Horizontal Transfer and Gene Loss Shaped the Evolution of Alpha-Amylases in Bilaterians

Andrea Desiderato,*† Marcos Barbeitos,* Clément Gilbert,† and Jean-Luc Da Lage‡,*

*Graduate Program in Zoology, Zoology Department, Federal University of Paraná, CP 19020, Curitiba, Paraná 81531-980, Brazil, †Department of Functional Ecology, Alfred Wegener Institute & Helmholtz Centre for Polar and Marine Research, Am Handelshafen 12, 27570 Bremerhaven, Germany, and ‡Évolution, Génomes, Comportement, Écologie. CNRS, IRD, Université Paris-Sud. Université Paris-Saclay. F-91198 Gif-sur-Yvette, France

ORCID IDs: 0000-0003-3284-5529 (A.D.); 0000-0001-6488-8254 (M.B.); 0000-0002-2131-7467 (C.G.); 0000-0002-1372-6762 (J.-L.D.L.)

ABSTRACT The subfamily GH13_1 of alpha-amylases is typical of Fungi, but it is also found in some unicellular eukaryotes (e.g., Amoebozoa, choanoflagellates) and non-bilaterian Metazoa. Since a previous study in 2007, GH13_1 amylases were considered ancestral to the Unikonts, including animals, except Bilateria, such that it was thought to have been lost in the ancestor of this clade. The only alpha-amylases known to be present in Bilateria so far belong to the GH13_15 and 24 subfamilies (commonly called bilaterian alpha-amylases) and were likely acquired by horizontal transfer from a proteobacterium. The taxonomic scope of Eukaryota genomes in databases has been greatly increased ever since 2007. We have surveyed GH13_1 sequences in recent data from ca. 1600 bilaterian species, 60 non-bilaterian animals and also in unicellular eukaryotes. As expected, we found a number of those sequences in non-bilaterians: Anthozoa (Cnidaria) and in sponges, confirming the previous observations, but none in jellyfishes and in Ctenophora. Our main and unexpected finding is that such fungal (also called Dictyo-type) amylases were also consistently retrieved in several bilaterian phyla: hemichordates (deuterostomes), brachiopods and related phyla, some molluscs and some annelids (protostomes). We discuss evolutionary hypotheses possibly explaining the scattered distribution of GH13_1 across bilaterians, namely, the retention of the ancestral gene in those phyla only and/or horizontal transfers from non-bilaterian donors.

KEYWORDS alpha-amylase gene loss horizontal gene transfer hemichordates brachiopods phoronids bryozoans molluscs annelids Bilateria glycosyl hydrolase introns

Alpha-amylases are enzymes that are almost ubiquitous in the living world, where they perform the hydrolysis of starch and related polysaccharides into smaller molecules, to supply energy to the organism through digestion. They belong to glycosyl hydrolases, a very large group of enzymes which have been classified in a number of families according to their structures, sequences, catalytic activities and catalytic mechanisms (Henrissat and Davies 1997). Most alpha-amylases are members of the glycoside hydrolase family 13 (GH13), which includes enzymes that can either break down or synthesize α-1,4-, α-1,6- and, less commonly, α-1,2- and α-1,3-glycosidic linkages. Sucrose and trehalose are also substrates for enzymes of this family (MacGregor et al. 2001). The numerous family GH13 is divided into 42 subfamilies, of which only three occur in Metazoans: GH13_1, GH13_15 and GH13_24 (Stam et al. 2006; Da Lage et al. 2007; Lombard et al. 2014). The latter two include the common animal alpha-amylases, while the former was first described in Fungi for which it represents the canonical alpha-amylase (Stam et al. 2006). Da Lage et al. (2007) described the subfamilies GH13_15/24 as private to Bilateria among metazoans. In the same article, they retrieved sequences belonging to the subfamily GH13_1 from the sponge Amphimedon queenslandica (named Reniera sp. in their paper) and the sea anemone Nematostella vectensis, besides the unikont choanoflagellates and amoebozoans, and also excavates and ciliates. They dubbed “Dictyo-type” this alpha-amylase, referring to the slime mold Dictyostelium discoideum (Amoebozoa Mycetozoa). The authors proposed that this amylase, ancestral to the Unikont clade, is shared among non-bilaterian metazoans (e.g., sponges, sea anemones and corals, and Placozoa), but was replaced in Bilateria by
an alpha-amylase of bacterial origin, whose sequence is close to the typical animal amylases.

Given that a wealth of new genomes have been sequenced in the twelve years after that publication, we decided to explore again the diversification of this enzyme subfamily among the Eukaryota. We will focus mainly on Metazoa, in which we show unexpected situations of co-occurrence of both subfamilies GH13_1 and GH13_15/24 in the same genomes. We will discuss two mutually exclusive explanations that may be proposed: either the retention of the ancestral GH13_1 gene along with the typical bilaterian GH13_15/24 in some phyla, or horizontal transfer(s) from non-bilaterian animal donor(s) which would have to be identified.

MATERIALS AND METHODS

In order to further characterize the distribution of GH13_1 genes in Metazoa, we used the sequence of the sponge Amphimedon queenslandica GH13_1 (GenBank XP_019851448) as a query to perform BLASTP and TBLASTN searches on various online data bases available in Genbank (nr, proteins, genomes, assembly, SRA, TSA, WGS), and also in the more specialized databases compagen.org, marinegenomics.oist.jp, reefgenomics.org, marisma.obs-vlfr.fr, vectorbase.org, PdumBase (pдумbase.gdcb.iastate.edu), AmpuBase (https://www.comp.hku.edu.hk/~db/AmpuBase/index.php) (Ip et al. 2018), between October 2018 and August 2019. Fungi were not searched further in this study because they are known to have a GH13_1 member as the usual alpha-amylase. To increase the chances to retrieve potential cnidarian or ctenophoran sequences, the starlet sea anemone, *C. tenuis* (Janecek et al. 1994; Janecek et al. 2014; MacGregor 2001; van der Kaaij et al. 2006). The tree was drawn at the iTOL website (Letunic and Bork 2015). In some cases, for unannotated genes, the N-terminal and/or the C-terminal parts of the retrieved genomic sequences were uncertain, and were not retained in the analyses.

Alignments were performed using MUSCLE (Edgar 2004), as implemented in Geneious (Biomatters Ltd.). A maximum likelihood (ML) tree was built using PhyML’s (Guindon and Gascuel 2003) current implementation at the phylogeny.fr portal (Dereeper et al. 2008). To this end, we first trimmed the N-terminal protein sequences up to the first well conserved motif LLTDR. C-terminal parts were also truncated at the last well aligned stretch. Gaps were removed from the alignments and data were analyzed under WAG (Whelan and Goldman 2001) with among-site rate variation modeled by four discrete rate categories sampled from a gamma distribution. Both the alpha parameter and the proportion of invariable sites were estimated from the data. The robustness of the nodes was estimated using an approximate likelihood ratio test (aLRT) (Anisimova and Gascuel 2006). The tree was drawn at the iTOL website (Letunic and Bork 2016). Metazoans and choanoflagellates were clustered as the ingroup.

Data availability

The protein sequences, Fasta alignment and Newick-formatted tree are available at figshare: https://figshare.com/articles/GH13_1_metazoa/9959369. Supplemental material available at figshare: https://figshare.com/articles/Suppl_data_GH13_1/9975956.

RESULTS

The sequences retrieved from the databases are listed in Table 1. The metazoans investigated are listed in Tables S1 (non-bilaterians) and S2 (bilaterians) with indication of the current state of genome/transcriptome sequencing, the database, the presence or absence of GH13_1 sequences, and the number of gene copies, where possible. A general protein alignment of the sequences found in this study along with already known GH13_1 sequences is shown in Fig. S2.

**GH13_1 sequences retrieved from unicellular taxa**

We confirmed the presence of GH13_1 in dictyostelids, in ciliates and also in oomycetes, some representatives of which (but not all) are indicated in Table 1. In two oomycetes, *Saprolegnia diclina* and *Achlya hypogynia*, the GH13_1-like sequences were the C-terminal half of longer sequences, the N-terminal half of which was similar to unclassified GH13 sequences found in e.g., *Acanthamoeba histolytica* (GenBank accession BAN39582), according to the CAZy database.
Table 1  GH13_1-like sequences found after BLAST searches in online databases (not comprehensive for unicells, without the Fungi). *: sequences which have not been characterized as protein-coding, in sequenced genomes with long contigs; (1): from short DNA sequences (except Sequence reads archive); †: reported as GH13_1 in CAZy. Most of the SRA data are from transcriptome studies; see Tables S1 and S2.

| Phylum                        | Species                        | Database            | Accession          |
|-------------------------------|--------------------------------|---------------------|--------------------|
| NON BILATERIAN METAZOA        |                                |                     |                    |
| Porifera Demospongiae         | Amphimedon queenslandica        | GenBank proteins    | XP_019851448       |
| Heteroscleromorpha            |                                 |                     | m.29963 g.29963    |
| Porifera Demospongiae         | Ephydatia mueller (1)           | Compagen.org        |                    |
| Heteroscleromorpha            |                                 |                     |                    |
| Porifera Demospongiae         | Haliclona tubifera             | GenBank TSA         | GFAV01017079       |
| Heteroscleromorpha            |                                 |                     |                    |
| Porifera Demospongiae         | Spongilla lacustris             | GenBank SRA         | SRX470277          |
| Heteroscleromorpha            |                                 |                     |                    |
| Porifera Demospongiae         | Xestospongia testudinaria (1)   | Reefgenomics.org    | gntBIL_0RD_ID6299  |
| Heteroscleromorpha            |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Anthopleura elegantissima      | GenBank TSA         | GEVE01039432       |
| Actinaria                      |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Anthopleura buddemeieri (1)     | Reefgenomics.org    | c117986_g2_i1      |
| Actinaria tenebrosa           |                                 |                     | c88768_g1_i1       |
| Cnidaria Hexacorallaria       | Calliactis polypus (1)          | Reefgenomics.org    | ecm.model.scaffold_206.7 |
| Actinaria tenebrosa           |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Exaiphtasia pallida            | GenBank proteins    | XP_020895894       |
| Scyphozoa                      |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Nemastrella vectensis           | GenBank proteins    | XP_001629956       |
| Scyphozoa                      |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Stychochactyla heliantus        | GenBank TSA         | GKNY01117022       |
| Hexacorallaria                |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Aulactinia verata (1)           | Reefgenomics.org    | C36117_g1_i2       |
| Hexacorallaria                |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Pachyseris speciosa*            | Reefgenomics.org    | Sc0001227 74283-80000 |
| Corallimorpharia              |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Acropora digitifera             | GenBank proteins    | XP_015760547 partial |
| Scheritaria                   |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Acropora millepora              | GenBank proteins    | XP_029201467       |
| Scheritaria                   |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Acropora tenuis*                | GenBank proteins    | aten_0.1.m1.3059.m1 |
| Scheritaria                   |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Fungia sp.*                     | GenBank proteins    | fflu.1.m.15065.m1   |
| Scheritaria                   |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Goniastrea aspera*              | GenBank proteins    | gasp1.m.3.6500.m1   |
| Scheritaria                   |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Orbicella faveolata             | GenBank proteins    | TR260650_g2_i3     |
| Scheritaria                   |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Pocillopora damicornis           | GenBank genomes     | XP_020268431       |
| Scheritaria                   |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Portes lutea*                   | Reefgenomics.org    | plutz2.m.18168.m1   |
| Scheritaria                   |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Portes rus                      | GenBank genomes     | ORKRP01000157       |
| Scheritaria                   |                                 |                     |                    |
| Cnidaria Hexacorallaria       | Stylophora pistillata           | GenBank proteins    | XP_022802004       |
| Scheritaria                   |                                 |                     |                    |
| Cnidaria Octocorallaria       | Renilla koellikeri              | GenBank SRA         | SRX4346409         |
| Pennatulacea                   |                                 |                     |                    |
| Cnidaria Octocorallaria       | Renilla muelleri                | GenBank SRA         | SRX4717871         |
| Pennatulacea                   |                                 |                     |                    |
| Cnidaria Octocorallaria       | Renilla reniformis*             | GenBank genomes     | FXAL01159338       |
| Pennatulacea                   |                                 |                     | XP_002114911       |
| Placozoa                      | Trichoplax adhaerens            | GenBank proteins    |                    |
|                               |                                 |                     |                    |
| BILATERIA                      |                                |                     |                    |
| Brachiopoda Linguliformea     | Glottidia pyramidata           | GenBank SRA         | SRX731468          |
|                               |                                 |                     |                    |
| Brachiopoda Linguliformea     | Lingula anatina                 | GenBank proteins    | XP_013396432       |
|                               |                                 |                     |                    |
| Brachiopoda Linguliformea     | Lingula anatina                 | GenBank proteins    | XP_013378610       |
|                               |                                 |                     |                    |
| Brachiopoda Craniformea       | Novocrania anomalae            | GenBank SRA         | SRX731472          |
|                               |                                 |                     |                    |
| Brachiopoda Rhyynchonelliforme| Kraussina rubra                 | GenBank SRA         | SRX7312037         |
|                               |                                 |                     |                    |
| Brachiopoda Rhyynchonelliforme| Macandrevia cranium            | GenBank SRA         | SRX731471          |
|                               |                                 |                     |                    |
| Brachiopoda Rhyynchonelliforme| Hemithiris psittacea            | GenBank SRA         | SRX731469          |
|                               |                                 |                     |                    |
| Brachiopoda Rhyynchonelliforme| Terebratalia transversa         | GenBank SRA         | SRX1307070         |
|                               |                                 |                     |                    |
| Brachiopoda Phoroniformea or  | Phoronis australis              | marinegenomics      | g9986.t1           |
| Phoronida                      |                                 |                     |                    |
|                               |                                 |                     |                    |
| Brachiopoda Phoroniformea or  | Phoronis australis              | marinegenomics      | g16048.t1          |
| Phoronida                      |                                 |                     |                    |
|                               |                                 |                     |                    |
| Bryozoa Flustrina              | Bugula neritina                 | GenBank SRA         | SRX1121914         |
|                               |                                 |                     |                    |
| Bryozoa Flustrina              | Bugula stolonifera              | GenBank SRA         | SRX4648236         |
|                               |                                 |                     |                    |
| Bryozoa Ctenostomatida         | Flustellida corniculata         | GenBank SRA         | SRX4648237         |
|                               |                                 |                     |                    |
| Bryozoa Chelostomatida         | Membranipora membranacea        | GenBank SRA         | SRX1121923         |
|                               |                                 |                     |                    |
| Hemichordata Enteropneusta    | Ptychodera flava                | Marinegenomics      | pfl_40v0_9_20150316_1g2314.t1 |
|                               |                                 |                     | GenBank WGS         | LD343027 41534-50098 |
| Hemichordata Enteropneusta    |                                 |                     |                    |
|                               |                                 |                     |                    |
| Hemichordata Enteropneusta    | Ptychodera flava                | Marinegenomics      | pfl_40v0_9_20150316_1g6997.t1 |
|                               |                                 |                     | GenBank WGS         | LD343027 51007-66347 |
| Hemichordata Enteropneusta    |                                 |                     |                    |
|                               |                                 |                     |                    |
| Hemichordata Enteropneusta    | Saccoglossus kowalevskii        | GenBank proteins    | XAV816681         |
|                               |                                 |                     |                    |
| Hemichordata Enteropneusta    | Saccoglossus kowalevskii        | GenBank proteins    | XP_006819810        |
|                               |                                 |                     |                    |
| Hemichordata Enteropneusta    | Saccoglossus kowalevskii        | GenBank proteins    | XP_006819810        |
|                               |                                 |                     |                    |
| Hemichordata Enteropneusta    | Schizocardium californicum      | GenBank SRA         | SRX1436000         |

(continued)
| Phylum | Species | Database | Accession |
|--------|---------|----------|-----------|
| **Hemichordata Enteropneusta** | Torquatoridae antarctica | GenBank SRA | SRX798197 |
| **Hemichordata Pterobranchia** | Rhadobdopleura sp. | GenBank SRA | SRX879690 |
| **Mollusca Gastropoda Caenogastropoda** | Asolene platae | AmpuBase | Api52885 |
| **Mollusca Gastropoda Caenogastropoda** | Batillaria attra | GenBank SRA | SRX2957288 |
| **Mollusca Gastropoda Caenogastropoda** | Charonia tritonis | GenBank SRA | SRX2735455 |
| **Mollusca Gastropoda Caenogastropoda** | Conus tribelli | GenBank WGS | LFLW010536118 |
| **Mollusca Gastropoda Caenogastropoda** | Crepidula novicella | GenBank SRA | GELE01086894 |
| **Mollusca Gastropoda Caenogastropoda** | Glaussolax didyma | GenBank SRA | SRRX5277776 |
| **Mollusca Gastropoda Caenogastropoda** | Hemifusus tuba | GenBank SRA | ERX1318227 |
| **Mollusca Gastropoda Caenogastropoda** | Lanistes nyssanus | AmpuBase | Lny24710 |
| **Mollusca Gastropoda Caenogastropoda** | Marisa comuaires | AmpuBase | Mco2627 |
| **Mollusca Gastropoda Caenogastropoda** | Melainodium tuberculata | GenBank SRA | SRRX5832309 |
| **Mollusca Gastropoda Caenogastropoda** | Neveita didyma | GenBank TSA | GHHQ01002371 |
| **Mollusca Gastropoda Caenogastropoda** | Nucella lapillus | GenBank SRA | SRX4378318 |
| **Mollusca Gastropoda Caenogastropoda** | Onocmeleania hupensis | GenBank SRA | SRRX2735363 |
| **Mollusca Gastropoda Caenogastropoda** | Pila ampullacea | AmpuBase | Pila8267 |
| **Mollusca Gastropoda Caenogastropoda** | Pomacea canaliculata | GenBank proteins | XP_025109323 (incomplete) |
| **Mollusca Gastropoda Caenogastropoda** | Pomacea diffusa | AmpuBase | Pda16479 (partial) |
| **Mollusca Gastropoda Caenogastropoda** | Pomacea maculata | AmpuBase | Pma3939 (partial) |
| **Mollusca Gastropoda Caenogastropoda** | Pomacea scalaris | AmpuBase | Psc4690 |
| **Mollusca Gastropoda Caenogastropoda** | Rapania venosa | GenBank TSA | GDA01047641 |
| **Mollusca Gastropoda Caenogastropoda** | Semisulcospira coreana | GenBank TSA | GGNX01507307 |
| **Mollusca Gastropoda Vetigastropoda** | Haliothis laevigata | GenBank TSA | GFTT01038064 |
| **Mollusca Gastropoda Vetigastropoda** | Haliothis rubra | GenBank WGS | QXJH01001142 |
| **Mollusca Gastropoda Vetigastropoda** | Haliothis rubescens | GenBank WGS | QX0010103565 |
| **Mollusca Gastropoda Vetigastropoda** | Tegula atr | GenBank SRA | SRX958768 |
| **Mollusca Bivalvia Miyloloidi** | Bathymodiolus platirons | GenBank Assembly | MJU01033839 |
| **Mollusca Bivalvia Mytiloidi** | Limnopoma fortunei | GenBank Assembly | NFUK01006104 |
| **Mollusca Bivalvia Miyloloidi** | Lithophaga lithopha | GenBank SRA | SRX1940727 |
| **Mollusca Bivalvia Mytiloidi** | Modiolus philipinarum | GenBank Assembly | MIIII01021410 |
| **Mollusca Bivalvia Mytiloidi** | Mytilus gallogalprensis | GenBank Assembly | APJB01151127 |
| **Mollusca Bivalvia Mytiloidi** | Mytilus galloprovincialis | GenBank TSA | GHK0102503 |
| **Mollusca Bivalvia Mytiloidi** | Perina canaliculus | GenBank TSA | GGLA0150624 |
| **Mollusca Bivalvia Mytiloidi** | Septifer virgatus | GenBank TSA | GF010356711 |
| **Mollusca Bivalvia Mytiloidi** | Permutomythus purpuratus | GenBank SRA | SRX2210805 |
| **Mollusca Bivalvia Mytiloidi** | Xenostrobus secus | GenBank SRA | SRX4058936 |
| **Mollusca Bivalvia Pterygida** | Maluleus candeans | GenBank SRA | SRX1688295 |
| **Mollusca Bivalvia Pterygida** | Pinctada martensi | GenBank Assembly | CM008066 |
| **Mollusca Bivalvia Pterygida** | Pinctada fucata | Marinegenomics | pfu_aug1.0_41421.0_1638 |
| **Mollusca Bivalvia Pterygida** | Pteria penguin | GenBank TSA | GEM01011007 |
| **Mollusca Bivalvia Arcoida** | Anadara trapeza | GenBank SRA | SRX323049 |
| **Mollusca Bivalvia Arcoida** | Scapharca broughtoni | GenBank TSA | GE01004652 |
| **Mollusca Bivalvia Arcoida** | Tegillara granosa | GenBank SRA | SRX1334524 |
| **Mollusca Bivalvia Unionoidi** | Cristaria pilica | GenBank SRA | SRX1153631 |
| **Annelida Oligochaeta** | Drawia calebi | GenBank SRA | SRX5656293 |
| **Annelida Oligochaeta** | Glossoscoleus pustulata | GenBank TSA | GB1L01075477 |
| **Annelida Polychaeta** | Hydroies elegans | GenBank Assembly | LQRL0114159 |
| **Annelida Polychaeta** | Pygospio elegans | GenBank TSA | GFPL01035490 |
| **Annelida Polychaeta** | Spirobranchus lamarckii | GenBank TSA | GSGS01192599 |

**UNICELLULAR EUKARYOTES**

| Phylum | Species | Database | Accession |
|--------|---------|----------|-----------|
| **Amoebozoa Mycetoza** | Cavendoria fasciculata | GenBank proteins | XP_00435194 |
| **Amoebozoa Mycetoza** | Dictyostellium discoideum | GenBank proteins | XP_640516 |
| **Amoebozoa Mycetoza** | Polyphondylum pallidum | GenBank proteins | XP_020492468 |
| **Amoebozoa Discosea** | Acanthamoeba castellanii | GenBank proteins | XP_004368209 |
| **Choanoflagellida Salpingoecidae** | Monosiga brevicollis | GenBank proteins | XP_001742116 |
| **Choanoflagellida Salpingoecidae** | Salpingoeca rosetta | GenBank proteins | XP_00498636 |
| **Ciliata** | Ichthyophytilus multifiliis | GenBank proteins | XP_00427176 |

Downloaded from https://academic.oup.com/g3journal/article-10.1093/g3journal/nqaa127/6026274 by guest on 31 July 2021
In our general phylogenetic tree (Figure 1), these sequences were used as outgroups. In choanoflagellates, where *Monosiga brevicollis* was already known to harbor a GH13_1 sequence (Da Lage et al. 2007), we found a GH13_1 sequence in the genome of *Salpingoeca rosetta*. A partial sequence was also returned from incomplete genome data from *Monosiga ovata* (at Compagno, not shown).

**GH13_1 sequences retrieved from non-bilaterian animals**

In Cnidaria, a number of GH13_1 sequences were recovered from many Anthozoa species (sea anemones, corals and allies), from genome as well as transcriptome data, at the Reefgenomics database (Table S1). Interestingly, we found no alpha-amylase sequences at all in Medusozoa (jellyfishes, hydrosd) nor in Endocnidoidae (parasitic cnidarians). In the general tree (Figure 1), cnidian sequences form a clear cluster with two main branches, grouping Actiniaria (sea anemones) and Pennatulacea (soft corals) on one branch, and Scleractinia (hard corals) and Corallimorpharia (mushroom anemones) on the other branch.

In sponges (Porifera), data were less abundant. No alpha-amylase sequence was found in *Sycon ciliatum* (Calcarea) and *Oscarella carmela* (Homoscleromorpha). All the sequences we retrieved belonged to Demospomigae. Similarly, we found no amylase sequence at all in the phylum Ctenophora (*Mnemiopsis leidyi*, *Pleurobrachia bachei*), suggesting of this study, on which we will focus our attention, the divergence between paralogs is ancestral to these species, dating back at least to basal Cambrian, according to the TimeTree database (Kumar et al. 2017). GH13_1 sequences were found in other brachiopods and phoronids as sequence reads (SRA) from transcriptome data only, with no available genomic support (listed in Table 1 and S2). We must be cautious when only transcriptome data are available, as transcripts from contaminating symbionts or parasites may generate false positives (Borner and Burmester 2017) and/or the lack of expression of the targeted sequence in the investigated tissues may lead to false negatives. However, seven different brachiopod species returned positive hits, giving some robustness to our finding. Two phyla are related to Brachiopoda: Bryozoa and Nemertea (Kocot 2016; Luo et al. 2018, but see Marlétaz et al. 2019). We found clues for the presence of GH13_1 in four Bryozoan species, but only transcriptome reads were available. In contrast, in Nemertea, none of the 14 species investigated returned any GH13_1 sequence, including the annotated genome of *Natosperma genticulatus*.

Similarly, we found three gene copies in the genomes of the hemichordates *Saccoglossus kowalevskii* and *Pycthodera flavia*. In both species, two copies are close to each other (XP_006816581 and XP_006816582 in *S. kowalevskii*, and their counterparts in *P. flavia* as shown by the topology of the gene tree (Figure 1). This could suggest independent gene duplication in each species. However, we observed that the two duplicates were arranged in tandem in both species, which would rather suggest concerted evolution of two shared copies. In *P. flavia*, this genome region is erroneously annotated as a single gene at the OIST Marine Genomics database. The third paralog is very divergent from the two other copies, so its divergence from the ancestral copy probably occurred before the species split, as well. The three copies were therefore probably already present before the split of the two lineages, some 435 mya (Kumar et al. 2017). Three other hemichordate species, *Schizocardium californicum*, *Torquatoridaria antarctica* and *Rhabdopleura sp*., harbor a GH13_1 gene, as shown by SRA search in GenBank (Table 1). A positive result was also retrieved from the genome of *Glandiceps talaboti* (Héctor Escrivá, Oceanology Observatory at Banyuls-sur-mer, personal communication).

In molluscs, we found BLAST hits with significant e-values in gastropod species from two clades only, the Veticagnostopoda (*e.g.*, the abalone *Haliotis* sp.) and the Caenogastropoda (*e.g.*, Ampullariidae such as *Pomacea canaliculata*). We consistently found one copy in eight species belonging to the family Ampullariidae. In *P. canaliculata*, the genome of which has been well annotated, the GH13_1 sequence (XP_025109323) lies well inside a 26 Mbp long scaffold (linkage group 10, NC_037599) and is surrounded by *bona fide* mollusc genes (Table S3). GH13_1 sequences were found in other Caenogastropoda from SRA or transcriptome databases (Table 1 and S2). We also found GH13_1 sequences in several bivalve clades: *Mytiloida* (*e.g.*, the mussel *Mytilus galloprovincialis*), *Pterioida* (*e.g.*, the pearl oyster *Pinctada imbricata*), *Arcoidea* (*e.g.*, *Scapharca broughtii*).
Figure 1 ML tree of GH13_1 protein sequences of metazoan and non-metazoan species. The tree was rooted by placing fungi and unicellular organisms, except choanoflagellates, as outgroups. The numbers at the nodes are the aLRT supports. Dark green: hemichordates; light blue: brachiozoans; red: cnidarians, dark blue: sponges; orange: placozoans; pink: choanoflagellates; purple: amoebozoans; brown: fungi; gray: molluscs; bright green: annelids; black: other protists.
and in the Unionoida Cristaria plicata. For sequences retrieved from the TSA or SRA databases (see Table 1), whose issues were mentioned above, we performed reciprocal BLAST in GenBank nr. Almost always Lingula anatina was recovered as the best hit. However, as an example of the necessary careful examination of results, we found a significant HSP in a transcriptome database of the sea hare Aplysia californica (TSA GBDA01069300). This sequence was not found in the well annotated A. californica genome, and turned out to be related to clades. We found no occurrence of GH13_1 in Veneroida, Pectinoida and Ostreoida, for which annotated and/or assembled genomes exist, nor in cephalopods.

In annullids, we found occurrences of GH13_1 genes in a few species, the genomes of which are still not fully assembled, namely the “polychaetes” Hydroides elegans, Pygospio elegans and Spirobranchus lamarcki but not from Amynthas corticis or Eisenia fetida. We found no GH13_1 sequences in Hirudinea (leeches). To summarize, in molluscs as well as in annulids, the presence of GH13_1 genes is scattered and patchy across and within lineages. Interestingly, we found that some of the mollusc GH13_1-like sequences, especially in bivalves, were much shorter, either truncated at the C-terminal, or this region was so divergent from the query sequence (L. anatina) that it was impossible to identify, assemble and align it with our data set (Fig. S2). In addition, we found that the annelid Hydroides elegans had an internal deletion, which precluded its inclusion in the phylogenetic analysis. This suggests that those sequences may not have alpha-amylase activity.

**Gene tree analysis: position of bilaterian sequences**

The goal of the gene tree analysis is to examine whether the occurrence of GH13_1 genes in bilaterian animals may be due to independent horizontal gene transfers (HGT) or if they descend from a GH13_1 alpha-amylase copy ancestral to Unikonts. In the first case, the bilaterians GH13_1 sequences are unlikely to cluster together and the gene tree topology will likely display one or more nodes that are inconsistent with the bilaterian phylogeny. In the second case, the bilaterian sequences are expected to recover a bilaterian clade and to have a cnidarian tree topology will likely display one or more nodes that are inconsistent with the emergence of the “polychaetes” Hydroides elegans, Pygospio elegans and Spirobranchus lamarcki but not from Amynthas corticis or Eisenia fetida. We found no GH13_1 sequences in Hirudinea (leeches). To summarize, in molluscs as well as in annulids, the presence of GH13_1 genes is scattered and patchy across and within lineages. Interestingly, we found that some of the mollusc GH13_1-like sequences, especially in bivalves, were much shorter, either truncated at the C-terminal, or this region was so divergent from the query sequence (L. anatina) that it was impossible to identify, assemble and align it with our data set (Fig. S2). In addition, we found that the annelid Hydroides elegans had an internal deletion, which precluded its inclusion in the phylogenetic analysis. This suggests that those sequences may not have alpha-amylase activity.

**Analysis of intron positions**

Intron positions may be valuable markers when reconstituting gene histories. We identified 56 intron positions from the subset of species of the general tree for which we could find data (Figure 2). Only one intron position is widely shared among these GH13_1 gene sequences. It is the first position reported in the alignment, and it lies just upstream to the first conserved part of the alignment. The main observation is the numerous conserved positions across bilaterian sequences (10 positions), and between bilaterian sequences and the sponge and the Placozoa (7 positions). In addition, three positions are conserved with Cnidaria, as phylogenetically expected in the case of a shared ancestral gene, as a robust cluster grouping one Brachiozoa (brachiopod/phoronid) copy, the molluscs and the annulids, which is consistent with the phylogeny. However, the tandem hemichordate duplicates and the other Brachiozoa genes are not included in the bilaterian clade, but remain ingroup relative to the sponge sequences.

Interestingly, the two remaining hemichordate sequences are the earliest diverging lineage of the Metazoa + Choanoflagellata cluster, since they are branched with the placozoan Trichoplax adhaerens sequence, this relationship being strongly supported whatever the tree reconstruction method employed (Figure 1, and data not shown). In order to check for the possibility of a long branch attraction (LBA), which would artificially cluster hemichordate and placozoan sequences, we performed Tajima’s relative rate tests (Tajima 1993) using MEGA7 (Kumar et al. 2016). The sequence of S. kowalevskii XP_006819810, suspected to evolve fast, was compared with its paralog XP_006816581, using five different outgroups, i.e., the three sponges and the two choanoflagellates. Unexpectedly, the χ² tests returned non-significant values in two tests and significant values in three tests (Table S4). Therefore, with our data, LBA cannot be entirely ruled out in this particular case.

**DISCUSSION**

The evolutionary scenario proposed by Da Lage et al. 2007, suggested that the GH13_1 alpha-amylase gene ancestral to Unikonts (Amoebozoa and Opisthokonts, i.e., Fungi and Metazoa/Choanoflagellata) was totally absent from Bilateria, due to its complete replacement by a new alpha-amylase, originating from a bacterium through HGT. Here, we have shown that a limited number of bilaterian lineages, all aquatic species, namely hemichordates, brachiozoans, bryozoans, and some sparse molluscs and annulids, actually do harbor GH13_1 alpha-amylase genes. Note that all those species also have at least one

ML tree scores was 53.01. According to Kass and Raftery (1995), there is very strong support for the highest likelihood hypothesis (in our case, the ML tree in Figure 1) when the double of this difference (i.e., 2 × 53.01 = 106.02) exceeds 10 log-likelihood units. Thus, there is significant inconsistency between the position of R. reniformis’ GH13_1 copy and the phylogenetic placement of this species. This may be due to a horizontal transfer event that would have occurred within Cnidaria, but additional data from well-sequetted Pennatulacea would be welcome to check this possibility. Nevertheless, it is noteworthy that the genome of Dendronephthya gigantea (Octocorallia, order Alcyonacea) returned no result. Most bilaterian sequences are clustered with Cnidaria, as phylogenetically expected in the case of a shared ancestral gene, as a robust cluster grouping one Brachiozoa (brachiopod/phoronid) copy, the molluscs and the annulids, which is consistent with the phylogeny. However, the tandem hemichordate duplicates and the other Brachiozoa genes are not included in the bilaterian clade, but remain ingroup relative to the sponge sequences.

Interestingly, the two remaining hemichordate sequences are the earliest diverging lineage of the Metazoa + Choanoflagellata cluster, since they are branched with the placozoan Trichoplax adhaerens sequence, this relationship being strongly supported whatever the tree reconstruction method employed (Figure 1, and data not shown). In order to check for the possibility of a long branch attraction (LBA), which would artificially cluster hemichordate and placozoan sequences, we performed Tajima’s relative rate tests (Tajima 1993) using MEGA7 (Kumar et al. 2016). The sequence of S. kowalevskii XP_006819810, suspected to evolve fast, was compared with its paralog XP_006816581, using five different outgroups, i.e., the three sponges and the two choanoflagellates. Unexpectedly, the χ² tests returned non-significant values in two tests and significant values in three tests (Table S4). Therefore, with our data, LBA cannot be entirely ruled out in this particular case.

**Analysis of intron positions**

Intron positions may be valuable markers when reconstituting gene histories. We identified 56 intron positions from the subset of species of the general tree for which we could find data (Figure 2). Only one intron position is widely shared among these GH13_1 gene sequences. It is the first position reported in the alignment, and it lies just upstream to the first conserved part of the alignment. The main observation is the numerous conserved positions across bilaterian sequences (10 positions), and between bilaterian sequences and the sponge and the Placozoa (7 positions). In addition, three positions are conserved with Cnidaria, as phylogenetically expected in the case of a shared ancestral gene, as a robust cluster grouping one Brachiozoa (brachiopod/phoronid) copy, the molluscs and the annulids, which is consistent with the phylogeny. However, the tandem hemichordate duplicates and the other Brachiozoa genes are not included in the bilaterian clade, but remain ingroup relative to the sponge sequences.

Interestingly, the two remaining hemichordate sequences are the earliest diverging lineage of the Metazoa + Choanoflagellata cluster, since they are branched with the placozoan Trichoplax adhaerens sequence, this relationship being strongly supported whatever the tree reconstruction method employed (Figure 1, and data not shown). In order to check for the possibility of a long branch attraction (LBA), which would artificially cluster hemichordate and placozoan sequences, we performed Tajima’s relative rate tests (Tajima 1993) using MEGA7 (Kumar et al. 2016). The sequence of S. kowalevskii XP_006819810, suspected to evolve fast, was compared with its paralog XP_006816581, using five different outgroups, i.e., the three sponges and the two choanoflagellates. Unexpectedly, the χ² tests returned non-significant values in two tests and significant values in three tests (Table S4). Therefore, with our data, LBA cannot be entirely ruled out in this particular case.

**DISCUSSION**

The evolutionary scenario proposed by Da Lage et al. 2007, suggested that the GH13_1 alpha-amylase gene ancestral to Unikonts (Amoebozoa and Opisthokonts, i.e., Fungi and Metazoa/Choanoflagellata) was totally absent from Bilateria, due to its complete replacement by a new alpha-amylase, originating from a bacterium through HGT. Here, we have shown that a limited number of bilaterian lineages, all aquatic species, namely hemichordates, brachiozoans, bryozoans, and some sparse molluscs and annulids, actually do harbor GH13_1 alpha-amylase genes. Note that all those species also have at least one
metazoans. It implies that HGTs obviously happened after the split of Annelida. Brachiozoa, Bryozoa, and in scatter end lineages across Mollusca and versi case of Lophotrochozoa, this would have occurred before the di-most phyla, like in the alternative hypothesis. More precisely, in the stomes, otherwise the transferred copies should have been lost in lineages, but remained (given the current data) in hemichordates, tion is that the ancestral GH13_1 gene was not lost in all bilaterian ex-

The hypothesis of HGT requires several such events between metazoans. It implies that HGTs obviously happened after the split of the two main branches of bilaterians, protostomes and deute-

Figure 2 Intron positions compared across the sampled GH13_1 genes. The intron positions found in the studied parts of the sequences were numbered from 1 to 56. Pink: phase zero introns; green: phase 1 introns; blue: phase 2 introns. The black horizontal bar separates bilaterians from species where GH13_1 alpha-amylases are considered native. The color code for species is the same as in Figure 1.

classical animal alpha-amylase of the GH13_15/24 subfamilies. Several species with whole genome well sequenced and annotated were found to harbor such genes in each phylum Hemichordata, Brachiozoa and molluscs. They were investigated in more details, especially regarding the genomic environment of their GH13_1 genes. We are quite confident that the GH13_1 sequences we found are not due to contaminating DNA. First, the bilaterian sequences retrieved from annotated genomes were inside long contigs, and mostly surrounded by genes showing bilaterian best BLAST hits (Table S3). However, the S. kowalevskii XP_006819810 gene could appear somewhat dubious, since it is placed at the distal end of a contig, with only two other genes on the contig (Table S3), one of which has a placozoan best hit. But its P. flava counterpart is well inside a gene-rich contig. Therefore, these seemingly non-bilaterian genes are well in bilaterian genomic contexts. Second, a lot of additional sequences from other species belonging to these phyla were gathered from more sketchy data, i.e., lower-quality assembled genomes, transcriptome or sequence read archive databases, which added some support to the presence of these amylase genes. Although transcriptome and rough genomic data should be handled with care, this lends support to our observations. Moreover, reciprocal BLAST from the transcriptome hits always returned a bilaterian (L. anatina or S. kowalevskii) best hit, not fungal, protist or other non-bilaterian GH13_1 sequence.

The new data unveils an evolutionary story more complicated than previously supposed. There are two alternative explanations. The first explanation is that several HGTs occurred from non-bilaterian to both hemichordate and Lophotrochozoa ancestors. The second explanation is that the ancestral GH13_1 gene was not lost in all bilaterian lineages, but remained (given the current data) in hemichordates, Brachiozoa, Bryozoa, and in scattered lineages across Mollusca and Annelida.

The hypothesis of HGT requires several such events between metazoans. It implies that HGTs obviously happened after the split of the two main branches of bilaterians, protostomes and deuterostomes, otherwise the transferred copies should have been lost in most phyla, like in the alternative hypothesis. More precisely, in the case of Lophotrochozoa, this would have occurred during the diversification of this clade and after its divergence from the Platyzoa, some 700 mya (Kumar et al. 2017); in the case of hemichordates, after diverging from their common ancestor with the echinoderms, and before the divergence between S. kowalevskii and Ptychodera flava, i.e., between 657 and ca. 435 mya (Kumar et al. 2017). Therefore, we may infer at least two HGTs, each early in the evolution of the phyla, with a number of subsequent losses in Lophotrochozoa (Figure 3A). The donor species, given the sequence clustering in the trees, could be related to cnidarians. However, we have underlined that the intron-exon structures of the bilaterian sequences were most similar to the one of the sponge, and that the cnidarian GH13_1 amylases had very different structures. This may be possible if the donors were related to cnidarians, perhaps an extinct phylum or an ancestor of extant Cnidaria, but had conserved the ancestral structures exemplified by the sponge and the placozoan. Indeed, if the structure shared by the sponge, the placozoan and the bilaterians reflects the ancestral state, cnidarians must have undergone a drastic rearrangement of the intron-exon structure of this gene. This would be in line with the long internal branch leading to this clade in the trees (Figure 1), which suggests accelerated evolution.

The alternative hypothesis of massive GH13_1 gene loss in most phyla except the ones where we found such sequences seems no more parsimonious. It requires many losses, the number of which depends on the phylogeny used. For instance, considering the phylogeny shown in Figure 3B, regarding deuterostomes, one loss occurred in echinoderms and another one in chordates. In protostomes, one GH13_1 loss in edyszoa, and independent losses in Platyza and in several lophotrochozoan lineages would be required to produce the observed pattern.

However, although not parsimonious in terms of number of events, we would rather favor the gene loss hypothesis, because this is a common phenomenon, especially given how ubiquitous co-option is (Flores and Livingstone 2017; Hejnol and Martindale 2008). In this respect, the GH13_15/24 gene that was acquired from a bacterium is a type of horizontal transfer akin to what Husnik and McCutcheon called a “maintenance transfer” since it allowed the original function to be maintained while the primitive GH13_1 gene became free to evolve or even to be lost (Husnik and McCutcheon 2018) (see also Da Lage et al. 2013). In contrast, while numerous cases of HGT from bacteria to metazoans, or from fungi to metazoans have been reported (e.g., Wybouw et al. 2016; Dunning Hotopp 2011, 2018; Haegeman et al. 2011; Crisp et al. 2015; Cordaux and Gilbert 2017), very few HGT events have been inferred that involve a metazoan donor and a meta-

Figure 3A Intron positions found in the sampled GH13_1 genes. The intron positions found in the studied parts of the sequences were numbered from 1 to 56. Pink: phase zero introns; green: phase 1 introns; blue: phase 2 introns. The black horizontal bar separates bilaterians from species where GH13_1 alpha-amylases are considered native. The color code for species is the same as in Figure 1.
families. For instance, HMM search, such as in PFAM (pfam.xfam.org), shows that the domain composition of e.g., the Lingula anatina sequence XP_013396432 consists in an alpha-amylase domain linked to a DUFF1966 domain (DUFF1266 is also present in several fungal proteins, including obviously the GH13_1 amylase). The alpha-amylase domain is actually present in many glycosyl hydrolase families. Interestingly, the sequences found in some molluscs do not have a complete alpha-amylase domain, because they are shorter than usual (see Results). We assumed nonetheless that all the sequences we recovered belong to the GH13_1 subfamily, due to sequence similarities, as shown by the easy sequence alignment. Further, some of them have been assigned to this subfamily in the reference database CAZy.org (see Table 1). In addition, if we add sequences from the closest subfamilies, namely GH13_2 or GH13_19 (Stam et al. 2006) in the alignment and in the phylogenetic tree, the putative GH13_1 and the ascertained GH13_1 remain well clustered together (not shown). It is possible that modifications of a few amino acid positions could bring a change in the substrate or catalytic activity. For instance, concerning the substrate affinity, when the genome of L. anatina was released, the authors hypothesized a biminerolization pathway that involves acid proteins, as found in scleractinians and molluscs (Marin et al. 2007; Ramos-Silva et al. 2013). Given the calcium binding activity of alpha-amylases (Boel et al. 1990; Grossman and James 1993; Svensson 1994; Pujadas and Palau 2001), the presence of both GH13_1 and GH13_15/24 subfamilies in L. anatina opens the possibility for the neofunctionalization of one of them in the biminerolization process. In the analyses performed by those authors, no amylase was found in the shell matrix, but this does not exclude the possibility of its presence in the pathway. Moreover, the fact that in some molluscs, the sequences are incomplete compared to the brachiopod query or to the sponge and cnidarian GH13_1 amylases, and therefore probably devoid of an amylolytic function, would add credence to another function, especially considering that they are transcribed. This conjecture requires further investigation. On the other hand, the full-size GH13_1 sequences only present in a few bilaterians could have remained true alpha-amylases with the classical function, but this would make even more enigmatic why they have been conserved, either by descent or by horizontal transfer.

ACKNOWLEDGMENTS

We want to thank Pedro E. Vieira for introducing AD to molecular analyses. We are grateful to Didier Casane and Emmanuelle Renard for fruitful advise and discussion and two anonymous reviewers for critical reading of the manuscript. We thank Héctor Escrivá for sharing sequence data. This work was funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (process no. 141565/2017-9) to AD and regular funding of the CNRS to JLDL and CG. CG was also supported by a grant from Agence Nationale de la Recherche (ANR-15-CE32-0011-01 TransVir).

LITERATURE CITED

Anisimova, M., and O. Gascuel, 2006 Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst. Biol. 55: 539–552. https://doi.org/10.1080/10635150600755453
Barbeitos, M. S., S. L. Romano, and H. R. Lasker, 2010 Repeated loss of acitvity, when the genome of L. anatina was released, the authors hypothesis for the neofunctionalization of one of them in the biminerolization process. In the analyses performed by those authors, no amylase was found in the shell matrix, but this does not exclude the possibility of its presence in the pathway. Moreover, the fact that in some molluscs, the sequences are incomplete compared to the brachiopod query or to the sponge and cnidarian GH13_1 amylases, and therefore probably devoid of an amylolytic function, would add credence to another function, especially considering that they are transcribed. This conjecture requires further investigation. On the other hand, the full-size GH13_1 sequences only present in a few bilaterians could have remained true alpha-amylases with the classical function, but this would make even more enigmatic why they have been conserved, either by descent or by horizontal transfer.

ACKNOWLEDGMENTS

We want to thank Pedro E. Vieira for introducing AD to molecular analyses. We are grateful to Didier Casane and Emmanuelle Renard for fruitful advise and discussion and two anonymous reviewers for critical reading of the manuscript. We thank Héctor Escrivá for sharing sequence data. This work was funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (process no. 141565/2017-9) to AD and regular funding of the CNRS to JLDL and CG. CG was also supported by a grant from Agence Nationale de la Recherche (ANR-15-CE32-0011-01 TransVir).

LITERATURE CITED

Anisimova, M., and O. Gascuel, 2006 Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst. Biol. 55: 539–552. https://doi.org/10.1080/10635150600755453
Barbeitos, M. S., S. L. Romano, and H. R. Lasker, 2010 Repeated loss of calcium-binding in alpha-amylases: an X-ray diffraction study at 2.1 Å resolution of two enzymes from Aspergillus. Biochemistry 29: 6244–6249. https://doi.org/10.1021/bi00478a019
Borner, J., and T. Burmester, 2017 Parasite infection of public databases: a data mining approach to identify apicomplexan contaminations in

G3 Genomes | Genomics | Genetics

Volume 10 February 2020 I Origin of Animal Amylases I 717

Downloaded from https://academic.oup.com/g3journal/article/10/2/709/6026274 by guest on 31 July 2021
Pujadas, G., and J. Palau, 2001 Evolution of α-amylase: architectural features and key residues in the stabilization of the (β/α)8 scaffold. Mol. Biol. Evol. 18: 38–54. https://doi.org/10.1093/oxfordjournals.molbev.a003718

Ramos-Silva, P., J. Kaandorp, L. Huisman, B. Marie, I. Zanella-Cléon et al., 2013 The skeletal proteome of the coral Acropora millepora: the evolution of calcification by co-option and domain shuffling. Mol. Biol. Evol. 30: 2099–2112. https://doi.org/10.1093/molbev/mst109

Rodríguez, E., M. S. Barbeitos, M. R. Brugler, L. M. Crowley, A. Grajales et al., 2014 Hidden among Sea Anemones: The First Comprehensive Phylogenetic Reconstruction of the Order Actiniaria (Cnidaria, Anthozoa, Hexacorallia) Reveals a Novel Group of Hexacorals. PLoS One 9: e96998. https://doi.org/10.1371/journal.pone.0096998

Simion, P., H. Philippe, D. Baurain, M. Jager, D. J. Richter et al., 2017 A large and consistent phylogeny dataset supports sponges as the sister group to all other animals. Curr. Biol. 27: 958–967. https://doi.org/10.1016/j.cub.2017.02.031

Srivastava, M., E. Begovic, J. Chapman, N. H. Putnam, U. Hellsten et al., 2008 The Trichoplax genome and the nature of placozoans. Nature 454: 955–960. https://doi.org/10.1038/nature07191

Srivastava, M., O. Simakov, J. Chapman, B. Fahey, M. E. A. Gauthier et al., 2010 The Amphimedon queenslandica genome and the evolution of animal complexity. Nature 466: 720–726. https://doi.org/10.1038/nature09201

Stam, M. R., E. G. J. Danchin, C. Rancurel, P. M. Coutinho, and B. Henrissat, 2006 Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase-related proteins. Protein Eng. Des. Sel. 19: 555–562. https://doi.org/10.1093/protein/gzl044

Sullivan, J. C., A. M. Reitzel, and J. R. Finnerty, 2006 A high percentage of introns in human genes were present early in animal evolution: evidence from the basal metazoan Nematostella vectensis. Genome Informatics 17: 219–229.

Svensson, B., 1994 Protein engineering in the alpha-amylase family: catalytic mechanism, substrate specificity, and stability. Plant Mol. Biol. 25: 141–157. https://doi.org/10.1007/BF00023233

Tajima, F., 1993 Simple methods for testing the molecular evolutionary clock hypothesis. Genetics 135: 599–607.

Uribe, J., Y. Kano, J. Templado, and R. Zardoya, 2016 Mitogenomics of Vetigastropoda: insights into the evolution of pallial symmetry. Zool. Scr. 45: 145–159. https://doi.org/10.1111/zsc.12146

van der Kaaij, R. M., S. Janecek, M. J. E. C. van der Maarel, and L. Dijkhuizen, 2007 Phylogenetic and biochemical characterization of a novel cluster of intracellular fungal alpha-amylase enzymes. Microbiology 153: 4003–4015. https://doi.org/10.1099/mic.0.2007/008697-0

Whelan, N. V., K. M. Kocot, T. P. Moroz, K. Mukherjee, P. Williams et al., 2017 Ctenophore relationships and their placement as the sister group to all other animals. Nat. Ecol. Evol. 1: 1737–1746. Erratum: 1783. https://doi.org/10.1038/s41559-017-0331-3

Whelan, S., and N. Goldman, 2001 A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol. Biol. Evol. 18: 691–699. https://doi.org/10.1093/oxfordjournals.molbev.a003851

Wybouw, N., Y. Pauchet, D. G. Heckel, and T. Van Leuuwen, 2016 Horizontal gene transfer contributes to the evolution of arthropod herbivory. Genome Biol. Evol. 8: 3594–3613.

Zapata, F., F. E. Goetz, S. A. Smith, M. Howison, S. Siebert et al., 2015 Phylogenomic Analyses Support Traditional Relationships within Cnidaria. PLoS One 10: e0139068. https://doi.org/10.1371/journal.pone.0139068

Communicating editor: H. Tachida