EggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–D314 (2019).
90. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
91. Derelle, R., Philippe, H. & Colbourne, J. K. Broccoli: combining phylogenetic and network analyses for orthology assignment. *Mol. Biol. Evol.* **37**, 3389–3396 (2020).
92. Jones, P. et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
93. Käll, L., Krogh, A. & Sonnhammer, E. L. L. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucleic Acids Res.* **35**, W429–W432 (2007).
94. Katoh, K., Kuma, K., Toh, H. & Miyata, T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **33**, 511–518 (2005).
95. Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
96. Steenwyk, J. L. et al. PhyKIT: a broadly applicable UNIX shell toolkit for processing and analyzing phylogenomic data. *Bioinformatics* **37**, 2325–2331 (2021).
97. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
98. Duchemin, W. et al. RecPhyloXML: a format for reconciled gene trees. *Bioinformatics* **34**, 3646–3652 (2018).
99. Dunn, C. W. et al. Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* **452**, 745–749 (2008).
100. Li, Y., Shen, X.-X., Evans, B., Dunn, C. W. & Rokas, A. Rooting the animal tree of life. *Mol. Biol. Evol.* **38**, 4322–4333 (2021).

101. Feuda, R. et al. Improved modeling of compositional heterogeneity supports sponges as sister to all other animals. *Curr. Biol.* **27**, 3864–3870.e4 (2017).
102. Pisani, D. et al. Genomic data do not support comb jellies as the sister group to all other animals. *Proc. Natl Acad. Sci.* **112**, 15402–15407 (2015).
103. Giacomelli, M., Rossi, M. E., Lozano-Fernandez, J., Feuda, R. & Pisani, D. Resolving tricky nodes in the tree of life through amino acid recoding. *iScience* **25**, 105594 (2022).
104. Wilkinson, M. Identifying stable reference taxa for phylogenetic nomenclature. *Zool. Scr.* **35**, 109–112 (2006).
105. dos Reis, M. et al. Uncertainty in the timing of origin of animals and the limits of precision in molecular timescales. *Curr. Biol.* **25**, 2939–2950 (2015).
106. Cohen, K. M., Finney, S. C., Gibbard, P. L. & Fan, J.-X. The ICS International Chronostratigraphic Chart. *Episodes J. Int. Geosci.* **36**, 199–204 (2013).

## Acknowledgements

This study was supported by a Royal Society University Research Fellowship and Grants (UF160226 and RGF\EA\180052) to R.F., by the Swiss National Science Foundation (Grants IZCOZO\_182957 and 310030\_188471) to S.G.S., by the European Union's Horizon 2020 research and innovation programme (Marie Skłodowska-Curie grant agreement No 765937) to E.R. and by a PhD studentship from the University of Leicester to M.G. This research used the ALICE High-Performance Computing Facility at the University of Leicester.

## Author contributions

R.F. and M.G. conceived the study. M.G. and G.B. performed the analyses. R.F., M.G., G.B., E.R. and S.S. performed the data interpretation. R.F. and M.G. wrote the main text with the help of G.B., E.R. and S.S.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-39030-2>.

**Correspondence** and requests for materials should be addressed to Roberto Feuda.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023