

# The genome of the contractile demosponge *Tethya wilhelma* and the evolution of metazoan neural signalling pathways

Warren R. Francis<sup>1</sup>, Michael Eitel<sup>1</sup>, Sergio Vargas<sup>1</sup>, Marcin Adamski<sup>2</sup>, Steven H.D. Haddock<sup>3</sup>, Stefan Krebs<sup>4</sup>, Helmut Blum<sup>4</sup>, Dirk Erpenbeck<sup>1,5</sup>, Gert Wörheide<sup>1,5,6</sup>

<sup>1</sup>Department of Earth and Environmental Sciences, Paleontology and Geobiology, Ludwig-Maximilians-Universität München

Richard-Wagner Straße 10, 80333 Munich, Germany

<sup>2</sup>Research School of Biology, College of Medicine, Biology & Environment, Australian National University, Canberra ACT 0200 Australia

<sup>3</sup>Monterey Bay Aquarium Research Institute, Moss Landing, CA 95039, USA

<sup>4</sup>Laboratory for Functional Genome Analysis (LAFUGA), Gene Center, Ludwig-Maximilians-University Munich, Munich, Germany

<sup>5</sup>GeoBio-Center, Ludwig-Maximilians-Universität München, Munich, Germany

<sup>6</sup>Bavarian State Collection for Paleontology and Geology, Munich, Germany

## Abstract

Porifera are a diverse animal phylum with species performing important ecological roles in aquatic ecosystems, and have become models for multicellularity and early-animal evolution. Demosponges form the largest class in sponges, but previous studies have relied on the only draft demosponge genome of *Amphimedon queenslandica*. Here we present the 125-megabase draft genome of a contractile laboratory demosponge *Tethya wilhelma*, sequenced to almost 150x coverage. We explore the genetic repertoire of transporters, receptors, and neurotransmitter metabolism across early-branching metazoans in the context of the evolution of these gene families. Presence of many genes is highly variable across animal groups, with many gene family expansions and losses. Three sponge classes show lineage-specific expansions of GABA-B receptors, far exceeding the gene number in vertebrates, while ctenophores appear to have secondarily lost most genes in the GABA pathway. Both GABA and glutamate receptors show lineage-specific domain rearrangements, making it difficult to trace the evolution of these gene families. Gene sets in the examined taxa suggest that nervous systems evolved independently at least twice and either changed function or were lost in sponges. Changes in gene content are consistent with the view that ctenophores and sponges are the earliest-branching metazoan lineages and provide additional support for the proposed clade of Placozoa/Cnidaria/Bilateria.

## Introduction

The presence of neurons is a defining character of animals, and is symbolic of their alleged superiority over all other life on earth. Nonetheless, the four non-bilaterian phyla, Porifera, Placozoa, Ctenophora and Cnidaria, are most different from other animals in their sensory systems and are often mistakenly referred to as “lower” animals in common parlance, despite the fact that, like bilaterians, non-bilaterians exist at the tips of the tree of life. Indeed, animals such as corals and sponges appear immobile or often unresponsive, challenging early theorists in their ideas of what is and is not an animal. Yet we now know that representatives from all four non-bilaterian phyla demonstrate dynamic responses to outside stimuli.

28 Neural evolution has been discussed previously in the context of paleontology (reviewed in [Wray et al.,  
29 2015]) and metazoan phylogeny (reviewed in [Jékely et al., 2015]). Indeed, it has been suggested that many  
30 features of bilaterian neurons and nervous systems represent separate, parallel evolutionary events from a  
31 “simple” nervous system. A simple nervous system then must arise from proto-neurons [Schierwater et al.,  
32 2009], however it is unclear what that might look like.

33  
34 Several qualities can be used to define neurons or proto-neurons [Leys, 2015, Nickel, 2010] such as synapses,  
35 electrical excitability, membrane potential, or secretory functions, though no single quality (and ultimately  
36 gene set) solely defines such cells as neurons. Two non-bilaterian groups, ctenophores and cnidarians, are  
37 thought to have true neurons. When considering the remaining two non-bilaterian phyla, sponges and  
38 placozoans, many components of neural cells are found without any neuron-like cells having been identi-  
39 fied [Srivastava et al., 2010, Riesgo et al., 2014a, Leys, 2015], although synapse-like structures have been  
40 identified in placozoan fiber cells that show vesicles close to an osmophile contact [Grell and Benwitz, 1974].

41  
42 Comparative analyses revealed a gradient of neural-like qualities indicating that “neuron-or-not” classi-  
43 fications are not straightforward. While ctenophores, cnidarians, and bilaterians have true neurons, struc-  
44 tural and biochemical differences, [Moroz et al., 2014, Moroz, 2015] led to the proposition that neurons in  
45 ctenophores and cnidarians may not be homologous, but rather separate evolutionary outcomes from neural-  
46 like precursor cells. Potentially, in the case of independent evolutions, neurons are “easy” to evolve, since it  
47 involves co-expression of various pan-metazoan genetic modules in the same cell type. Alternatively, early  
48 rudimentary signaling systems may have been energetically costly and not especially useful in pre-Cambrian  
49 oceans, and in such cases, it may have been comparatively easy to lose such genes and with them neuronal-  
50 type cells.

51  
52 Interpretation of neural evolution requires an accurate metazoan phylogeny, and the phylogenetic relation-  
53 ships of early-branching metazoans have been a topic of continued controversy. Some analyses support the  
54 traditional phylogenetic position of sponges as sister group to all other metazoans (“Porifera-sister”) [Philippe  
55 et al., 2009, Pick et al., 2010, Nosenko et al., 2013, Pisani et al., 2015, Simion et al., 2017] while others suggest  
56 that Ctenophora are the sister group to all other animals (“Ctenophora-sister”) [Dunn et al., 2008, Ryan  
57 et al., 2013, Whelan et al., 2015], and some analyses also recover the classical view, a Coelenterata clade  
58 uniting Cnidaria and Ctenophora [Philippe et al., 2009, Simion et al., 2017]. Importantly, phylogenomic  
59 analyses can be prone to systematic artifacts under some circumstances, depending on taxon sampling [Pick  
60 et al., 2010, Philippe et al., 2011], gene set [Nosenko et al., 2013], phylogenetic model [Pisani et al., 2015], or  
61 use of nucleotides instead of proteins [Jarvis et al., 2014]. Other methods based on presence or absence of  
62 the genes themselves have been proposed to provide a sequence-independent inference of phylogeny [Ryan  
63 et al., 2010, Ryan et al., 2013, Pisani et al., 2015], relying on the assumption that gene loss is a rare event.  
64 However, non-bilaterians have the additional problem that basic knowledge of many aspects of their biology  
65 is absent [Dunn et al., 2015], and so the biological context that may separate or unite groups is limited.

66  
67 In the context of phylogeny, the branching order critically affects whether neurons evolved multiple times  
68 or were lost (see schematic in Figure 1). Given the gradient of neural-like qualities, the actual evolutionary  
69 scenario may be somewhere in between a simple gain-loss of neurons. While some previous studies have  
70 focused on neural evolution in ctenophores [Ryan et al., 2013, Alberstein et al., 2015, Li et al., 2015] or  
71 analysing the genomic data from *A. queenslandica* [Krishnan et al., 2014], these alone do not provide a  
72 comprehensive picture of all animals.

73  
74 Here we have sequenced the genome of the contractile laboratory demosponge *Tethya wilhelma* [Sara  
75 et al., 2001] and examined the protein repertoire in the context of genes mediating the contraction, and  
76 other neural-like functions. Many metabolic genes show unique expansions in different sponge clades, as well  
77 as other phyla, making it challenging to clearly assign functions based on similarity to human proteins. We  
78 consider these expansions in the context of phylogenetics, showing that even though sponges lack neurons,  
79 signaling pathways have still expanded. This gives support to the hypothesis that early neural-like cells have  
80 become neurons multiple times in the history of animals.

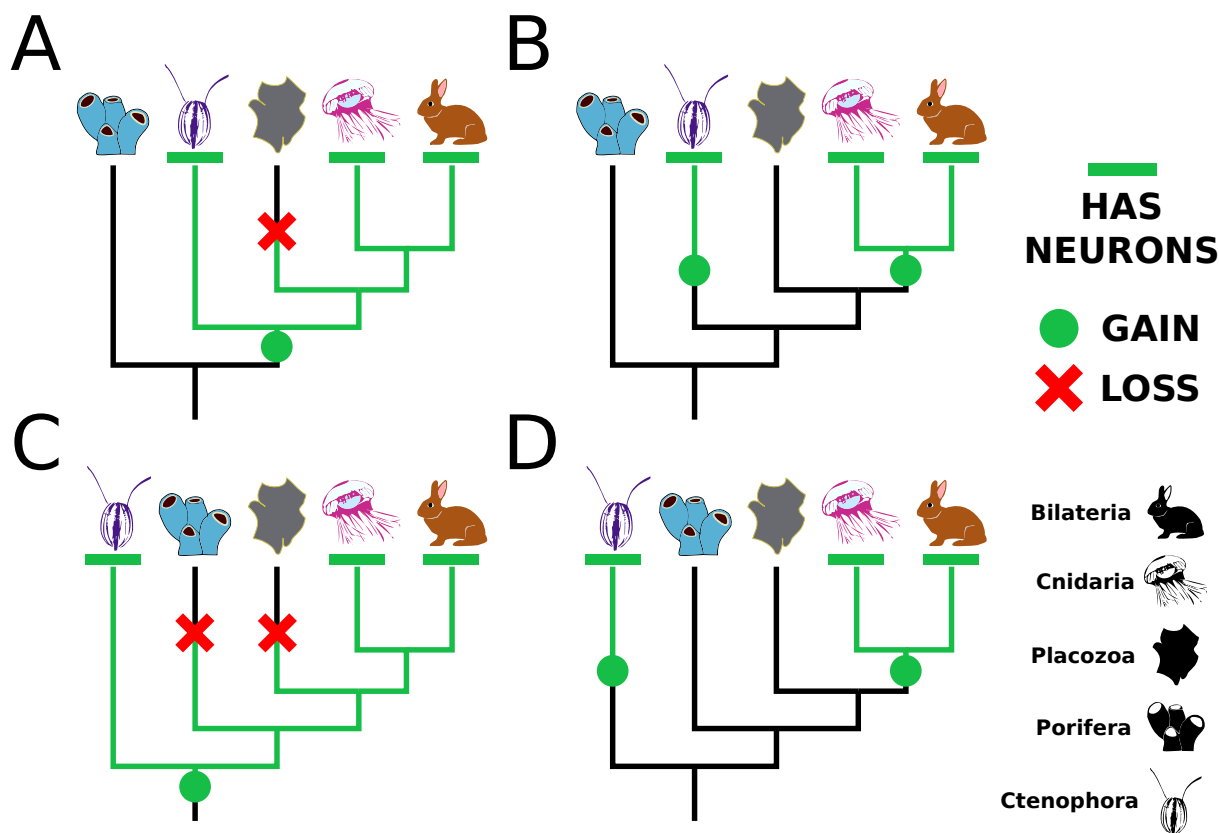


Figure 1: **Schematic of neural evolution depending on metazoan phylogeny** The presence of neurons or neural-like cells in ctenophores, cnidarians, and bilaterians can be viewed differently depending on the phylogeny. Two different metazoan phylogenies based on recent multi-gene phylogenetic analyses are the source of the Porifera-sister (A,B) [Philippe et al., 2009, Pisani et al., 2015, Simion et al., 2017] and Ctenophora-sister (C,D) [Ryan et al., 2013, Whelan et al., 2015] scenarios. Neurons can either have evolved once requiring a secondary loss in sponges, placozoans, or both (A,C), or evolved twice, in ctenophores and in cnidarians/bilaterians (B,D).

## 81 Results

### 82 Genome assembly and annotation

83 We generated a total of 61 gigabases of paired-end reads from a whole specimen of *T. wilhelma* (Figure 2)  
 84 and all associated bacteria. Because of a close association with microbes, some contigs were expected to  
 85 have derived from bacteria, as many reads have unexpectedly high GC content (Supplemental Figures 1-4).  
 86 After assembly and filtration of bacterial contigs, the final assembly was 125Mb, similar to *A. queenslandica*,  
 87 with a N50 value of 70kb (Supplemental Table 1). Gene annotation was done with a combination of a  
 88 deeply-sequenced RNAseq library from an adult sponge and *ab initio* gene predictions. Because of high  
 89 density of genes, extensive manual curation was often necessary to correct genes of the same strand that  
 90 were erroneously merged. After correction and filtering of the *ab initio* predictions, we counted 37,416  
 91 predicted genes, comparable with the counts in *A. queenslandica* (40,122) [Fernandez-Valverde et al., 2015]  
 92 and *S. ciliatum* (40,504) [Fortunato et al., 2014].

93 General trends in splice variation were similar between *T. wilhelma* and *A. queenslandica* (Supplemental  
 94 Tables 2 and 3), suggesting similar underlying biology or genome structure. One-to-one orthologs from *T.*  
 95 *wilhelma* and *A. queenslandica* had relatively low identity (Supplemental Figure 5), with the average identity  
 96 of 57.8%, showing a high genetic diversity within Porifera. The average identity is lower when compared to

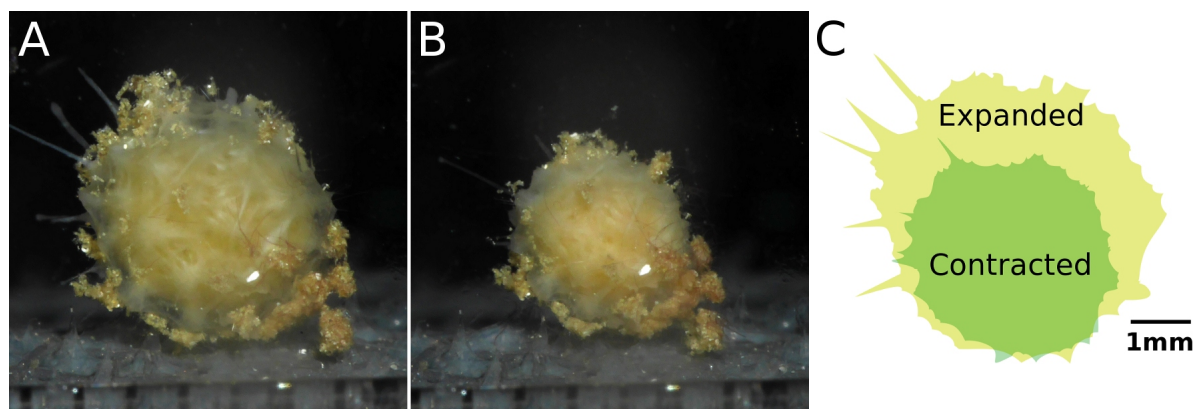


Figure 2: **Contraction of a normal specimen** (A) Maximally expanded state of *Tethya wilhelma*. (B) The same specimen approximately an hour later in the most contracted state. (C) Cartoon view of most contracted versus expanded state. Scale bar applies to all images. Photos courtesy of Dan B. Mills.

97 *S. ciliatum* (49.7%), *N. vectensis* (53.5%) and human (52.0%), which is not surprising given that *A. queens-*  
98 *landica* and *T. wilhelma* are both demosponges. Although both genomes are too fragmented to find syntenic  
99 chromosomal regions, ordered blocks of genes are still identifiable between *T. wilhelma* and *A. queenslandica*  
100 (Supplemental Figure 6), though not with *S. ciliatum*.

101

## 102 Neurotransmitter metabolism across early-branching metazoans

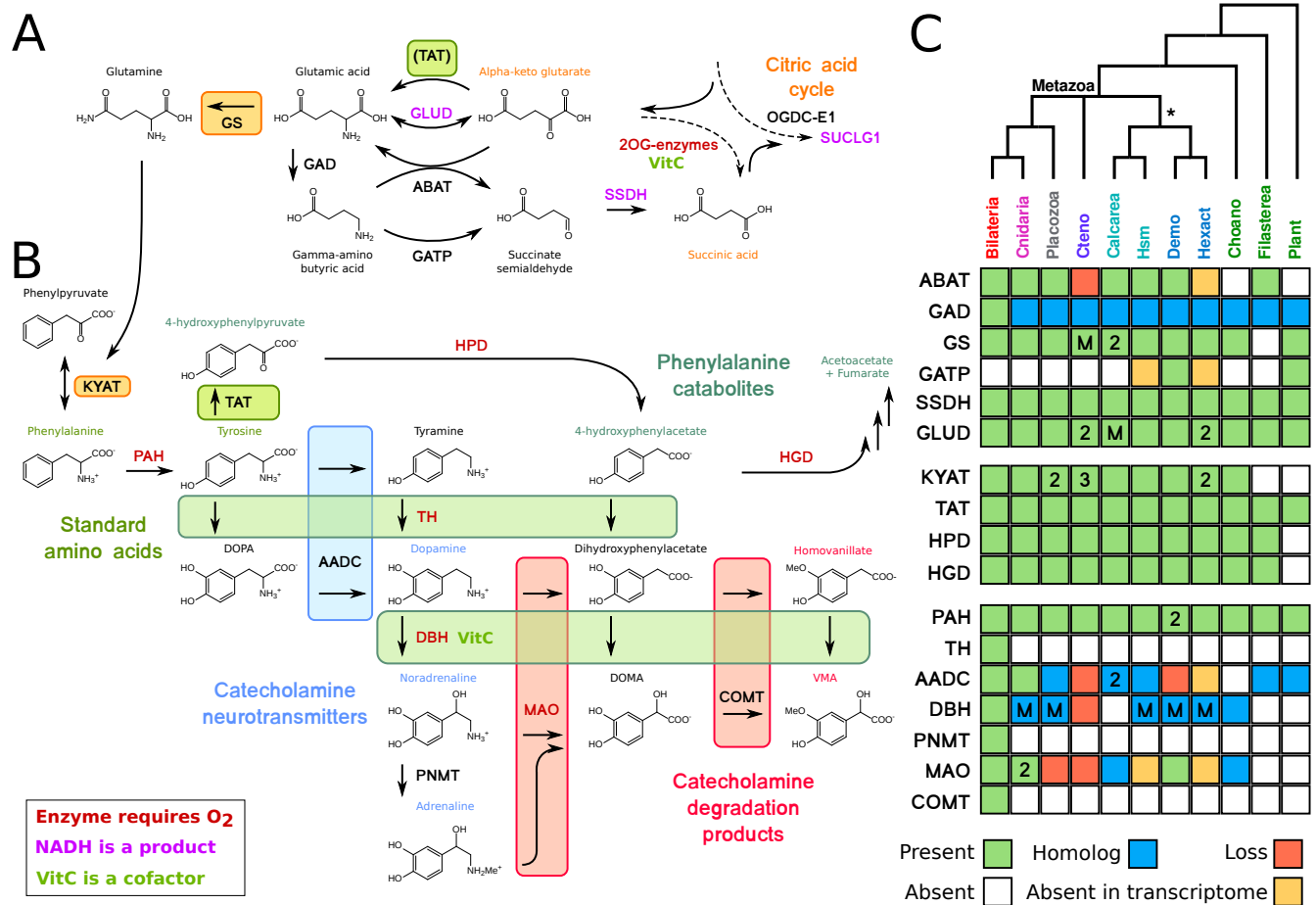
103 Compared to most other metazoans, sponges have a limited set of behaviors (contraction, closure of osculum  
104 or choanocyte chambers to control flow), yet respond to many signaling molecules present in bilaterians [Ell-  
105 wanger and Nickel, 2006, Ellwanger et al., 2007]. Some genes involved in vertebrate-like neurotransmitter  
106 metabolism have been found in sponges [Riesgo et al., 2014a, Krishnan and Schiöth, 2015], although many  
107 display a sister-group relationship to homologs found in other animals and appear to have a complex evo-  
108 lutionary history with duplications in sponges and other non-bilaterian animals (Figure 3, Supplemental  
109 Figures 7-9), making the prediction of their functions difficult. Implicitly, presence of a gene usually means  
110 a one-to-one orthology relationship with a functionally annotated protein, probably a human protein. Since  
111 many of the non-bilaterian proteins in our set are many-to-many orthologs to human proteins with known  
112 functions, declaring presence or absence of any individual gene or genetic module is not correct in the strictest  
113 sense, as one-to-many or many-to-many orthologs are not the same gene. In such cases, it is not currently  
114 possible to computationally predict which, if any, of the sponge orthologs shares its function with a human  
115 protein.

116

117 For instance, biosynthesis of monoamine neurotransmitters (dopamine, serotonin, etc.) requires two  
118 enzymes, tryptophan hydroxylase and tyrosine hydroxylase. These two enzymes appear to have arisen in  
119 bilaterians from duplications of an ancestral phenylalanine hydroxylase [Cao et al., 2010], though evidence  
120 is lacking as to whether this ancestral protein had multiple functions that specialized after duplication (sub-  
121 functionalization) or developed new functions (neofunctionalization) post-duplication. The absence of these  
122 proteins in non-bilaterians seems to be ancestral; in other words, they had not evolved yet when these groups  
123 split and diversified.

124

125 Among other non-bilaterians, some monoamine neurotransmitters are found in cnidarians [Carlberg and  
126 Rosengren, 1985], but are mostly absent in ctenophores (or at detection limit) [Moroz et al., 2014]. Indeed,  
127 previous studies were unable to find homologs of DOPA decarboxylase (AADC, Supplemental Figure 8),  
128 dopamine  $\beta$ -hydroxylase (DBH, Supplemental Figure 7), monoamine oxidase (MAO, Supplemental Fig-  
129 ure 9), or tyrosine hydroxylase (TH) in the genome of the ctenophore *M. leidyi* or any available ctenophore  
130 transcriptome, and it was suggested that some of these proteins were absent in sponges as well (see Supple-



**Figure 3: Neurotransmitter overview across metazoans** Summary schematic of neurotransmitter biosynthesis and degradation pathways across early-branching metazoans. Each of the four non-bilaterian groups is presumed to be monophyletic, although some individual trees of genes or gene families may display alternate topologies. Bold letters refer to the enzymes. Individual gene trees that display the orthology of the clades are found in the supplemental information. Arrows are shown in one direction though many reactions can be reversible. (A) Glutamate and GABA metabolism from the citric acid cycle through the “GABA shunt” pathway. Because ABAT is absent in ctenophores, GLUD and TAT are potentially alternatives to convert  $\alpha$ -ketoglutarate to glutamate. GATP can convert GABA to succinate semialdehyde, but this enzyme was only found in some demosponges and plants. (B) Monoamine metabolism, excluding tryptophan. (C) Table of presence-absence for genes in parts A and B. Presence (green) refers to a 1-to-1 ortholog where orthology is clear from the tree position. Homolog (blue) refers to a sister group position in trees before duplications with different or unknown functions. Secondary loss (red) refers to the gene missing in the clade, but homologs are found in non-metazoan phyla. Numbers inside the boxes indicate copy number specific to that group, M refers to multiple duplications within the group where the copy number is variable among species. Abbreviations for clades are as follows: Cteno, ctenophores; Hsm, homoscleromorphs; Demo, demosponges; Hexact, hexactinellids; Choano, choanoflagellates.

131 mentary Tables 17 and 19 in [Ryan et al., 2013]). However, we found orthologs of MAO and homologs of  
 132 AADC and DBH in several sponges, though it is unclear if they perform the same function as the human  
 133 proteins. Additionally, homologs of four enzymes, AADC, MAO, DBH, and ABAT, are present in single-  
 134 celled eukaryotes but not ctenophores, implying a secondary loss of these protein families in this phylum.

135



## 136 GABA receptors

137 The neurotransmitter gamma-amino butyric acid (GABA) has been shown to affect contraction in *T. wil-*  
138 *helma* [Ellwanger et al., 2007] and the freshwater sponge *E. muelleri* [Elliott and Leys, 2010]. The genome of  
139 *T. wilhelma* contains metabotropic receptors (GABA-B, mGABARs), but not ionotropic GABA receptors  
140 (GABA-A, iGABARs). While humans have two mGABARs and the ctenophore *M. leidyi* has only one, the  
141 *T. wilhelma* genome has nine. Sponges appear to have undergone a large expansion of this protein family  
142 (Figure 4), similar to the expansion of glutamate-binding GPCRs previously observed in sponges [Krishnan  
143 et al., 2014]. Based on the structure of the binding pocket of human GABAR-B1 [Geng et al., 2013], many  
144 differences are observed across the mGABAR protein family, even showing that many residues involved in  
145 coordination of GABA are not conserved between the two human proteins or all other animals (Supple-  
146 mental Figure 11). Contrary to previous reports [Ramoino et al., 2010], we were unable to find normal  
147 mGABARs in the two calcareous sponges *S. ciliatum* and *L. complicata*. Instead, in these two species, the  
148 best BLAST hits from human GABA-B receptors (the putative mGABARs) had the best reciprocal hits to  
149 Insulin-like growth-factor receptors. Structurally, this was due to the normal seven-transmembrane domain  
150 being swapped with a C-terminal protein kinase domain (Figure 5), meaning these are not true metabotropic  
151 GABA receptors. Similarly, in the filasterean *Capsaspora owczarzaki*, the N-terminal ligand binding domain  
152 is also exchanged with other domains, suggesting as well that these are not true metabotropic GABA recep-  
153 tors.

## 155 Glutamate receptors

156 Glutamate is of particular interest as it is a key metabolic intermediate and the main excitatory neurotrans-  
157 mitter in animal nervous systems, acting on two types of receptors: the metabotropic glutamate receptors  
158 (mGluRs) and the ionotropic ones (iGluRs). Some sponge species possess iGluRs, though these receptors  
159 were absent in the transcriptomes of several demosponges [Riesgo et al., 2014a]. We were unable to find  
160 iGluRs in the genome of *T. wilhelma*, in the genome and transcriptomes of any other demosponge, or in the  
161 genomes of two choanoflagellates (*M. brevicollis* and *S. rosetta*). The top BLAST hits in demosponges have  
162 a GPCR domain instead of the ion channel domain, indicating that these are not true iGluRs (Supplemental  
163 Figure 13). Because the domain structure in plants is the same as most animal iGluRs, the ligand binding  
164 domain was swapped out in demosponges.

165  
166 The homoscleromorpha/calcareous clade appears to have an independent expansion of iGluRs (Supplemen-  
167 tal Figure 12), though the normal ion transporter domain is switched with a SBP-bac-3 domain (PFAM  
168 domain PF00497) compared to all other iGluRs (Supplemental Figure 13). Additionally, ctenophores and  
169 placozoans appeared to have dramatic expansions of this protein family as well [Ryan et al., 2013, Moroz  
170 et al., 2014, Alberstein et al., 2015], suggesting that a small set of iGluRs was present in the common ancestor  
171 of eukaryotes and have diversified multiple times in both plants and animals, while other clades appear to  
172 have modified or lost these proteins.

## 174 Vesicular transporters

175 Secretory systems are a common feature of all eukaryotes, as most cells have endoplasmic reticulum to secrete  
176 proteins or make membrane proteins. Neurons secrete peptides (conceptually identical to any other protein)  
177 or small-molecule neurotransmitters in a paracrine fashion, specifically to other neural cells. Compared to  
178 peptides, small-molecule neurotransmitters need to be loaded into vesicles by dedicated transport proteins.  
179 Vesicular glutamate transporters (VGluTs, SLC17A6-8) are part of a superfamily of transporters [Sreedha-  
180 ran et al., 2010] that carry glutamate, aspartate, and nucleotides. The position of sponge proteins in the tree  
181 is inconsistent with a clear role in glutamate transport (Supplemental Figure 14), as several sponge clades  
182 and ctenophores occur as sister group to multiple duplications. Transporters in sponges, ctenophores, and  
183 choanoflagellates may well act upon glutamate or other amino acids, but this needs to be experimentally  
184 investigated.

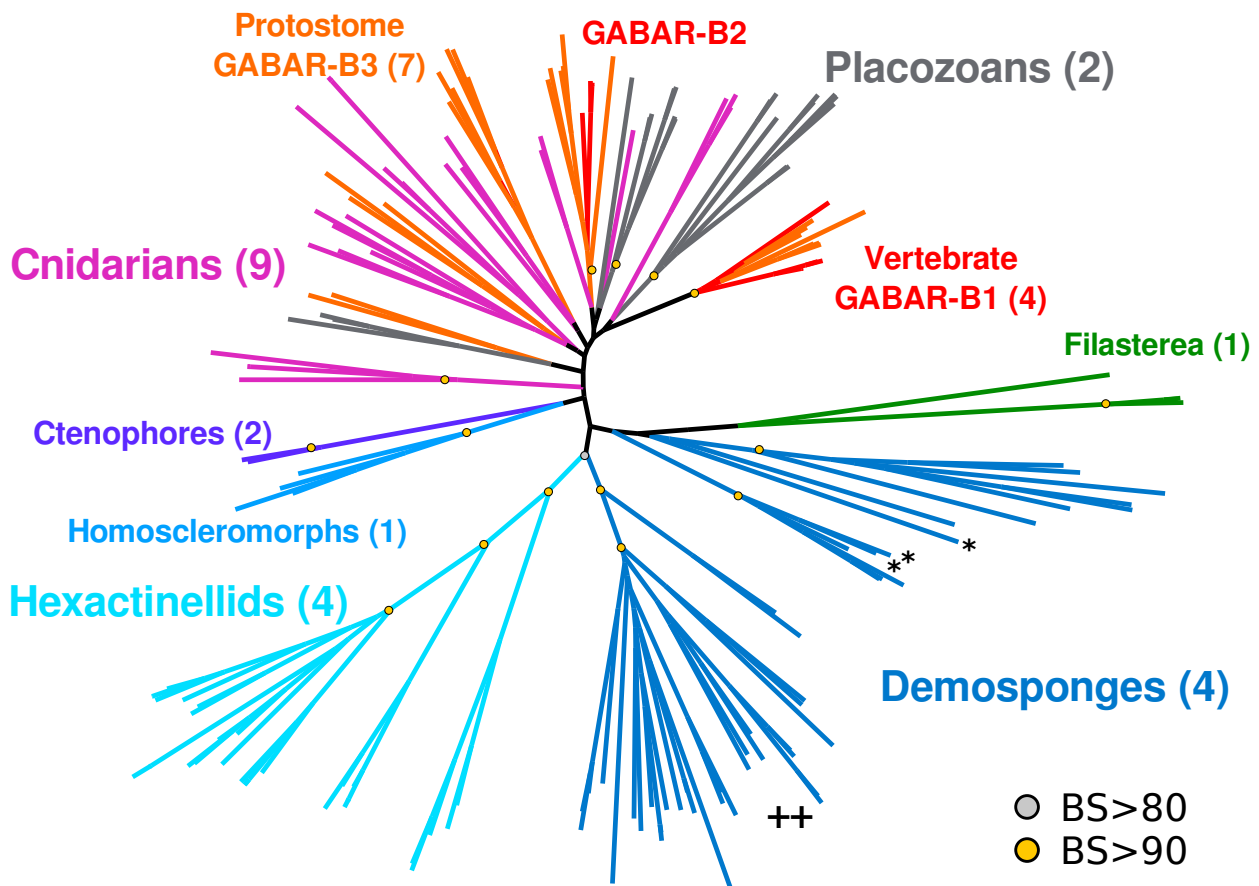


Figure 4: **Metabotropic GABA receptors (GABA-B type) across metazoans** Protein tree generated with RAxML. Numbers in parentheses indicate the number of species from that group, so the 46 demosponge mGABARs come from 4 species. Key bootstrap values are summarized as yellow or gray dots, for values of 90 or more, or 80 or more, respectively. Single star indicates sequences annotated as mGABARs in [Krishnan et al., 2014], double plus indicates the clade annotated as “sponge specific expansion” in [Krishnan et al., 2014]. For complete version with protein names and all bootstrap values, see Supplemental Figure 10

186 Similar to glutamate, GABA is loaded into vesicles with the vesicular inhibitory amino acid transporter  
 187 (VIAAT). Ctenophores, sponges, and placozoans lack one-to-one homologs of VIAAT (Supplemental Fig-  
 188 ure 15). Several other transporters are thought to transport GABA (ANTL or SLC6 class) and many other  
 189 amino acids. SLC6-class transporters, which transport diverse amino acids, are found in all non-bilaterian  
 190 groups, so the function of VIAAT may be redundant.

#### 191 Glycine receptors

192 Glycine is known to affect the contraction of *T. wilhelma* [Ellwanger and Nickel, 2006]. Some ctenophore  
 193 iGluRs have been shown to bind glycine [Alberstein et al., 2015] due to the substitution of serine for arginine  
 194 (S687 in human GluN1), though this appears to be specific only to ctenophores, as essentially all other  
 195 iGluRs have the conserved serine/threonine at this position. Because no ionotropic glycine receptors could  
 196 be identified in the *T. wilhelma* genome (or any other sponges, ctenophores or placozoans), other proteins  
 197 may be responsible for mediating this effect.  
 198

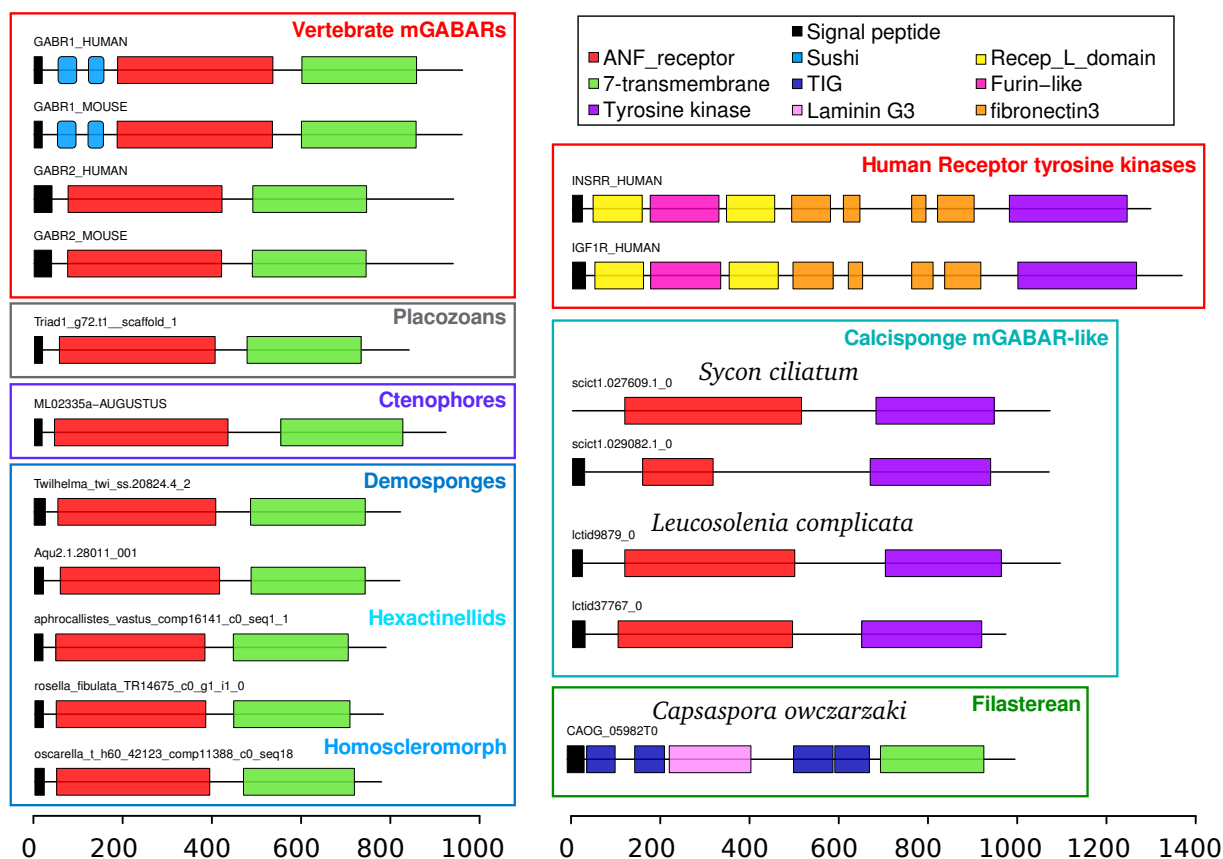


Figure 5: Domain organization of GABA-B type receptors across metazoans

Scale bar displays number of amino acids. Top reciprocal BLAST hits to human for putative mGABARs in calcareous sponges are INSRR and IGF1R, due to high-scoring hits to the tyrosine kinase domain. All mGABARs from demosponges, glass sponges, and the homoscleromorph *Oscarella carmela* share the 7-transmembrane domain (green) with mGABARs from other animals, while calcarean proteins have the same ligand-binding domain (red) but instead have a protein tyrosine kinase domain (purple) at the C-terminus, similar to growth-factor receptors. The filasterean *Capsaspora owczarzaki* has alternate domains at the N-terminus.

## 199 Mechanical receptors

200 Some sponges can contract in response to mechanical agitation, as reported for the demosponges *E. muel-*  
 201 *leri* [Elliott and Leys, 2007] and *T. wilhelma* [Nickel, 2010]. Several diverse protein families appear to  
 202 be responsible for the sense of touch [Árnadóttir and Chalfie, 2010]. A subgroup of the TRP (transient  
 203 receptor-potential) channels, TRP-N, thought to mediate mechanosensation was determined to be absent  
 204 in sponges [Schuler et al., 2015], and we were unable to identify any in either *T. wilhelma* or *S. ciliatum*,  
 205 although other TRP-class channels were found [Ludeman et al., 2014, Schuler et al., 2015]. Because the  
 206 mechanosensory function of TRP channels may be redundant, we analysed for the presence of PIEZO, a  
 207 280kDa trimeric protein [Ge et al., 2015] involved in touch sensation in mammals [Coste et al., 2012]. Al-  
 208 though two homologs were found in vertebrates, we found one copy in all other animals (Supplemental  
 209 Figure 16) as well as fungi, plants and most other eukaryotic groups, suggesting an ancient and conserved  
 210 function of this protein.

211



## 212 Voltage-gated channels

213 Voltage-gated ion channels are necessary for the propagation of electrical signals down axons and dendrites  
214 [Zakon, 2012, Moran et al., 2015], and have specificities for sodium, potassium, or calcium. Previous analyses  
215 were unable to clearly identify potassium or sodium channels in sponges [Liebeskind et al., 2011]; only one  
216 partial potassium channel was found in the transcriptome of the homoscleromorph *Corticium candelabrum*  
217 [Riesgo et al., 2014a, Li et al., 2015]. We were unable to find any voltage-gated sodium or potassium channels  
218 in the genome or transcriptome of any sponge. We then examine voltage-gated hydrogen channels (*hvcn1*), as  
219 these proteins have been found in a number of single-celled eukaryotes [Smith et al., 2011], and are extremely  
220 conserved. These channels were found in all sponge groups, although the high protein identity resulted in a  
221 poorly-resolved tree (Supplemental Figure 19).

222 Reports of action potentials in hexactinellids [Leys et al., 1999, Leys et al., 2007, Nickel, 2010] showed that  
223 sponge action potentials were inhibited by divalent cations [Leys et al., 1999], suggesting a role of calcium  
224 channels instead. Because voltage-gated sodium and calcium channels arose from a duplication event [Gur  
225 Barzilai et al., 2012], the ion selectivity may be variable within this protein family. Most sponges have only a  
226 single CaV-channel (Supplemental Figure 18) and several Hv-channels, and no voltage-gated channel of any  
227 kind was found in any glass sponge. However, all glass sponge sequences are from transcriptomes, therefore  
228 either the expression level of the true channels is low in glass sponges, or they have independently evolved  
229 another mechanism to propagate action potentials.

230

## 231 Discussion

### 232 Gene content variation of metazoa

233 Among the thousands of genes in the genome, we focused on genes that may be mediating contractile be-  
 234 havior in *T. wilhelma*, and the interactions of those genes within broader metabolic pathways. Many of the  
 235 “housekeeping” genes in our study have lineage-specific duplications in at least one animal phylum. Consid-  
 236 ering the importance of “single-copy” proteins in phylogenetic analyses, as taxon sampling improves, it may  
 237 be found that very few or no genes are single copy across most or all animal phyla. Many other genes that  
 238 are critical for neural functioning in bilaterians have independent losses in other animal lineages (Figure 6).  
 239

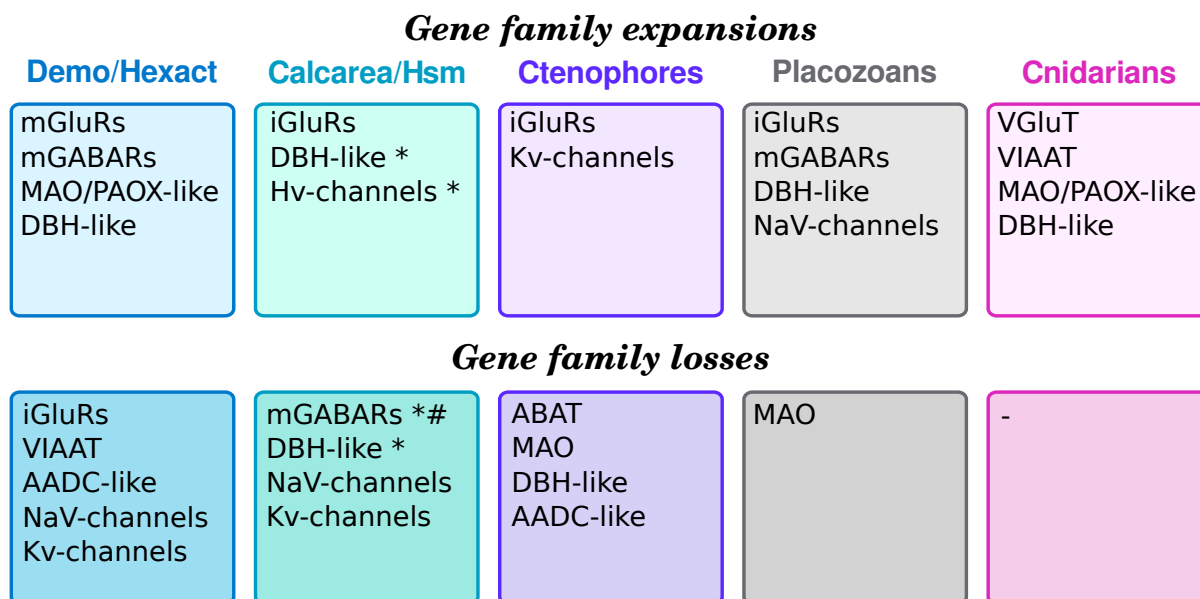


Figure 6: Summary of gene expansions and losses

Demo/Hexact refers to the clade of demosponges and glass sponges. Calcarea/Hsm refers to the unnamed clade of calcareous sponges and homoscleromorphs. The star indicates that expansion or loss was found in one class but not the other. Number sign indicates the domain rearrangement in Calcarea.

### 240 Glutamate and GABA receptor evolution

241 There is stark contrast in the relative abundance of mGABARs and iGluRs in sponges and ctenophores.  
 242 The relative dearth of mGABARs in ctenophores may reflect the apparently absence of amino-butyrate  
 243 amino-transferase (ABAT) in ctenophores, suggesting that ctenophores use an alternate pathway to produce  
 244 glutamate or metabolize GABA (Figure 3), rarely use GABA as a neurotransmitter, or simply are missing  
 245 this pathway. Other aminotransferases such as GLUD or TAT may perform some of the exchange between  
 246  $\alpha$ -ketoglutarate and glutamate, particularly as ctenophores have two copies of GLUD while most animals  
 247 have only one. Ctenophores also have multiple (variable) copies of glutamate synthase and three copies of  
 248 KYAT one of which may serve to balance glutamate metabolism in these animals.

249  
 250 There are two explanations for the diversity of mGABARs in sponges. Given the high variability of  
 251 amino acids in the mGABAR binding pocket (Supplemental Figure 11), it is plausible that many of these  
 252 receptors do not bind GABA at all, and have diversified for other ligands. There is precedent for this as it  
 253 was shown that the independent expansion of ctenophore iGluRs also included several key mutations to the  
 254 binding pocket which changed the ligand specificity of these proteins [Alberstein et al., 2015]. For the other

255 hypothesis, all of the receptors could bind GABA, essentially mediating the same contraction signal, but  
256 their kinetics could differ and be influenced by factors such as, for instance, temperature. Because sponges  
257 are mostly immobile, they often can be subject to environment variation in terms of light, oxygen, and  
258 temperature. The possession of a set of proteins capable of triggering the same response (e.g. contraction)  
259 with varying daily or seasonal environmental conditions (e.g. temperature) would be beneficial and may  
260 explain the diverse set of receptors observed in sponges. Experimental characterization of these binding  
261 domains is necessary and may even show that a combination of these hypotheses explains the diversification  
262 of mGABARs in Porifera.

263  
264 The apparent absence of true mGABARs in calcareous sponges (the genome of *S. ciliatum* and transcrip-  
265 tome of *L. complicata*) conflicts with a previous study that identified key proteins in the GABA pathway  
266 by immunostaining [Ramoino et al., 2010]. The best mGABARs BLAST hits found in the two calcareous  
267 sponges display a conserved ligand binding domain but the seven-transmembrane domain has been swapped  
268 with a tyrosine kinase domain (Figure 5). Structural similarity of the conserved N-terminal domain may  
269 result in a false-positive signal in studies using immunostaining with standard antibodies [Ramoino et al.,  
270 2010]. On the other hand, compared to ctenophores, which apparently lack ABAT, this enzyme was found  
271 in both of the calcareous sponges analyzed. Thus it would be surprising if these sponges had no capacity  
272 to create or respond to GABA. Since true vertebrate-like mGABARs are found in all other sponge classes,  
273 and our study could only examine two calcareous sponges, it could be that mGABAR presence is variable in  
274 this class. The genome of *S. ciliatum* contains 40 proteins annotated as mGluRs [Fortunato et al., 2014], so  
275 a third possibility is that even in the absence of true mGABARs, some of these proteins may have evolved  
276 affinity for GABA and mediate its signaling in calcareous sponges.

277  
278 Although a putative iGluR was identified in the transcriptome of the demosponge *Ircinia fasciculata*,  
279 this sequence was only a fragment, so the glutamate affinity and domain structure could not be determined.  
280 As with the mGABARs, the domain structure is different between the sponge classes. Otherwise, it appears  
281 that only calcareous sponges and homoscleromorphs have NMDA/AMPA-like iGluRs. The presence of these  
282 proteins in plants and other single-celled eukaryotes suggests that at least iGluRs were present in the com-  
283 mon ancestor of all eukaryotes, and their absence in demosponges is likely the product of secondary losses.  
284 In the context of contractions of *T. wilhelma*, the abundance of mGluRs and mGABARs could plausibly  
285 work in antagonistic ways via the action of different G-proteins making ionotropic channels not necessary  
286 for the modulation of this behavior.

## 287 288 **Variation in neurotransmitter metabolism**

289 Many of the oxidative enzymes in the monoamine pathway require molecular oxygen, suggesting an im-  
290 portant role of this molecule both the synthesis of the neurotransmitters (with PAH, TH, and DBH) and  
291 their inactivation (with MAO). Two catabolic pathways arise from tyrosine (Figure 3) and require oxygen  
292 at nearly all steps. It is unclear why intermediate products of one of these two pathways, the catecholamine  
293 pathway, became neurotransmitters and the other did not, particularly as hydroxyphenylpyruvate pathway  
294 is universally found in animals and catabolic intermediates are likely to be ubiquitous.

295  
296 MAO was found in most animal groups, but we were unable to find any in placozoans or ctenophores.  
297 The topology of the MAO phylogenetic tree suggests a secondary loss of this protein in these phyla (Sup-  
298 plemental Figure 9). Related genes (PAOX, polyamine oxidase) were found in placozoans with several  
299 placozoan-specific duplications, and again, potentially one of these may catalyze the oxidation of aromatic  
300 amines. The analysis of these proteins also uncovered a clade including sponges, cnidarians, and lancelets,  
301 though the function of these proteins cannot be predicted based on homology searches. *In vitro* charac-  
302 terization of these enzymes may reveal the function to provide evidence as to how these could have been  
303 important for metabolism in early animals, and was subsequently replaced or lost in most other metazoan  
304 lineages.

305  
306 Remarkably, the DBH group has independent expansions in three sponge classes as well as placozoans

307 and cnidarians (Supplemental Figure 7). No DBH homologs were identified in calcareous sponges or in  
308 ctenophores. A putative homolog of this group was found in the choanoflagellate *M. brevicollis* but not in  
309 any other non-metazoan. The alignment and the phylogenetic position of the *M. brevicollis* protein suggest  
310 that it may be a member of the copper-binding oxygenase superfamily, rather than a true homolog of DBH  
311 (see Supplemental Alignment).

312  
313 The presence of DBH-like and AADC-like enzymes in most animal groups suggests the possibility to  
314 make phenylethanolamines (like octopamine or noradrenaline) from tyrosine, and then subsequently inacti-  
315 vate them with MAO. All demosponges appear to lack AADC, and ctenophores appear to lack both of  
316 these enzymes calling into question a previous report of the detectability of monoamine neurotransmitters  
317 in ctenophores [Carlberg and Rosengren, 1985].  
318

## 319 Conserved properties of neurons

320 Neurons are generally defined by the presence of five key aspects: membrane potential, voltage-gated ion  
321 channels, secretory pathways, ligand-gated ion channels, and cell-cell junctions to form synapses. Voltage-  
322 gated channels, secretory systems, and ligand-gated ion channels are discussed above. Membrane potential  
323 is maintained in animal cells by sodium-potassium pumps (ATP-ases), which are a class of cation pumps  
324 exclusively found in animals [Stein, 1995, Sáez et al., 2009]. It is thought that such pumps are necessary  
325 because animals are the only multicellular group that lacks any kind of cell wall, thus careful control of  
326 ionic balance is necessary to resist osmotic stress [Stein, 1995]. For non-bilaterian animals, cell layers were  
327 in direct contact with water, so potentially all cells needed this protein to function normally. Therefore  
328 having neuron-like functionality is unlikely to rest upon the gain or loss of this gene. The last feature is  
329 the presence of cell-cell connections. Many proteins involved in synapse structure or neurotransmission are  
330 found in sponges, [Srivastava et al., 2010, Riesgo et al., 2014a, Moran et al., 2015, Leys, 2015] though it is not  
331 clear which genes are necessary for neural functioning, or may have evolved independently.

## 332 Neural evolution and losses

333 Based on recent phylogenies, both Porifera-sister and Ctenophora-sister evolutionary scenarios require either  
334 at least one loss of neurons or two independent gains (Figure 1) of this cell-type. The only scenario that  
335 allows for a single evolution of neurons and no losses is the “Coelenterata” hypothesis (reviewed in [Jékely  
336 et al., 2015]), which joins cnidarians and ctenophores in a clade. However, many molecular datasets [Dunn  
337 et al., 2008, Ryan et al., 2013, Whelan et al., 2015, Pisani et al., 2015] and morphological evidence [Harbison,  
338 1985] argue against this scenario (but also see [Philippe et al., 2009] and [Simion et al., 2017]). One other  
339 alternative is that placozoans have an unidentified neuron-like cell in a Porifera-sister context, which would  
340 therefore allow for a single origin of neural systems in animals and no losses.

341  
342 What do the two different scenarios mean for evolution of neuronal cells? Considering the basic properties  
343 of neurons related to electrical signaling or secretory pathways, it had been shown before that many of the  
344 genes involved are universally found in animals. A single origin and multiple losses implies that the genetic  
345 toolkit necessary for all of these functions was present in the same single-celled organism or the same cell  
346 type (an hypothetical proto-neuron) of the last common ancestor of crown-group Metazoa, and either that  
347 cell type was lost or its functions were split up.

348  
349 Sponges and ctenophores both appear to have lost several gene families (Figure 6), though ctenophores  
350 nonetheless have neural cells. Thus, the losses of the GABA or monoamine pathways are not critical for the  
351 functioning of neural cell types overall. However, voltage-gated potassium and sodium channels are thought  
352 to be essential for the propagation of electrical signals down axons and dendrites and have been found in  
353 all animal groups except sponges [Moran et al., 2015]. The NaV-channel tree shows a single origin of this  
354 protein family (Supplemental Figure 17), and presence of these channels in choanoflagellates suggests they  
355 were present in the common ancestor of all animals; the apparent absence in sponges therefore is probably a  
356 secondary loss. By comparison, ctenophores have a mostly-unique expansion of Kv-channels relative to the  
357 rest of metazoans [Li et al., 2015] and a duplication in NaV channels. Together with the loss of this protein

358 family in sponges, the gene content argues for a combination of both multiple, independent gains and a loss  
359 of neural-type cells and their associated functions across animals.

360  
361 Properties of the earliest metazoans are unknown, including life cycle or number of cell types, but it is  
362 most parsimonious that the first obligate multicellular animals did not have anything resembling a modern  
363 bilaterian nervous system [Wray et al., 2015]. Yet, the genomic evidence shows that these animals have cel-  
364 lular capacity to respond to environmental or paracrine signals, regulate the cell internal ion concentrations  
365 and respond to changes in their concentrations, and secrete small molecules that could serve as effectors  
366 in unconnected (but proximal) cells. Thus, the earliest animals likely had the capacity to develop nerve  
367 cells using the genetic toolkit they possessed, though the number of times this occurred is unclear. This  
368 capacity appears to have been lost in sponges with the loss of voltage-gated channels. As we were unable to  
369 find putative genes to mediate action potentials in glass sponges, either all of the four transcriptomes were  
370 incomplete or the unique action potentials of glass sponges may represent a third case of the evolution of  
371 neural-like functions in Metazoa.

372

## 373 **Methods**

### 374 **Sequence data**

375 Project overview can be found at [spongebase.net](http://spongebase.net). Reference data from the demosponge *Tethya wilhelma*  
376 are available at: [https://bitbucket.org/molpalmuc/tethya\\_wilhelma-genome](https://bitbucket.org/molpalmuc/tethya_wilhelma-genome)

377

378 Raw genomic reads for *T. wilhelma* are available on NCBI SRA under accession numbers SRR2163223  
379 (genomic reads), SRR2296844 (mate pairs), SRR5369934 (DNA Molecule), and SRR4255675 (RNAseq).

380

### 381 **Genome assembly**

#### 382 **Processing and assembly**

383 We generated 25Gb of 100bp paired-end Illumina reads of genomic DNA and 35Gb of 125bp Illumina gel-free  
384 mate-pair reads. Contigs were assembled with SOAPdenovo2 [Luo et al., 2012] using a kmer of 83bp. We  
385 also generated 436Mb of Molecule synthetic long reads. Because both haplotypes are represented in the  
386 Molecule reads, we merged the Molecule reads using HaploMerger [Huang et al., 2012]. Contigs and merged  
387 Molecule reads were then scaffolded using the gel-free mate-pairs with SSPACE [Boetzer et al., 2011] and  
388 BESST [Sahlin et al., 2014]. The first draft assembly had 7,947 contigs, totaling 145 megabases.

389

#### 390 **Removal of low-coverage contigs**

391 To examine the completeness of the genome, we generated a plot of kmer coverage against GC percentage for  
392 the contigs (Supplemental Figure 1) using custom Python scripts (available at <http://github.org/wrf/lavaLampPlot>).  
393 This revealed 1,040 contigs with a coverage of zero that were carried over from the Molecule reads and were  
394 not assembled (Supplemental Figure 2), accounting for 6 megabases. As these reads likely derived from  
395 bacterial contamination in the aquarium water, these contigs were removed, leaving 6,907 contigs totalling  
396 138 megabases.

397

#### 398 **Separation of bacterial contigs**

399 Additionally, the plot revealed many contigs with lower coverage (20x-90x) and high GC content (50-75%)  
400 suggesting the presence of bacteria (Supplemental Figure 3). Because many of these contigs were shorter  
401 than 10kb, separation of the bacterial contigs was done through several steps. We found 4,858 contigs with  
402 mapped RNAseq reads and GC content under 50%, as expected of metazoans. These contigs accounted for

403 88% of the sponge assembly, or 121 megabases. For the 2,014 contigs with no mapped RNAseq, we used  
404 blastn to search the contigs against the *A. queenslandica* scaffolds and all complete bacterial genomes from  
405 Genbank (5,242 sequences). Based on subtraction of bitscores, 62 contigs were identified as sponge and 565  
406 were identified as bacterial. For the remaining 1,387 contigs, most of which were under 10kb, we repeated  
407 the search with tblastx against *A. queenslandica* scaffolds and the genomes of *Sinorhizobium medicae* and  
408 *Roseobacter litoralis*, which were the most similar complete genomes to the two bacterial 16S rRNAs identi-  
409 fied in the contigs. After all sorting, 798 putative bacterial contigs accounted for 12.7 megabases and were  
410 separated to bring the total to 6,109 sponge contigs. Contigs for the two bacteria were binned by tetranu-  
411 cleotide frequency using MetaWatt [Strous et al., 2012] (Supplemental Figure 4).

## 413 Genome coverage and completeness

414 Coverage was estimated two ways: kmer frequency and read mapping. Kmers of 31bp were counted using  
415 the Jellyfish kmer counter [Marçais and Kingsford, 2011] and analyzed using custom Python and R scripts  
416 (“fastqDumps2histo.py” and “jellyfish\_gc\_coverage\_blob\_plot\_v2.R”, available at [http://github.org/wrf/](http://github.org/wrf/lavaLampPlot)  
417 [lavaLampPlot](http://github.org/wrf/lavaLampPlot)). As expected, the kmer distribution showed two peaks, one for kmers at heterozygous  
418 positions and one for homozygous positions, whereupon the coverage peak was at 131-fold coverage for ho-  
419 mozygous positions. Because of sequencing errors, this method often underestimates coverage, and so to  
420 confirm this estimate we then mapped all reads to the genome using Bowtie2 [Langmead and Salzberg, 2012].  
421 The sum of mapped reads divided by the total length provided an estimated coverage of 159-fold physical  
422 coverage.

423  
424 Of the original reads, 185 million (86.5%) mapped back to the assembled sponge contigs. Completeness  
425 for gene content was assessed with BUSCO [Simão et al., 2015], whereupon we found 728 (86%) complete  
426 genes and 42 (4.9%) predicted-incomplete genes. Overall, these data suggest that the genome assembly is  
427 adequate for downstream analyses.

## 429 Genome annotation

### 430 Transcriptome versions

431 The transcriptome for *T. wilhelma* was assembled *de novo* using Trinity (release r20140717) [Grabherr et al.,  
432 2011, Haas et al., 2013]. Default parameters were used, except for strand specific assembly, *in silico* read  
433 normalization, and trimming (-SS\_lib\_type RF -normalize\_reads -trimmomatic). This produced 127,012  
434 transcripts with an average length of 913bp. Assembled transcripts were mapped to the genomic assembly  
435 using GMAP [Wu and Watanabe, 2005] to produce a GFF file of the transcript mapping. Of these, 114,744  
436 transcripts were mapped 166,847 times, allowing for multiple mappings.

437  
438 For the genome-guided transcriptome, strand-specific RNAseq reads were mapped against the genome  
439 build using Tophat2 v2.0.13 [Kim et al., 2013] using strand-specific mapping (option -library-type fr-  
440 firststrand) and otherwise default parameters. Mapped reads were then joined into transcripts using StringTie  
441 v1.0.2 [Pertea et al., 2015] with default parameters.

442  
443 Additionally, *ab initio* gene models were predicted using AUGUSTUS [Stanke et al., 2008]. AUGUSTUS  
444 was trained on the webAugustus server [Hoff and Stanke, 2013] using the highest expression transcripts for  
445 each Trinity component and the assembled contigs. This identified 27,551 putative genes. The majority of  
446 these overlapped partially or completely with a predicted gene based on the Trinity mapping or Stringtie  
447 genes. However, 3,866 genes (4,321 transcripts) had no overlap with any predicted exon from either the  
448 Trinity or StringTie set, and were kept for the final set. Considering the possibility that some of these may  
449 be pseudogenes, we aligned these proteins to the SwissProt database with BLASTP [Camacho et al., 2009].  
450 Of these, only 759 had reliable hits (E-value < 10<sup>-5</sup>) to 688 proteins. The annotated functions were diverse,  
451 including proteins similar to many receptors and large structural proteins such as fibrillin (potentially any  
452 protein with EGF repeats), dynein heavy chain, and titin; because very large proteins may be split across



453 multiple contigs, the predicted genes may be only fragments of the full gene. Only 42 of the hits were against  
454 transposable elements.

### 455 **Filtering of the final gene set**

456 Because assembly of transcripts for both StringTie and Trinity relies on overlaps in the genome or RNAseq  
457 reads, genes that overlap in the untranslated regions (UTRs) can sometimes be erroneously fused. For  
458 StringTie, we developed a custom Python script to separate non-overlapping transcripts belonging to the  
459 same “gene” (`stringtiesplitgenes.py`, available at <https://bitbucket.org/wrf/sequences/>). Tandem du-  
460 plications can lead to RNAseq reads bridging the two tandem copies and result in both copies being called  
461 the same gene. The original StringTie set contained 46,572 transcripts for 32,112 genes, while the corrected  
462 set contained 33,200 genes and identified 1,088 new non overlapping genes.

463  
464 Positional errors in the genome or allelic variations may result in some RNAseq reads not mapping  
465 to the genome, so some genes are fragmented in the genome-guided transcriptome but not the *de novo*  
466 assembly. Making use of the protein predictions from TransDecoder, we compared the predicted pro-  
467 teins between the two transcriptomes using a custom Python script (`transdecodersplitgenes.py`, available  
468 at <https://bitbucket.org/wrf/sequences/>). This identified 406 StringTie transcripts that were better  
469 modeled by Trinity transcripts.

### 471 **Functional gene annotation**

472 Many genes of functional importance were examined manually, and the best transcript from StringTie, Trin-  
473 ity, or AUGUSTUS was retained for the final gene set. In the GFF and fasta versions of the transcriptomes,  
474 names of protein functions were assigned several ways. Target genes that were manually curated and edited,  
475 such as those used in all trees, are named by the generic function or the annotated function of the closest  
476 human protein. For instance, the dopamine beta-hydroxylase (DBH) homolog in *T. wilhelma* was manu-  
477 ally corrected, and the position in a phylogenetic tree demonstrated that demosponges diverged before the  
478 duplication which created DBH and the two DBH-like proteins in humans, thus the *T. wilhelma* protein  
479 is annotated as DBH-like. Secondly, automated ortholog finding pipelines (HaMStR [Ebersberger et al.,  
480 2009]) used for phylogeny [Cannon et al., 2016] have identified homologs in *T. wilhelma*, which have been  
481 manually checked based on positions in the phylogenetic trees. Thirdly, single-direction BLAST results were  
482 kept as annotations provided that the BLAST hit had a bitscore over 1000, or a bitscore over 300 and the  
483 *T. wilhelma* protein covered at least 75% of the best hit against the human protein dataset from SwissProt.  
484 The bitscore and length cutoffs were applied to reduce the number of annotations based on a single domain.

### 486 **Analysis of splice variation**

487 Using the transcriptome from StringTie, splice variation was assessed using a custom Python script (`splice-`  
488 `variantstats.py`, available at <https://bitbucket.org/wrf/sequences/>). In this script, several ambiguous  
489 definitions were clarified to define the different splice types. Firstly, single exon genes with no variants are  
490 distinguished from single exon genes with variations, that is, a gene with two exons can have a variant with  
491 one exon. For loci with only two transcripts, the canonical or main transcript is defined as the one with  
492 the higher expression level, as measured by the higher FPKM value reported from StringTie. For loci with  
493 three or more transcripts, main or canonical exons are those included in at least two transcripts. A cassette  
494 exon must occur in less than 50% of the transcripts for a locus, otherwise such case is defined as a skipped  
495 exon. A retained intron is any portion that exactly spans two other exons; for highly expressed transcripts  
496 this may include erroneously retained introns due to intermediates in splicing. A summary of the splicing  
497 types is displayed in Supplemental Table 2.

498  
499 Intron retention was recently reported to be a common mode of alternative splicing in *A. queens-*  
500 *landica* [Fernandez-Valverde et al., 2015]. We found 3,295 transcripts with 3,565 retained intron events  
501 (Supplemental Table 2). We then analyzed the length of the retained introns and found the phase of the

502 retained piece to be randomly distributed (unlike cassette exons, Supplemental Table 3), suggesting that  
503 many of the retained introns result from incomplete splicing rather than functional retention.

## 505 Microsynteny across sponges

506 Putative synteny blocks were identified using a custom Python script (microsynteny.py, available at <https://bitbucket.org/wrf/sequences/>). Briefly, the script combines the gene positions on scaffolds for both  
507 the query and the reference with BLASTX hits for the query against the reference. If a minimum of three  
508 genes in a row on a query scaffold match to different genes on the same reference scaffold, the group is  
509 kept. By default, this mandated a gap of no more than five genes before discarding the block, and that the  
510 next gene must occur within 30kb. This method was designed to work for highly fragmented genomes with  
511 thousands of scaffolds, so the order and direction of the corresponding genes on the reference scaffold do not  
512 need to match those of the query scaffold.

514  
515 StringTie transcripts for *T. wilhelma* were aligned against the *A. queenslandica* v2.0 protein set with  
516 BLASTX [Camacho et al., 2009], and positions were taken from the accompanying *A. queenslandica* v2.0  
517 GTF. The same procedure was attempted against the *S. ciliatum* gene models, though essentially no syn-  
518 tenic blocks were detected, indicating either substantial differences in gene content or gene order between  
519 demosponges and calcareous sponges.

## 521 Collection of reference data

522 Proteins for *Oikopleura dioica* [Denoeud et al., 2010] were downloaded from Genoscope. Gene models  
523 for *Ciona intestinalis* [Dehal et al., 2002], *Branchiostoma floridae* [Putnam et al., 2008], *Trichoplax ad-*  
524 *herens* [Srivastava et al., 2008], *Capitella teleta*, *Lottia gigantea*, *Helobdella robusta* [Simakov et al., 2013],  
525 *Saccoglossus kowalevskii* [Simakov et al., 2015], and *Monosiga brevicollis* [King et al., 2008] were downloaded  
526 from the JGI genome portal. Gene models for *Sphaeroforma arctica*, *Capsaspora owczarzaki* [Suga et al.,  
527 2013] and *Salpingoeca rosetta* [Fairclough et al., 2013] were downloaded from the Broad Institute.

528  
529 We used genomic data of the cnidarians *Nematostella vectensis* [Moran et al., 2014], *Exaiptasia pall-*  
530 *ida* [Baumgarten et al., 2015], and *Hydra magnipapillata* as well as transcriptomes from 33 other cnidari-  
531 ans [Bhattacharya et al., 2016, Zapata et al., 2015, Pralong et al., 2015, Brinkman et al., 2015, Ponce et al.,  
532 2016], mostly corals.

533  
534 For demosponges, we used the genome of *Amphimedon queenslandica* [Srivastava et al., 2010, Fernandez-  
535 Valverde et al., 2015] and transcriptomic data from: *Mycale phyllophila* [Qiu et al., 2015], *Petrosia fici-*  
536 *formis* [Riesgo et al., 2014a], *Crambe crambe* [Versluis et al., 2015], *Cliona varians* [Riesgo et al., 2014b], *Hal-*  
537 *isarca dujardini* [Borisenko et al., 2016], *Crella elegans* [Pérez-Porro et al., 2013], *Stylissa carteri*, *Xestospon-*  
538 *gia testudinaria* [Ryu et al., 2016], *Scopalina* sp., and *Tedania anhelens*. We used data from the genome of the  
539 calcareous sponge *Sycon ciliatum* [Fortunato et al., 2014] and the transcriptome of *Leucosolenia complicata*.  
540 For hexactinellids (glass sponges), we used transcriptome data from *Aphrocallistes vastus* [Ludeman et al.,  
541 2014], *Hyalonema populiferum*, *Rosella fibulata*, and *Sympagella nux* [Whelan et al., 2015]. For homosclero-  
542 morphs, we used two transcriptomes from *Oscarella carmela* and *Corticium candelabrum* [Ludeman et al.,  
543 2014].

544  
545 We used data from the two published draft genomes of ctenophores [Ryan et al., 2013, Moroz et al., 2014],  
546 as well as transcriptome data from 11 additional ctenophores: *Bathocyroe fosteri*, *Bathocyroe chuni*, *Beroe*  
547 *abyssicola*, *Bolinopsis infundibulum*, *Charistephane fugiens*, *Dryodora glandiformis*, *Euplokamis dunlapae*,  
548 *Hormiphora californensis*, *Lamnea lactea*, *Thalassocalyce inconstans*, and *Velamen parallelum*.

549  
550 We used data from the unpublished draft genome of a novel placozoan species, designated H13.

551

## 552 Gene trees

553 For protein trees, candidate proteins were identified by reciprocal BLAST alignment using blastp or tblastn.  
554 All BLAST searches were done using the NCBI BLAST 2.2.29+ package [Camacho et al., 2009]. Because  
555 most functions were described for human, mouse, or fruit fly proteins, these served as the queries for all  
556 datasets. Candidate homologs were kept for analysis if they reciprocally aligned by blastp to a query pro-  
557 tein, usually human. Alignments for protein sequences were created using MAFFT v7.029b, with L-INS-i  
558 parameters for accurate alignments [Katoh and Standley, 2013]. Phylogenetic trees were generated using ei-  
559 ther FastTree [Price et al., 2010] with default parameters or RAxML-HPC-PTHREADS v8.1.3 [Stamatakis,  
560 2014], using the PROTGAMMALG model for proteins and 100 bootstrap replicates with the “rapid boot-  
561 strap” (-f a) algorithm and a random seed of 1234.

## 563 Domain annotation

564 Domains for individual protein trees were annotated with “hmmScan” v3.1b1 from the HHMER pack-  
565 age [Eddy, 2011] using the PFAM-A database v27.0 [Finn et al., 2016] as queries. Signal peptides were pre-  
566 dicted using the stand-alone version of SignalP v4.1 [Petersen et al., 2011]. Domain structures were visualized  
567 using a custom Python script, “pfampipeline.py”, available at <https://github.com/wrf/genomeGTFtools>.

## 569 Acknowledgments

570 W.R.F would like to thank K. Achim, M. Nickel, J. Musser, J. Ryan, and I. Oldenburg for helpful comments.  
571 G.W and D.E. would like to thank M. Nickel for providing the initial *T. wilhelma* specimens to set up the  
572 culture in Munich. This work was supported by a LMUexcellent grant (Project MODELSPONGE) to D.E.  
573 and G.W. as part of the German Excellence Initiative, partially by research grant 9278 (“Early evolution  
574 of multicellular sponges”) from VILLUM FONDEN to G.W., and NIH grant NIGMS-5-R01-GM087198 to  
575 S.H.D.H. The authors declare no competing interests.

## 576 References

### 577 References

- 578 [Alberstein et al., 2015] Alberstein, R., Grey, R., Zimmet, A., Simmons, D. K., and Mayer, M. L.  
579 (2015). Glycine activated ion channel subunits encoded by ctenophore glutamate receptor genes.  
580 *Proceedings of the National Academy of Sciences*, 112(44):E6048–E6057.
- 581 [Árnadóttir and Chalfie, 2010] Árnadóttir, J. and Chalfie, M. (2010). Eukaryotic Mechanosensitive  
582 Channels. *Annual Review of Biophysics*, 39(1):111–137.
- 583 [Baumgarten et al., 2015] Baumgarten, S., Simakov, O., Esherrick, L. Y., Liew, Y. J., Lehnert, E. M.,  
584 Mitchell, C. T., Li, Y., Hambleton, E. a., Guse, A., Oates, M. E., Gough, J., Weis, V. M., Aranda,  
585 M., Pringle, J. R., and Voolstra, C. R. (2015). The genome of *Aiptasia*, a sea anemone model for  
586 coral symbiosis. *Proceedings of the National Academy of Sciences*, page 201513318.
- 587 [Bhattacharya et al., 2016] Bhattacharya, D., Agrawal, S., Aranda, M., Baumgarten, S., Belcaid, M.,  
588 Drake, J. L., Erwin, D., Foret, S., Gates, R. D., Gruber, D. F., Kamel, B., Lesser, M. P., Levy,  
589 O., Liew, Y. J., MacManes, M., Mass, T., Medina, M., Mehr, S., Meyer, E., Price, D. C., Putnam,  
590 H. M., Qiu, H., Shinzato, C., Shoguchi, E., Stokes, A. J., Tambutté, S., Tchernov, D., Voolstra,  
591 C. R., Wagner, N., Walker, C. W., Weber, A. P., Weis, V., Zelzion, E., Zoccola, D., and Falkowski,  
592 P. G. (2016). Comparative genomics explains the evolutionary success of reef-forming corals. *eLife*,  
593 5:1–26.
- 594 [Boetzer et al., 2011] Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D., and Pirovano, W. (2011).  
595 Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*, 27(4):578–579.

- 596 [Borisenko et al., 2016] Borisenko, I., Adamski, M., Ereskovsky, A., and Adamska, M. (2016). Sur-  
597 prisingly rich repertoire of Wnt genes in the demosponge *Halisarca dujardini*. *BMC evolutionary*  
598 *biology*, 16(1):123.
- 599 [Brinkman et al., 2015] Brinkman, D. L., Jia, X., Potriquet, J., Kumar, D., Dash, D., Kvaskoff, D.,  
600 and Mulvenna, J. (2015). Transcriptome and venom proteome of the box jellyfish *Chironex fleckeri*.  
601 *BMC Genomics*, 16(1):407.
- 602 [Camacho et al., 2009] Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer,  
603 K., and Madden, T. L. (2009). BLAST+: architecture and applications. *BMC bioinformatics*,  
604 10:421.
- 605 [Cannon et al., 2016] Cannon, J. T., Vellutini, B. C., Smith, J., Ronquist, F., Jondelius, U., and  
606 Hejnol, A. (2016). Xenacoelomorpha is the sister group to Nephrozoa. *Nature*, 530(7588):89–93.
- 607 [Cao et al., 2010] Cao, J., Shi, F., Liu, X., Huang, G., and Zhou, M. (2010). Phylogenetic analysis  
608 and evolution of aromatic amino acid hydroxylase. *FEBS Letters*, 584(23):4775–4782.
- 609 [Carlberg and Rosengren, 1985] Carlberg, M. and Rosengren, E. (1985). Biochemical basis for adren-  
610 ergic neurotransmission in coelenterates. *Journal of Comparative Physiology B*, 155(2):251–255.
- 611 [Coste et al., 2012] Coste, B., Xiao, B., Santos, J. S., Syeda, R., Grandl, J., Spencer, K. S., Kim, S. E.,  
612 Schmidt, M., Mathur, J., Dubin, A. E., Montal, M., and Patapoutian, A. (2012). Piezo proteins are  
613 pore-forming subunits of mechanically activated channels. *Nature*, 483(7388):176–181.
- 614 [Dehal et al., 2002] Dehal, P., Satou, Y., Campbell, R. K., Chapman, J., Degnan, B., De Tomaso, A.,  
615 Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein, D. M., Harafuji, N., Hastings, K. E. M., Ho,  
616 I., Hotta, K., Huang, W., Kawashima, T., Lemaire, P., Martinez, D., Meinertzhagen, I. a., Nacula,  
617 S., Nonaka, M., Putnam, N., Rash, S., Saiga, H., Satake, M., Terry, A., Yamada, L., Wang, H.-G.,  
618 Awazu, S., Azumi, K., Boore, J., Branno, M., Chin-Bow, S., DeSantis, R., Doyle, S., Francino, P.,  
619 Keys, D. N., Haga, S., Hayashi, H., Hino, K., Imai, K. S., Inaba, K., Kano, S., Kobayashi, K.,  
620 Kobayashi, M., Lee, B.-I., Makabe, K. W., Manohar, C., Matassi, G., Medina, M., Mochizuki, Y.,  
621 Mount, S., Morishita, T., Miura, S., Nakayama, A., Nishizaka, S., Nomoto, H., Ohta, F., Oishi, K.,  
622 Rigoutsos, I., Sano, M., Sasaki, A., Sasakura, Y., Shoguchi, E., Shin-i, T., Spagnuolo, A., Stainier,  
623 D., Suzuki, M. M., Tassy, O., Takatori, N., Tokuoka, M., Yagi, K., Yoshizaki, F., Wada, S., Zhang,  
624 C., Hyatt, P. D., Larimer, F., Detter, C., Doggett, N., Glavina, T., Hawkins, T., Richardson,  
625 P., Lucas, S., Kohara, Y., Levine, M., Satoh, N., and Rokhsar, D. S. (2002). The draft genome  
626 of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science (New York, N.Y.)*,  
627 298(5601):2157–2167.
- 628 [Denoeud et al., 2010] Denoeud, F., Henriot, S., Mungpakdee, S., Aury, J.-M., Da Silva, C.,  
629 Brinkmann, H., Mikhaleva, J., Olsen, L. C., Jubin, C., Canestro, C., Bouquet, J.-M., Danks, G.,  
630 Poulain, J., Campsteijn, C., Adamski, M., Cross, I., Yadetie, F., Muffato, M., Louis, A., Butcher,  
631 S., Tsagkogeorga, G., Konrad, A., Singh, S., Jensen, M. F., Cong, E. H., Eikeseth-Otteraa, H.,  
632 Noel, B., Anthouard, V., Porcel, B. M., Kachouri-Lafond, R., Nishino, A., Ugolini, M., Chourrout,  
633 P., Nishida, H., Aasland, R., Huzurbazar, S., Westhof, E., Delsuc, F., Lehrach, H., Reinhardt, R.,  
634 Weissenbach, J., Roy, S. W., Artiguenave, F., Postlethwait, J. H., Manak, J. R., Thompson, E. M.,  
635 Jaillon, O., Du Pasquier, L., Boudinot, P., Liberles, D. a., Volf, J.-N., Philippe, H., Lenhard, B.,  
636 Crollius, H. R., Wincker, P., and Chourrout, D. (2010). Plasticity of Animal Genome Architecture  
637 Unmasked by Rapid Evolution of a Pelagic Tunicate. *Science*, 1381(2010).
- 638 [Dunn et al., 2008] Dunn, C. W., Hejnol, A., Matus, D. Q., Pang, K., Browne, W. E., Smith, S. a.,  
639 Seaver, E., Rouse, G. W., Obst, M., Edgecombe, G. D., Sørensen, M. V., Haddock, S. H. D.,  
640 Schmidt-Rhaesa, A., Okusu, A., Kristensen, R. M., Wheeler, W. C., Martindale, M. Q., and Giribet,  
641 G. (2008). Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature*,  
642 452(7188):745–9.
- 643 [Dunn et al., 2015] Dunn, C. W., Leys, S. P., and Haddock, S. H. D. (2015). The hidden biology of  
644 sponges and ctenophores. *Trends in Ecology & Evolution*, pages 1–10.
- 645 [Ebersberger et al., 2009] Ebersberger, I., Strauss, S., and von Haeseler, A. (2009). HaMStR: profile  
646 hidden markov model based search for orthologs in ESTs. *BMC evolutionary biology*, 9:157.

- 647 [Eddy, 2011] Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Computational Biology*,  
648 7(10).
- 649 [Elliott and Leys, 2007] Elliott, G. R. D. and Leys, S. P. (2007). Coordinated contractions effectively  
650 expel water from the aquiferous system of a freshwater sponge. *Journal of Experimental Biology*,  
651 210(21):3736–3748.
- 652 [Elliott and Leys, 2010] Elliott, G. R. D. and Leys, S. P. (2010). Evidence for glutamate, GABA and  
653 NO in coordinating behaviour in the sponge, *Ephydatia muelleri* (Demospongiae, Spongillidae). *The*  
654 *Journal of experimental biology*, 213:2310–2321.
- 655 [Ellwanger et al., 2007] Ellwanger, K., Eich, A., and Nickel, M. (2007). GABA and glutamate specif-  
656 ically induce contractions in the sponge *Tethya wilhelma*. *Journal of Comparative Physiology A:*  
657 *Neuroethology, Sensory, Neural, and Behavioral Physiology*, 193(1):1–11.
- 658 [Ellwanger and Nickel, 2006] Ellwanger, K. and Nickel, M. (2006). Neuroactive substances specifically  
659 modulate rhythmic body contractions in the nerveless metazoan *Tethya wilhelma* (Demospongiae,  
660 Porifera). *Frontiers in zoology*, 3:7.
- 661 [Fairclough et al., 2013] Fairclough, S. R., Chen, Z., Kramer, E., Zeng, Q., Young, S., Robertson,  
662 H. M., Begovic, E., Richter, D. J., Russ, C., Westbrook, M. J., Manning, G., Lang, B. F., Haas,  
663 B., Nusbaum, C., and King, N. (2013). Premetazoan genome evolution and the regulation of cell  
664 differentiation in the choanoflagellate *Salpingoeca rosetta*. *Genome biology*, 14(2):R15.
- 665 [Fernandez-Valverde et al., 2015] Fernandez-Valverde, S. L., Calcino, A. D., and Degnan, B. M. (2015).  
666 Deep developmental transcriptome sequencing uncovers numerous new genes and enhances gene  
667 annotation in the sponge *Amphimedon queenslandica*. *BMC Genomics*, 16(1):1–11.
- 668 [Finn et al., 2016] Finn, R. D., Coghill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L.,  
669 Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J., and Bateman,  
670 A. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids*  
671 *Research*, 44(D1):D279–D285.
- 672 [Fortunato et al., 2014] Fortunato, S. a. V., Adamski, M., Ramos, O. M., Leininger, S., Liu, J., Ferrier,  
673 D. E. K., and Adamska, M. (2014). Calcisponges have a ParaHox gene and dynamic expression of  
674 dispersed NK homeobox genes. *Nature*, 514(7524):620–623.
- 675 [Ge et al., 2015] Ge, J., Li, W., Zhao, Q., Li, N., Chen, M., Zhi, P., Li, R., Gao, N., Xiao, B., and Yang,  
676 M. (2015). Architecture of the mammalian mechanosensitive Piezo1 channel. *Nature*, 527(5):64–69.
- 677 [Geng et al., 2013] Geng, Y., Bush, M., Mosyak, L., Wang, F., and Fan, Q. R. (2013). Structural  
678 mechanism of ligand activation in human GABA(B) receptor. *Nature*, 504(7479):254–9.
- 679 [Grabherr et al., 2011] Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. a.,  
680 Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N.,  
681 Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N.,  
682 and Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference  
683 genome. *Nature biotechnology*, 29(7):644–52.
- 684 [Grell and Benwitz, 1974] Grell, K. G. and Benwitz, G. (1974). Spezifische Verbindungsstrukturen der  
685 Faserzellen von *Trichoplax adhaerens* F.E. Schulze. *Z. Naturforsch.*, 29c:790.
- 686 [Gur Barzilai et al., 2012] Gur Barzilai, M., Reitzel, A. M., Kraus, J. E. M., Gordon, D., Technau, U.,  
687 Gurevitz, M., and Moran, Y. (2012). Convergent Evolution of Sodium Ion Selectivity in Metazoan  
688 Neuronal Signaling. *Cell Reports*, 2(2):242–248.
- 689 [Haas et al., 2013] Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden,  
690 J., Couger, M. B., Eccles, D., Li, B., Lieber, M., Macmanes, M. D., Ott, M., Orvis, J., Pochet,  
691 N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., Henschel, R., Leduc, R. D.,  
692 Friedman, N., and Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq  
693 using the Trinity platform for reference generation and analysis. *Nature protocols*, 8(8):1494–512.
- 694 [Harbison, 1985] Harbison, G. R. (1985). On the classification and evolution of the Ctenophora. In  
695 Conway Morris, S. C., George, J. D., Gibson, R., and Platt, H. M., editors, *The Origin and Rela-*  
696 *tionships of Lower Invertebrates*, pages 78–100. Clarendon Press, Oxford.



- 697 [Hoff and Stanke, 2013] Hoff, K. J. and Stanke, M. (2013). WebAUGUSTUS—a web service for training  
698 AUGUSTUS and predicting genes in eukaryotes. *Nucleic Acids Research*, 41(W1):W123–W128.
- 699 [Huang et al., 2012] Huang, S., Chen, Z., Huang, G., Yu, T., Yang, P., Li, J., Fu, Y., Yuan, S., Chen,  
700 S., and Xu, A. (2012). HaploMerger: Reconstructing allelic relationships for polymorphic diploid  
701 genome assemblies. *Genome Research*, 22(8):1581–1588.
- 702 [Jarvis et al., 2014] Jarvis, E. D., Mirarab, S., Aberer, A. J., Li, B., Houde, P., Li, C., Ho, S. Y. W.,  
703 Faircloth, B. C., Nabholz, B., Howard, J. T., Suh, A., Weber, C. C., da Fonseca, R. R., Li, J., Zhang,  
704 F., Li, H., Zhou, L., Narula, N., Liu, L., Ganapathy, G., Boussau, B., Bayzid, M. S., Zavidovych, V.,  
705 Subramanian, S., Gabaldon, T., Capella-Gutierrez, S., Huerta-Cepas, J., Rekepalli, B., Munch, K.,  
706 Schierup, M., Lindow, B., Warren, W. C., Ray, D., Green, R. E., Bruford, M. W., Zhan, X., Dixon,  
707 A., Li, S., Li, N., Huang, Y., Derryberry, E. P., Bertelsen, M. F., Sheldon, F. H., Brumfield, R. T.,  
708 Mello, C. V., Lovell, P. V., Wirthlin, M., Schneider, M. P. C., Prosdocimi, F., Samaniego, J. A.,  
709 Velazquez, A. M. V., Alfaro-Nunez, A., Campos, P. F., Petersen, B., Sicheritz-Ponten, T., Pas, A.,  
710 Bailey, T., Scofield, P., Bunce, M., Lambert, D. M., Zhou, Q., Perelman, P., Driskell, A. C., Shapiro,  
711 B., Xiong, Z., Zeng, Y., Liu, S., Li, Z., Liu, B., Wu, K., Xiao, J., Yinqi, X., Zheng, Q., Zhang, Y.,  
712 Yang, H., Wang, J., Smeds, L., Rheindt, F. E., Braun, M., Fjeldsa, J., Orlando, L., Barker, F. K.,  
713 Jonsson, K. A., Johnson, W., Koepfli, K.-P., O’Brien, S., Haussler, D., Ryder, O. A., Rahbek, C.,  
714 Willerslev, E., Graves, G. R., Glenn, T. C., McCormack, J., Burt, D., Ellegren, H., Alstrom, P.,  
715 Edwards, S. V., Stamatakis, A., Mindell, D. P., Cracraft, J., Braun, E. L., Warnow, T., Jun, W.,  
716 Gilbert, M. T. P., and Zhang, G. (2014). Whole-genome analyses resolve early branches in the tree  
717 of life of modern birds. *Science*, 346(6215):1320–1331.
- 718 [Jékely et al., 2015] Jékely, G., Paps, J., and Nielsen, C. (2015). The phylogenetic position of  
719 ctenophores and the origin(s) of nervous systems. *EvoDevo*, 6(1):1.
- 720 [Kato and Standley, 2013] Kato, K. and Standley, D. M. (2013). MAFFT multiple sequence align-  
721 ment software version 7: improvements in performance and usability. *Molecular biology and evolu-  
722 tion*, 30(4):772–80.
- 723 [Kim et al., 2013] Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L.  
724 (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and  
725 gene fusions. *Genome biology*, 14(4):R36.
- 726 [King et al., 2008] King, N., Westbrook, M. J., Young, S. L., Kuo, A., Abedin, M., Chapman, J.,  
727 Fairclough, S., Hellsten, U., Isogai, Y., Letunic, I., Marr, M., Pincus, D., Putnam, N., Rokas, A.,  
728 Wright, K. J., Zuzow, R., Dirks, W., Good, M., Goodstein, D., Lemons, D., Li, W., Lyons, J. B.,  
729 Morris, A., Nichols, S., Richter, D. J., Salamov, A., Sequencing, J. G. I., Bork, P., Lim, W. a.,  
730 Manning, G., Miller, W. T., McGinnis, W., Shapiro, H., Tjian, R., Grigoriev, I. V., and Rokhsar,  
731 D. (2008). The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans.  
732 *Nature*, 451(7180):783–8.
- 733 [Krishnan et al., 2014] Krishnan, A., Dnyansagar, R., Almén, M. S., Williams, M. J., Fredriksson,  
734 R., Manoj, N., and Schiöth, H. B. (2014). The GPCR repertoire in the demosponge *Amphimedon*  
735 *queenslandica*: insights into the GPCR system at the early divergence of animals. *BMC Evolutionary*  
736 *Biology*, 14(1):270.
- 737 [Krishnan and Schiöth, 2015] Krishnan, A. and Schiöth, H. B. (2015). The role of G protein-coupled  
738 receptors in the early evolution of neurotransmission and the nervous system. *The Journal of*  
739 *experimental biology*, 218(Pt 4):562–571.
- 740 [Langmead and Salzberg, 2012] Langmead, B. and Salzberg, S. L. (2012). Fast gapped-read alignment  
741 with Bowtie 2.
- 742 [Leys, 2015] Leys, S. P. (2015). Elements of a ‘nervous system’ in sponges. *The Journal of experimen-  
743 tal biology*, 218(Pt 4):581–91.
- 744 [Leys et al., 1999] Leys, S. P., Mackie, G. O., and Meech, R. W. (1999). Impulse conduction in a  
745 sponge. *The Journal of experimental biology*, 202 (Pt 9)(June 1997):1139–1150.
- 746 [Leys et al., 2007] Leys, S. P., Mackie, G. O., and Reiswig, H. M. (2007). The Biology of Glass Sponges.  
747 *Advances in Marine Biology*, 52(06):1–145.



- 748 [Li et al., 2015] Li, X., Liu, H., Chu Luo, J., Rhodes, S. a., Trigg, L. M., van Rossum, D. B., Anishkin,  
749 A., Diatta, F. H., Sassic, J. K., Simmons, D. K., Kamel, B., Medina, M., Martindale, M. Q., and  
750 Jegla, T. (2015). Major diversification of voltage-gated  $K^{+}$  channels occurred in ancestral  
751 parahoxozoans. *Proceedings of the National Academy of Sciences*, page 201422941.
- 752 [Liebeskind et al., 2011] Liebeskind, B. J., Hillis, D. M., and Zakon, H. H. (2011). Evolution of sodium  
753 channels predates the origin of nervous systems in animals. *Proceedings of the National Academy of  
754 Sciences of the United States of America*, 108(22):9154–9159.
- 755 [Ludeman et al., 2014] Ludeman, D. A., Farrar, N., Riesgo, A., Paps, J., and Leys, S. P. (2014).  
756 Evolutionary origins of sensation in metazoans: functional evidence for a new sensory organ in  
757 sponges. *BMC Evolutionary Biology*, 14(3):1–11.
- 758 [Luo et al., 2012] Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q.,  
759 Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C., Cheung,  
760 D. W., Yiu, S.-M., Peng, S., Xiaoqian, Z., Liu, G., Liao, X., Li, Y., Yang, H., Wang, J., Lam, T.-W.,  
761 and Wang, J. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo  
762 assembler. *GigaScience*, 1(1):18.
- 763 [Marçais and Kingsford, 2011] Marçais, G. and Kingsford, C. (2011). A fast, lock-free approach for  
764 efficient parallel counting of occurrences of k-mers. *Bioinformatics*, 27(6):764–770.
- 765 [Moran et al., 2015] Moran, Y., Barzilai, M. G., Liebeskind, B. J., and Zakon, H. H. (2015). Evolu-  
766 tion of voltage-gated ion channels at the emergence of Metazoa. *Journal of Experimental Biology*,  
767 218:515–525.
- 768 [Moran et al., 2014] Moran, Y., Fredman, D., Praher, D., Li, X. Z., Wee, L. M., Rentzsch, F., Zamore,  
769 P. D., Technau, U., and Seitz, H. (2014). Cnidarian microRNAs frequently regulate targets by  
770 cleavage. *Genome Research*, 24(4):651–663.
- 771 [Moroz, 2015] Moroz, L. L. (2015). Convergent evolution of neural systems in ctenophores. *Journal  
772 of Experimental Biology*, 218:598–611.
- 773 [Moroz et al., 2014] Moroz, L. L., Kocot, K. M., Citarella, M. R., Dosung, S., Norekian, T. P., Po-  
774 volotskaya, I. S., Grigorenko, A. P., Dailey, C., Berezikov, E., Buckley, K. M., Ptitsyn, A., Reshetov,  
775 D., Mukherjee, K., Moroz, T. P., Bobkova, Y., Yu, F., Kapitonov, V. V., Jurka, J., Bobkov, Y. V.,  
776 Swore, J. J., Girardo, D. O., Fodor, A., Gusev, F., Sanford, R., Bruders, R., Kittler, E., Mills, C. E.,  
777 Rast, J. P., Derelle, R., Solovyev, V. V., Kondrashov, F. a., Swalla, B. J., Sweedler, J. V., Rogaev,  
778 E. I., Halanych, K. M., and Kohn, A. B. (2014). The ctenophore genome and the evolutionary  
779 origins of neural systems. *Nature*, 17:1–123.
- 780 [Nickel, 2010] Nickel, M. (2010). Evolutionary emergence of synaptic nervous systems: what can we  
781 learn from the non-synaptic, nerveless Porifera? *Invertebrate Biology*, 129(1):1–16.
- 782 [Nosenko et al., 2013] Nosenko, T., Schreiber, F., Adamska, M., Adamski, M., Eitel, M., Hammel,  
783 J., Maldonado, M., Müller, W. E. G., Nickel, M., Schierwater, B., Vacelet, J., Wiens, M., and  
784 Wörheide, G. (2013). Deep metazoan phylogeny: when different genes tell different stories. *Molecular  
785 phylogenetics and evolution*, 67(1):223–33.
- 786 [Pérez-Porro et al., 2013] Pérez-Porro, a. R., Navarro-Gómez, D., Uriz, M. J., and Giribet, G. (2013).  
787 A NGS approach to the encrusting Mediterranean sponge *Crella elegans* (Porifera, Demospongiae,  
788 Poecilosclerida): transcriptome sequencing, characterization and overview of the gene expression  
789 along three life cycle stages. *Molecular ecology resources*, 454:494–509.
- 790 [Pertea et al., 2015] Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and  
791 Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq  
792 reads. *Nature Biotechnology*, 33(3).
- 793 [Petersen et al., 2011] Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP  
794 4.0: discriminating signal peptides from transmembrane regions. *Nature methods*, 8(10):785–6.
- 795 [Philippe et al., 2011] Philippe, H., Brinkmann, H., Lavrov, D. V., Littlewood, D. T. J., Manuel,  
796 M., Wörheide, G., and Baurain, D. (2011). Resolving difficult phylogenetic questions: why more  
797 sequences are not enough. *PLoS biology*, 9(3):e1000602.

- 798 [Philippe et al., 2009] Philippe, H., Derelle, R., Lopez, P., Pick, K., Borchiellini, C., Boury-Esnault,  
799 N., Vacelet, J., Renard, E., Houliston, E., Quéinnec, E., Da Silva, C., Wincker, P., Le Guyader, H.,  
800 Leys, S., Jackson, D. J., Schreiber, F., Erpenbeck, D., Morgenstern, B., Wörheide, G., and Manuel,  
801 M. (2009). Phylogenomics revives traditional views on deep animal relationships. *Current biology :  
802 CB*, 19(8):706–12.
- 803 [Pick et al., 2010] Pick, K. S., Philippe, H., Schreiber, F., Erpenbeck, D., Jackson, D. J., Wrede, P.,  
804 Wiens, M., Alié, A., Morgenstern, B., Manuel, M., and Wörheide, G. (2010). Improved phylogenomic  
805 taxon sampling noticeably affects nonbilaterian relationships. *Molecular Biology and Evolution*,  
806 27(9):1983–1987.
- 807 [Pisani et al., 2015] Pisani, D., Pett, W., Dohrmann, M., Feuda, R., Rota-Stabelli, O., Philippe, H.,  
808 Lartillot, N., and Wörheide, G. (2015). Genomic data do not support comb jellies as the sister group  
809 to all other animals. *Proceedings of the National Academy of Sciences*, 112(50):201518127.
- 810 [Ponce et al., 2016] Ponce, D., Brinkman, D. L., Potriquet, J., and Mulvenna, J. (2016). Tentacle tran-  
811 scriptome and venom proteome of the pacific sea nettle, *Chrysaora fuscescens* (Cnidaria: Scyphozoa).  
812 *Toxins*, 8(4).
- 813 [Pratlong et al., 2015] Pratlong, M., Haguenaue, A., Chabrol, O., Klopp, C., Pontarotti, P., and  
814 Aurelle, D. (2015). The red coral ( *Corallium rubrum* ) transcriptome: a new resource for population  
815 genetics and local adaptation studies. *Molecular Ecology Resources*, 15(5):1205–1215.
- 816 [Price et al., 2010] Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 - Approximately  
817 maximum-likelihood trees for large alignments. *PLoS ONE*, 5(3).
- 818 [Putnam et al., 2008] Putnam, N. H., Butts, T., Ferrier, D. E. K., Furlong, R. F., Hellsten, U.,  
819 Kawashima, T., Robinson-Rechavi, M., Shoguchi, E., Terry, A., Yu, J.-K., Benito-Gutiérrez, E. L.,  
820 Dubchak, I., Garcia-Fernández, J., Gibson-Brown, J. J., Grigoriev, I. V., Horton, A. C., de Jong,  
821 P. J., Jurka, J., Kapitonov, V. V., Kohara, Y., Kuroki, Y., Lindquist, E., Lucas, S., Osoegawa,  
822 K., Pennacchio, L. a., Salamov, A. a., Satou, Y., Sauka-Spengler, T., Schmutz, J., Shin-I, T., Toy-  
823 oda, A., Bronner-Fraser, M., Fujiyama, A., Holland, L. Z., Holland, P. W. H., Satoh, N., and  
824 Rokhsar, D. S. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature*,  
825 453(7198):1064–71.
- 826 [Qiu et al., 2015] Qiu, F., Ding, S., Ou, H., Wang, D., Chen, J., and Miyamoto, M. M. (2015). Tran-  
827 scriptome changes during the life cycle of the red sponge, *Mycale phyllophila* (Porifera, Demospon-  
828 giae, Poecilosclerida). *Genes*, 6(4):1023–1052.
- 829 [Ramoino et al., 2010] Ramoino, P., Ledda, F. D., Ferrando, S., Gallus, L., Bianchini, P., Diaspro, A.,  
830 Fato, M., Tagliaferro, G., and Manconi, R. (2010). Metabotropic ??-aminobutyric acid (GABAB)  
831 receptors modulate feeding behavior in the calcisponge *Leucandra aspera*. *Journal of Experimental  
832 Zoology Part A: Ecological Genetics and Physiology*, 313A(3):132–140.
- 833 [Riesgo et al., 2014a] Riesgo, A., Farrar, N., Windsor, P. J., Giribet, G., and Leys, S. P. (2014a). The  
834 Analysis of Eight Transcriptomes from All Poriferan Classes Reveals Surprising Genetic Complexity  
835 in Sponges. *Molecular biology and evolution*.
- 836 [Riesgo et al., 2014b] Riesgo, A., Peterson, K., Richardson, C., Heist, T., Strehlow, B., McCauley,  
837 M., Cotman, C., Hill, M., and Hill, A. (2014b). Transcriptomic analysis of differential host gene  
838 expression upon uptake of symbionts: a case study with *Symbiodinium* and the major bioeroding  
839 sponge *Cliona varians*. *BMC genomics*, 15(1):376.
- 840 [Ryan et al., 2010] Ryan, J. F., Pang, K., Mullikin, J. C., Martindale, M. Q., and Baxevanis,  
841 A. D. (2010). The homeodomain complement of the ctenophore *Mnemiopsis leidyi* suggests that  
842 Ctenophora and Porifera diverged prior to the ParaHoxozoa. *EvoDevo*, 1(1):9.
- 843 [Ryan et al., 2013] Ryan, J. F., Pang, K., Schnitzler, C. E., a. D. Nguyen, A.-d., Moreland, R. T.,  
844 Simmons, D. K., Koch, B. J., Francis, W. R., Havlak, P., Smith, S. a., Putnam, N. H., Haddock,  
845 S. H. D., Dunn, C. W., Wolfsberg, T. G., Mullikin, J. C., Martindale, M. Q., Baxevanis, A. D.,  
846 Comparative, N., and Program, S. (2013). The Genome of the Ctenophore *Mnemiopsis leidyi* and  
847 Its Implications for Cell Type Evolution. *Science*, 342(6164):1242592–1242592.

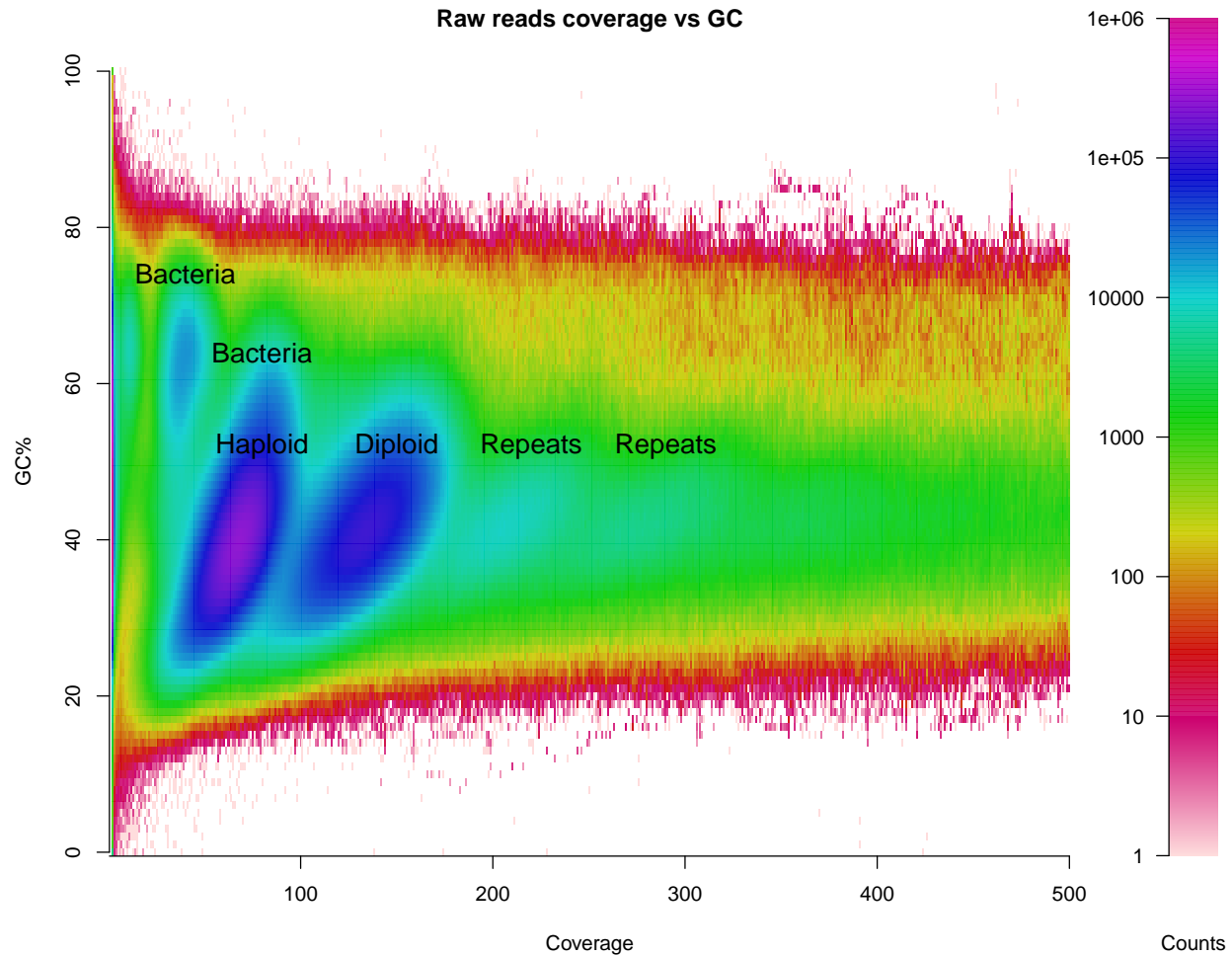
- 848 [Ryu et al., 2016] Ryu, T., Seridi, L., Moitinho-Silva, L., Oates, M., Liew, Y. J., Mavromatis, C.,  
849 Wang, X., Haywood, A., Lafi, F. F., Kupresanin, M., Sougrat, R., Alzahrani, M. A., Giles, E.,  
850 Ghosheh, Y., Schunter, C., Baumgarten, S., Berumen, M. L., Gao, X., Aranda, M., Foret, S.,  
851 Gough, J., Voolstra, C. R., Hentschel, U., and Ravasi, T. (2016). Hologenome analysis of two  
852 marine sponges with different microbiomes. *BMC genomics*, 17(1):158.
- 853 [Sáez et al., 2009] Sáez, A. G., Lozano, E., and Zaldívar-Riverón, A. (2009). Evolutionary history of  
854 Na,K-ATPases and their osmoregulatory role. *Genetica*, 136(3):479–490.
- 855 [Sahlin et al., 2014] Sahlin, K., Vezzi, F., Nystedt, B., Lundeberg, J., and Arvestad, L. (2014). BESST  
856 - Efficient scaffolding of large fragmented assemblies. *BMC Bioinformatics*, 15(1):281.
- 857 [Sara et al., 2001] Sara, M., Sara, A., Nickel, M., and Brümmer, F. (2001). Three New Species  
858 of Tethya (Porifera: Demospongia) from German Aquaria. *Stuttgarter Beiträge zur Naturkunde*,  
859 631(15S):1–16.
- 860 [Schierwater et al., 2009] Schierwater, B., Kolokotronis, S., Eitel, M., and DeSalle, R. (2009). The  
861 Diploblast-Bilateria Sister hypothesis. *Communicative & Integrative Biology*, 2(5):1–3.
- 862 [Schuler et al., 2015] Schuler, a., Schmitz, G., Reft, A., Ozbek, S., Thurm, U., and Bornberg-Bauer,  
863 E. (2015). The rise and fall of TRP-N, an ancient family of mechanogated ion channels, in Metazoa.  
864 *Genome Biology and Evolution*, 7(6):1–27.
- 865 [Simakov et al., 2015] Simakov, O., Kawashima, T., Marlétaz, F., Jenkins, J., Koyanagi, R., Mitros,  
866 T., Hisata, K., Bredeson, J., Shoguchi, E., Gyoja, F., Yue, J.-X., Chen, Y.-C., Freeman, R. M.,  
867 Sasaki, A., Hikosaka-Katayama, T., Sato, A., Fujie, M., Baughman, K. W., Levine, J., Gonzalez,  
868 P., Cameron, C., Fritzenwanker, J. H., Pani, A. M., Goto, H., Kanda, M., Arakaki, N., Yamasaki,  
869 S., Qu, J., Cree, A., Ding, Y., Dinh, H. H., Dugan, S., Holder, M., Jhangiani, S. N., Kovar, C. L.,  
870 Lee, S. L., Lewis, L. R., Morton, D., Nazareth, L. V., Okwuonu, G., Santibanez, J., Chen, R.,  
871 Richards, S., Muzny, D. M., Gillis, A., Peshkin, L., Wu, M., Humphreys, T., Su, Y.-H., Putnam,  
872 N. H., Schmutz, J., Fujiyama, A., Yu, J.-K., Tagawa, K., Worley, K. C., Gibbs, R. A., Kirschner,  
873 M. W., Lowe, C. J., Satoh, N., Rokhsar, D. S., and Gerhart, J. (2015). Hemichordate genomes and  
874 deuterostome origins. *Nature*, pages 1–19.
- 875 [Simakov et al., 2013] Simakov, O., Marletaz, F., Cho, S.-J., Edsinger-Gonzales, E., Havlak, P., Hell-  
876 sten, U., Kuo, D.-H., Larsson, T., Lv, J., Arendt, D., Savage, R., Osoegawa, K., de Jong, P.,  
877 Grimwood, J., Chapman, J. a., Shapiro, H., Aerts, A., Otilar, R. P., Terry, A. Y., Boore, J. L.,  
878 Grigoriev, I. V., Lindberg, D. R., Seaver, E. C., Weisblat, D. a., Putnam, N. H., and Rokhsar, D. S.  
879 (2013). Insights into bilaterian evolution from three spiralian genomes. *Nature*, 493(7433):526–31.
- 880 [Simão et al., 2015] Simão, F. A., Waterhouse, R. M., Ioannidis, P., and Kriventseva, E. V. (2015).  
881 BUSCO : assessing genome assembly and annotation completeness with single-copy orthologs.  
882 *Genome analysis*, 31(June):9–10.
- 883 [Simion et al., 2017] Simion, P., Philippe, H., Baurain, D., Jager, M., Richter, D. J., Di Franco, A.,  
884 Roure, B., Satoh, N., Quéinnec, É., Ereskovsky, A., Lapébie, P., Corre, E., Delsuc, F., King, N.,  
885 Wörheide, G., and Manuel, M. (2017). A Large and Consistent Phylogenomic Dataset Supports  
886 Sponges as the Sister Group to All Other Animals. *Current Biology*, pages 1–10.
- 887 [Smith et al., 2011] Smith, S. M. E., Morgan, D., Musset, B., Cherny, V. V., Place, A. R., Hastings,  
888 J. W., and DeCoursey, T. E. (2011). Voltage-gated proton channel in a dinoflagellate. *Proceedings*  
889 *of the National Academy of Sciences*, 108(44):18162–18167.
- 890 [Sreedharan et al., 2010] Sreedharan, S., Shaik, J. H. A., Olszewski, P. K., Levine, A. S., Schiöth, H. B.,  
891 and Fredriksson, R. (2010). Glutamate, aspartate and nucleotide transporters in the SLC17 family  
892 form four main phylogenetic clusters: evolution and tissue expression. *BMC genomics*, 11(iii):17.
- 893 [Srivastava et al., 2008] Srivastava, M., Begovic, E., Chapman, J., Putnam, N. H., Hellsten, U.,  
894 Kawashima, T., Kuo, A., Mitros, T., Salamov, A., Carpenter, M. L., Signorovitch, A. Y., Moreno,  
895 M. a., Kamm, K., Grimwood, J., Schmutz, J., Shapiro, H., Grigoriev, I. V., Buss, L. W., Schier-  
896 water, B., Dellaporta, S. L., and Rokhsar, D. S. (2008). The Trichoplax genome and the nature of  
897 placozoans. *Nature*, 454(7207):955–60.

- 898 [Srivastava et al., 2010] Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M. E. a.,  
899 Mitros, T., Richards, G. S., Conaco, C., Dacre, M., Hellsten, U., Larroux, C., Putnam, N. H.,  
900 Stanke, M., Adamska, M., Darling, A., Degnan, S. M., Oakley, T. H., Plachetzki, D. C., Zhai, Y.,  
901 Adamski, M., Calcino, A., Cummins, S. F., Goodstein, D. M., Harris, C., Jackson, D. J., Leys, S. P.,  
902 Shu, S., Woodcroft, B. J., Vervoort, M., Kosik, K. S., Manning, G., Degnan, B. M., and Rokhsar,  
903 D. S. (2010). The Amphimedon queenslandica genome and the evolution of animal complexity.  
904 *Nature*, 466(7307):720–6.
- 905 [Stamatakis, 2014] Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and  
906 post-analysis of large phylogenies. *Bioinformatics*, 30(9):1312–1313.
- 907 [Stanke et al., 2008] Stanke, M., Diekhans, M., Baertsch, R., and Haussler, D. (2008). Using native and  
908 syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics*, 24(5):637–  
909 644.
- 910 [Stein, 1995] Stein, W. D. (1995). The sodium pump in the evolution of animal cells. *Philosophical*  
911 *transactions of the Royal Society of London. Series B, Biological sciences*, 349(1329):263–9.
- 912 [Strous et al., 2012] Strous, M., Kraft, B., Bisdorf, R., and Tegetmeyer, H. E. (2012). The bin-  
913 ning of metagenomic contigs for microbial physiology of mixed cultures. *Frontiers in Microbiology*,  
914 3(DEC):1–11.
- 915 [Suga et al., 2013] Suga, H., Chen, Z., de Mendoza, A., Sebé-Pedrós, A., Brown, M. W., Kramer, E.,  
916 Carr, M., Kerner, P., Vervoort, M., Sánchez-Pons, N., Torruella, G., Derelle, R., Manning, G., Lang,  
917 B. F., Russ, C., Haas, B. J., Roger, A. J., Nusbaum, C., and Ruiz-Trillo, I. (2013). The Capsaspora  
918 genome reveals a complex unicellular prehistory of animals. *Nature communications*, 4:2325.
- 919 [Versluis et al., 2015] Versluis, D., D’Andrea, M. M., Ramiro Garcia, J., Leimena, M. M., Hugenholtz,  
920 F., Zhang, J., Öztürk, B., Nylund, L., Sipkema, D., van Schaik, W., de Vos, W. M., Kleerebezem,  
921 M., Smidt, H., and van Passel, M. W. J. (2015). Mining microbial metatranscriptomes for expression  
922 of antibiotic resistance genes under natural conditions. *Scientific reports*, 5(January):11981.
- 923 [Whelan et al., 2015] Whelan, N. V., Kocot, K. M., Moroz, L. L., and Halanych, K. M. (2015). Error  
924 , signal , and the placement of Ctenophora sister to all other animals. *Proceedings of the National*  
925 *Academy of Sciences*, 112(18):1–6.
- 926 [Wray et al., 2015] Wray, G. A., Smith, A., Peterson, K., Donoghue, P., Benton, M., Thackray, J.,  
927 Budd, G., Marshall, C., Shu, D., Isozaki, Y., Zhang, X., Han, J., Maruyama, S., Walcott, C., Knoll,  
928 A., Walter, M., Narbonne, G., Christie-Blick, N., Raymond, P., Cloud, P., Budd, G., Jensen, S.,  
929 Briggs, D., Fortey, R., Morris, S. C., Seilacher, A., Bose, P., Pfluger, F., Rasmussen, B., Bengtson,  
930 S., Fletcher, I., McNaughton, N., Pecoits, E., Konhauser, K., Aubet, N., Heaman, L., Veroslavsky,  
931 G., Stern, R., Gingras, M., Hultgren, T., Cunningham, J., Yin, C., Stampanoni, M., Marone, F.,  
932 Donoghue, P., Bengtson, S., Gaucher, C., Poire, D., Bossi, J., Bettucci, L., Beri, A., Fedonkin,  
933 M., Simonetta, A., Ivantsov, A., Sprigg, R., Glaessner, M., Gehling, J., Seilacher, A., Retallack,  
934 G., Narbonne, G., Seilacher, A., Grazhdankin, D., Legouta, A., Li, C., Chen, J., Hua, T., Maloof,  
935 A., Tang, F., Bengtson, S., Wang, Y., Wang, X., Yin, C., Yin, Z., Grant, S., Grotzinger, J.,  
936 Watters, W., Knoll, A., Fedonkin, M., Vickers-Rich, P., Swalla, B., Trusler, P., Hall, M., Xiao, S.,  
937 Zhang, Y., Knoll, A., Chen, J.-Y., Oliveri, P., Li, C., Zhou, G., Gao, F., Hagadorn, J., Peterson,  
938 K., Davidson, E., Hagadorn, J., Chen, L., Xiao, S., Pang, K., Zhou, C., Yuan, X., Bengtson, S.,  
939 Budd, G., Cunningham, J., Margoliash, E., Brown, R., Richardson, M., Boulter, D., Ramshaw,  
940 J., Jefferies, R., Runnegar, B., Knoll, A., Carroll, S., Wray, G., Levinton, J., Shapiro, L., Aris-  
941 Brosou, S., Yang, Z., Bromham, L., Rambaut, A., Fortey, R., Cooper, A., Penny, D., Welch, J.,  
942 Fontanillas, E., Bromham, L., Wheat, C., Wahlberg, N., Schulte, J., Erwin, D., Laflamme, M.,  
943 Tweedt, S., Sperling, E., Pisani, D., Peterson, K., Filipowski, A., Murillo, O., Freydenzon, A., Tamura,  
944 K., Kumar, S., Tamura, K., Battistuzzi, F., Billing-Ross, P., Murillo, O., Filipowski, A., Kumar,  
945 S., Hug, L., Roger, A., Ho, S., Phillips, M., Sanderson, M., Thorne, J., Kishino, H., Painter, I.,  
946 Sanderson, M., Britton, T., Anderson, C., Jacquet, D., Lundqvist, S., Bremer, K., Shaul, S., Graur,  
947 D., Battistuzzi, F., Billing-Ross, P., Murillo, O., Filipowski, A., Kumar, S., Mello, B., Schrago, C.,  
948 Rambaut, A., Bromham, L., Kishino, H., Thorne, J., Bruno, W., Douzery, E., Snell, E., Baptiste,  
949 E., Delsuc, F., Philippe, H., Battistuzzi, F., Filipowski, A., Hedges, S., Kumar, S., Schwartz, R.,

- 950 Mueller, R., Ho, S., Phillips, M., Drummond, A., Cooper, A., Blair, J., Hedges, S., Warnock, R.,  
951 Yang, Z., Donoghue, P., Battistuzzi, F., Billing-Ross, P., Murillo, O., Filipinski, A., Kumar, S., Ayala,  
952 F., Rzhetsky, A., Ayala, F., Peterson, K., Lyons, J., Nowak, K., Takacs, C., Wargo, M., McPeck,  
953 M., Cutler, D., Chernikova, D., Motamedi, S., Csuros, M., Koonin, E., Rogozin, I., Richter, D.,  
954 King, N., Dunn, C., Giribet, G., Edgecombe, G., Hejnol, A., Emes, R., Grant, S., Burkhardt, P.,  
955 Hejnol, A., Martindale, M., Balavoine, G., Adoutte, A., Hirth, F., Miller, D., Ball, E., Northcutt,  
956 R., Strausfeld, N., Hirth, F., Tosches, M., Arendt, D., Lyons, T., Reinhard, C., Planavsky, N.,  
957 Planavsky, N., Sperling, E., Frieder, C., Raman, A., Girguis, P., Levin, L., Knoll, A., Stanley, S.,  
958 Peterson, K., Cotton, J., Gehling, J., Pisani, D., Butterfield, N., Benito-Gutierrez, E., Arendt, D.,  
959 Holland, L., Carvalho, J., Escriva, H., Laudet, V., Schubert, M., Shimeld, S., Yu, J.-K., Turner,  
960 S., Young, J., Pisani, D., Poling, L., Lyons-Weiler, M., Hedges, S., Otsuka, J., Sugaya, N., Hedges,  
961 S., Blair, J., Venturi, M., Shoe, J., and Gu, X. (2015). Molecular clocks and the early evolution  
962 of metazoan nervous systems. *Philosophical transactions of the Royal Society of London. Series B,*  
963 *Biological sciences*, 370(1684):424–431.
- 964 [Wu and Watanabe, 2005] Wu, T. D. and Watanabe, C. K. (2005). GMAP: a genomic mapping and  
965 alignment program for mRNA and EST sequences. *Bioinformatics*, 21(9):1859–1875.
- 966 [Zakon, 2012] Zakon, H. H. (2012). Adaptive evolution of voltage-gated sodium channels: the first 800  
967 million years. *Proceedings of the National Academy of Sciences of the United States of America*, 109  
968 Suppl(Supplement\_1):10619–25.
- 969 [Zapata et al., 2015] Zapata, F., Goetz, F. E., Smith, S. A., Howison, M., Siebert, S., Church, S. H.,  
970 Sanders, S. M., Ames, C. L., McFadden, C. S., France, S. C., Daly, M., Collins, A. G., Haddock,  
971 S. H. D., Dunn, C. W., and Cartwright, P. (2015). Phylogenomic Analyses Support Traditional  
972 Relationships within Cnidaria. *Plos One*, 10(10):e0139068.



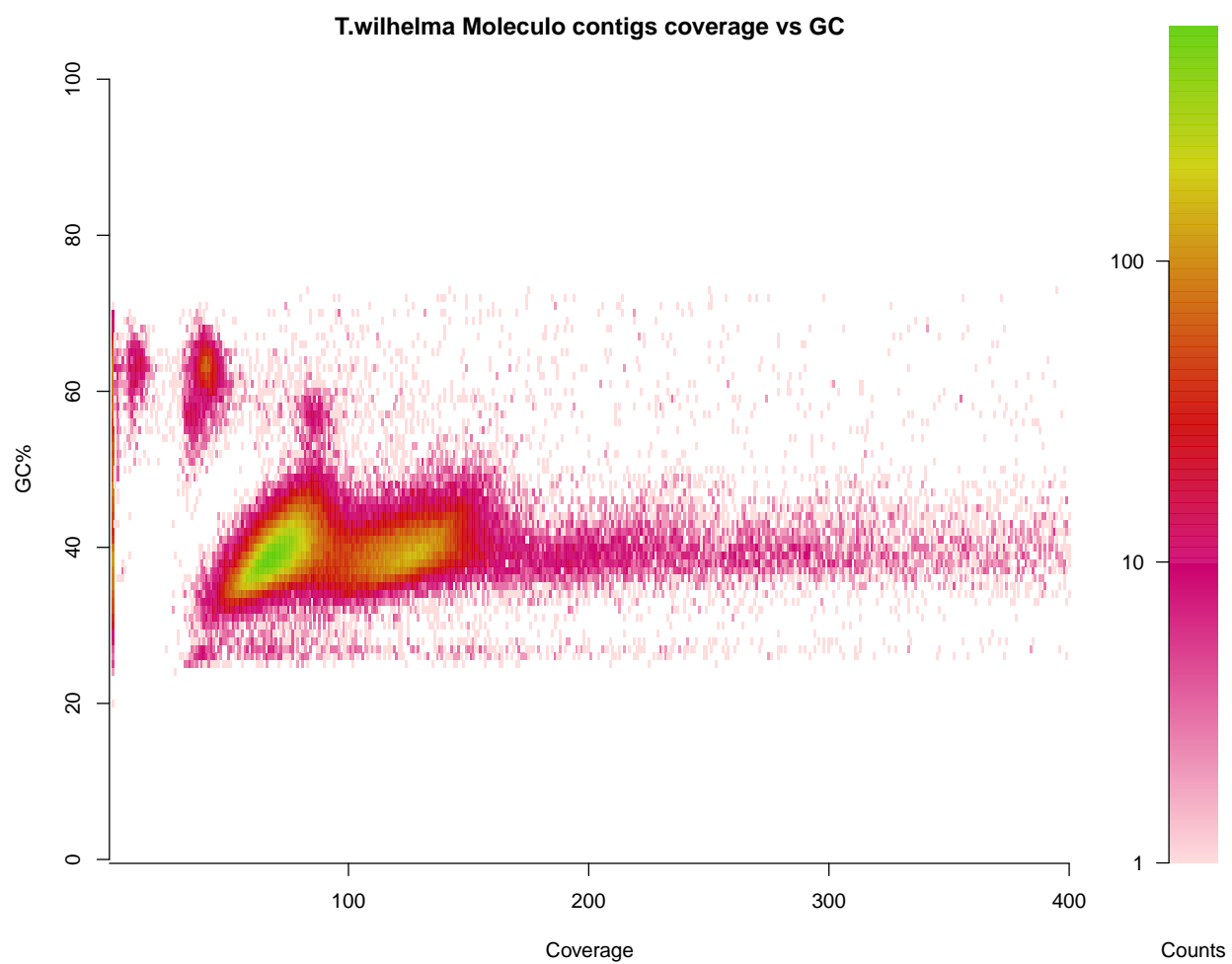
## 973 1 Supplemental Figures



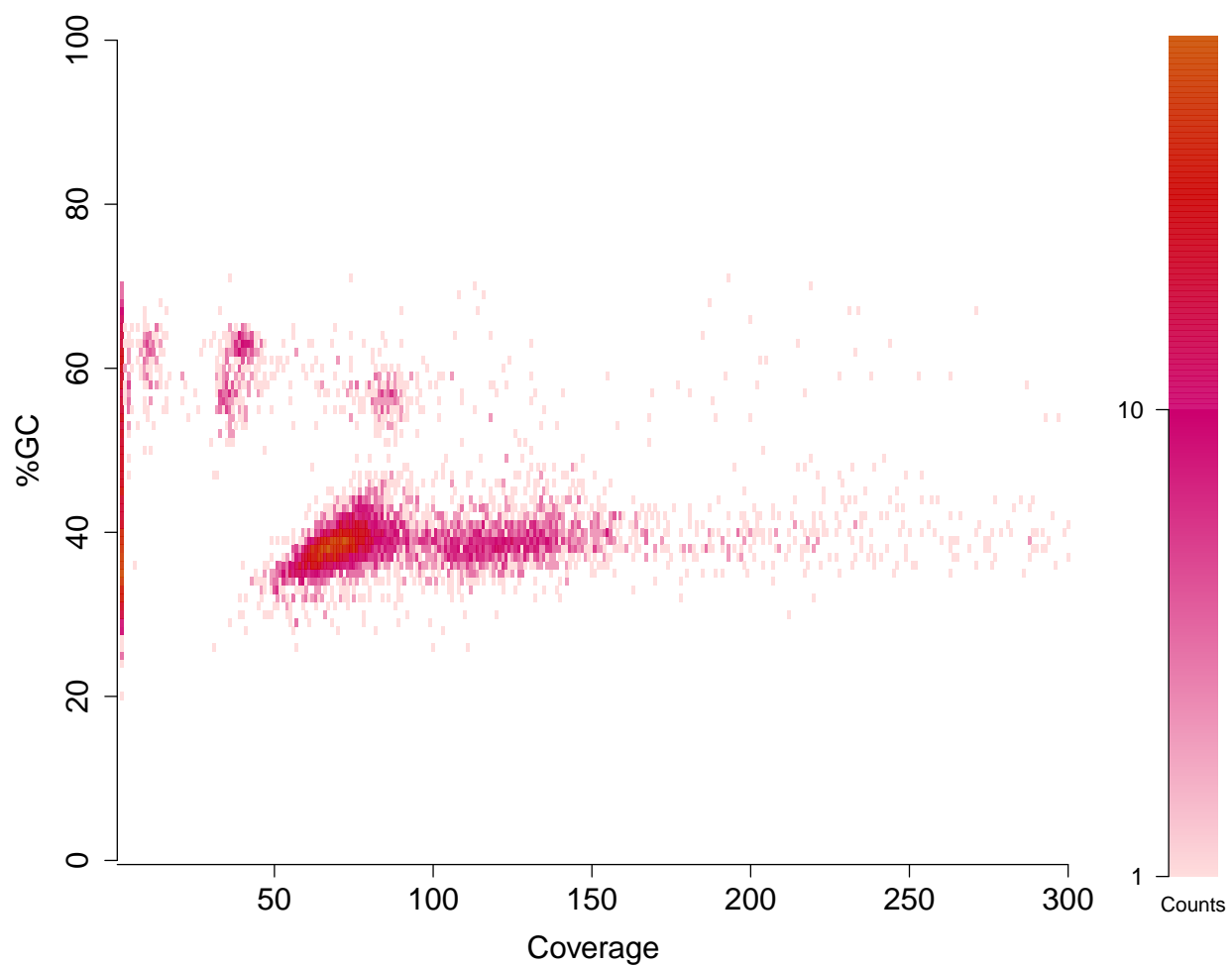
**Supplemental Figure 1: Coverage vs. GC content for reads**

Lava lamp plot of the unfiltered paired-end reads. Coverage was calculated as the median 31-mer coverage for each read. High-GC reads indicate the presence of bacteria in the raw reads.



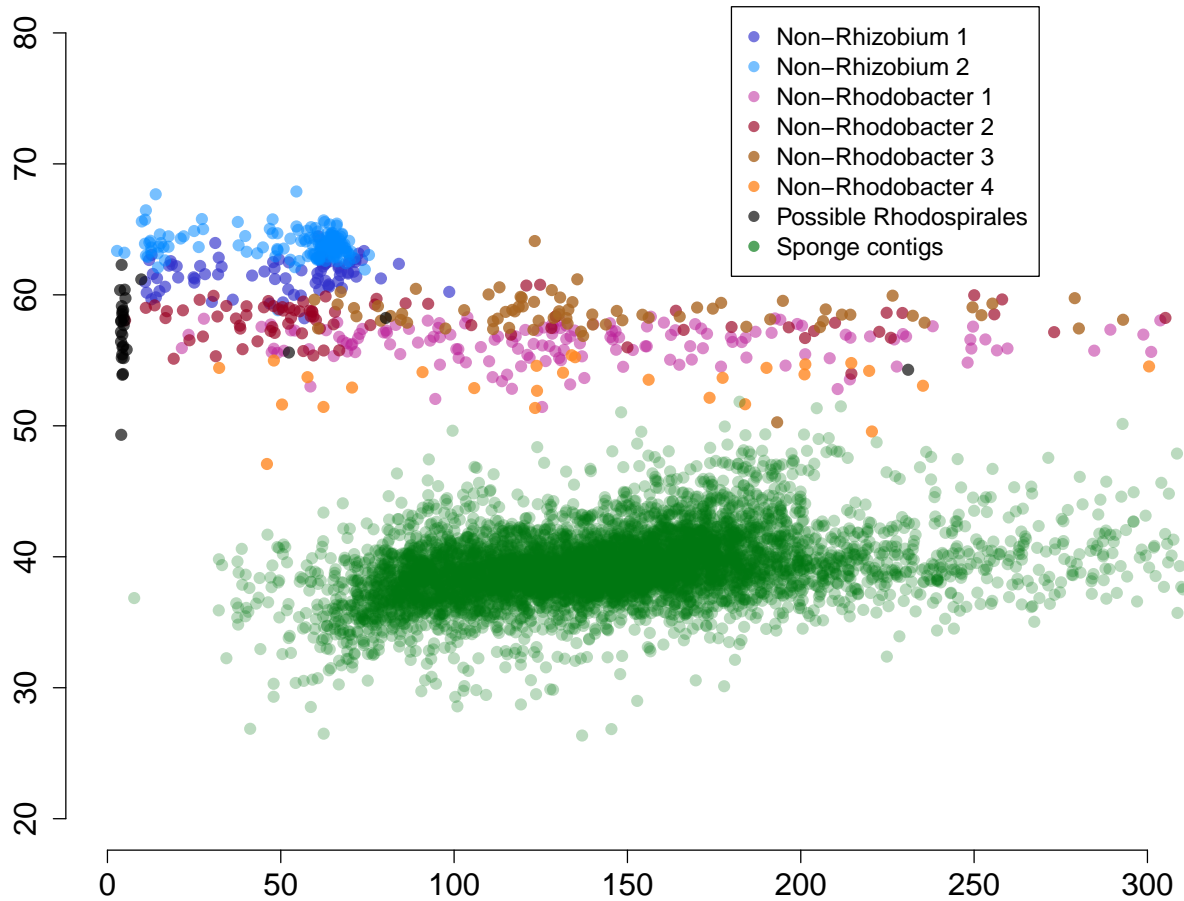


**Supplemental Figure 2: Coverage vs. GC content for genomic scaffolds**  
Heat map of percent GC versus coverage of reads for the all Moleculo reads.



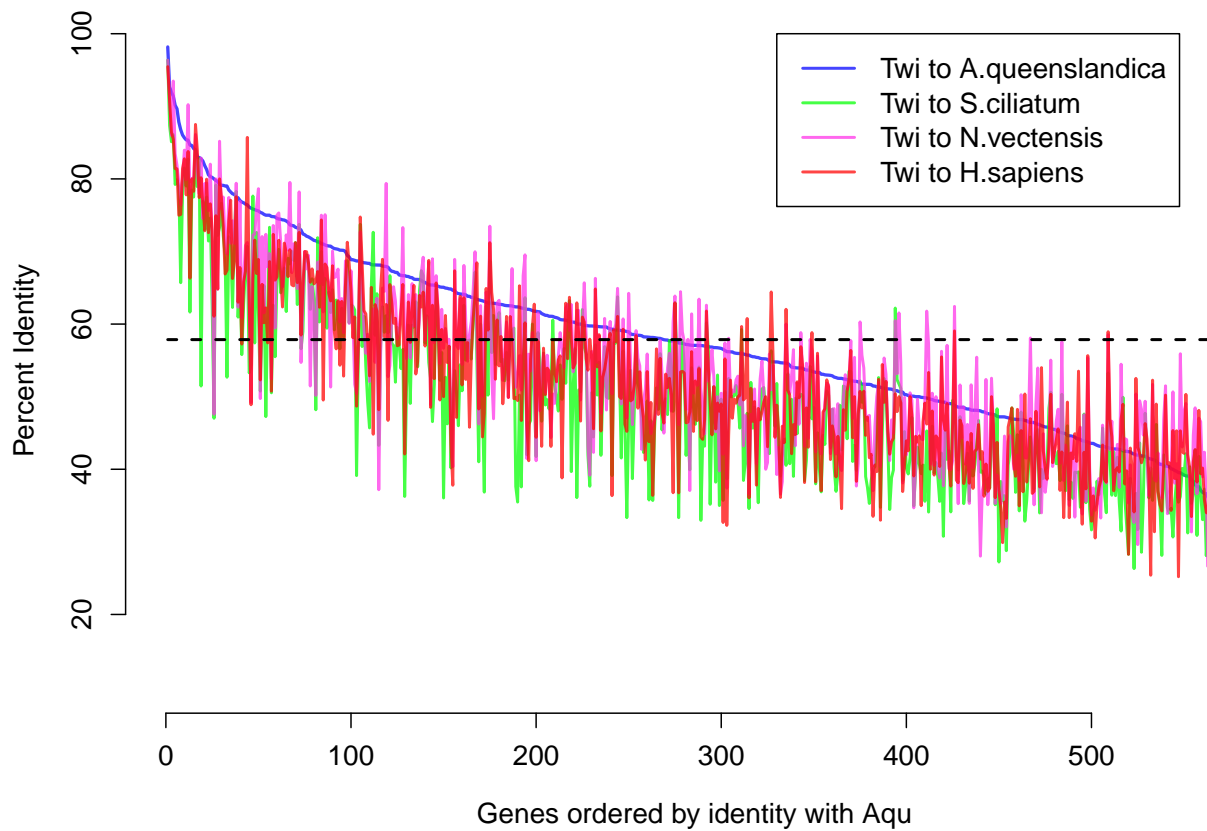
**Supplemental Figure 3: Coverage vs. GC content for genomic scaffolds**

Scatterplot of percent GC versus coverage of reads for the all scaffolds. The 1,040 contigs with zero coverage are carried over from low-coverage Moleculo reads, and likely derive from amplified contaminating DNA.



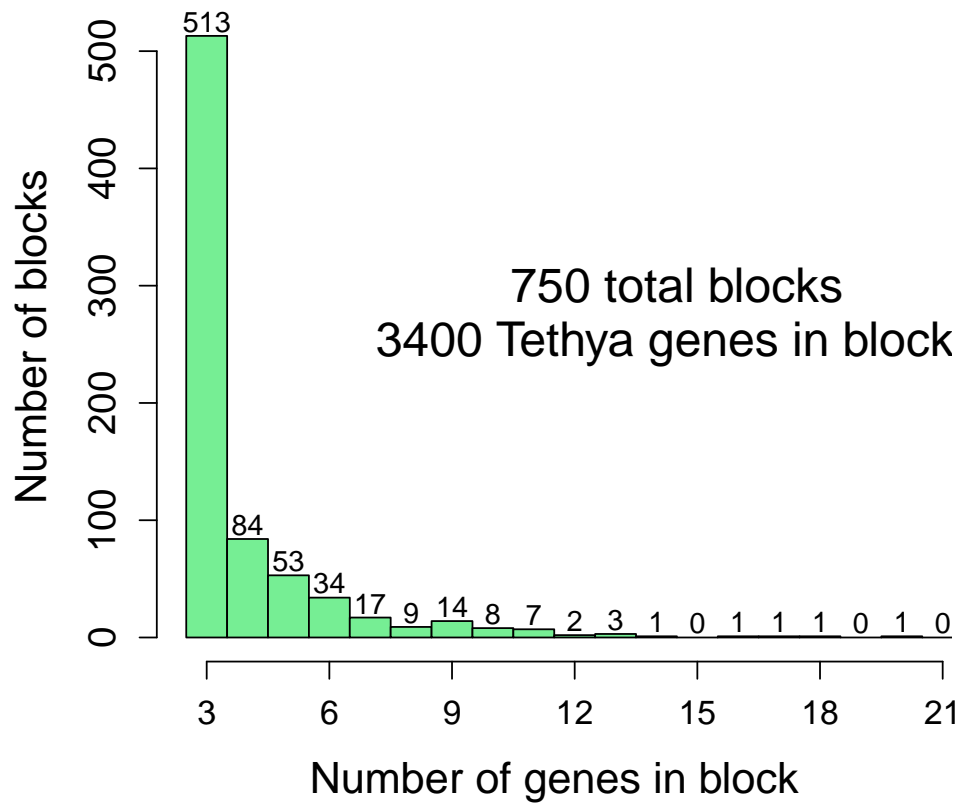
**Supplemental Figure 4: Separation of contigs derived from bacterial symbionts**

Sponge scaffolds are in green, while scaffolds assigned to bacteria are identified by blue (*Rhizobiales*) and pink (*Rhodobacter*). Low-coverage contigs were removed. Seven bins were identified with MetaWatt to separate the bacterial contigs, though several bins appeared to correspond to the same bacteria.



**Supplemental Figure 5: Percent identity between sponges**

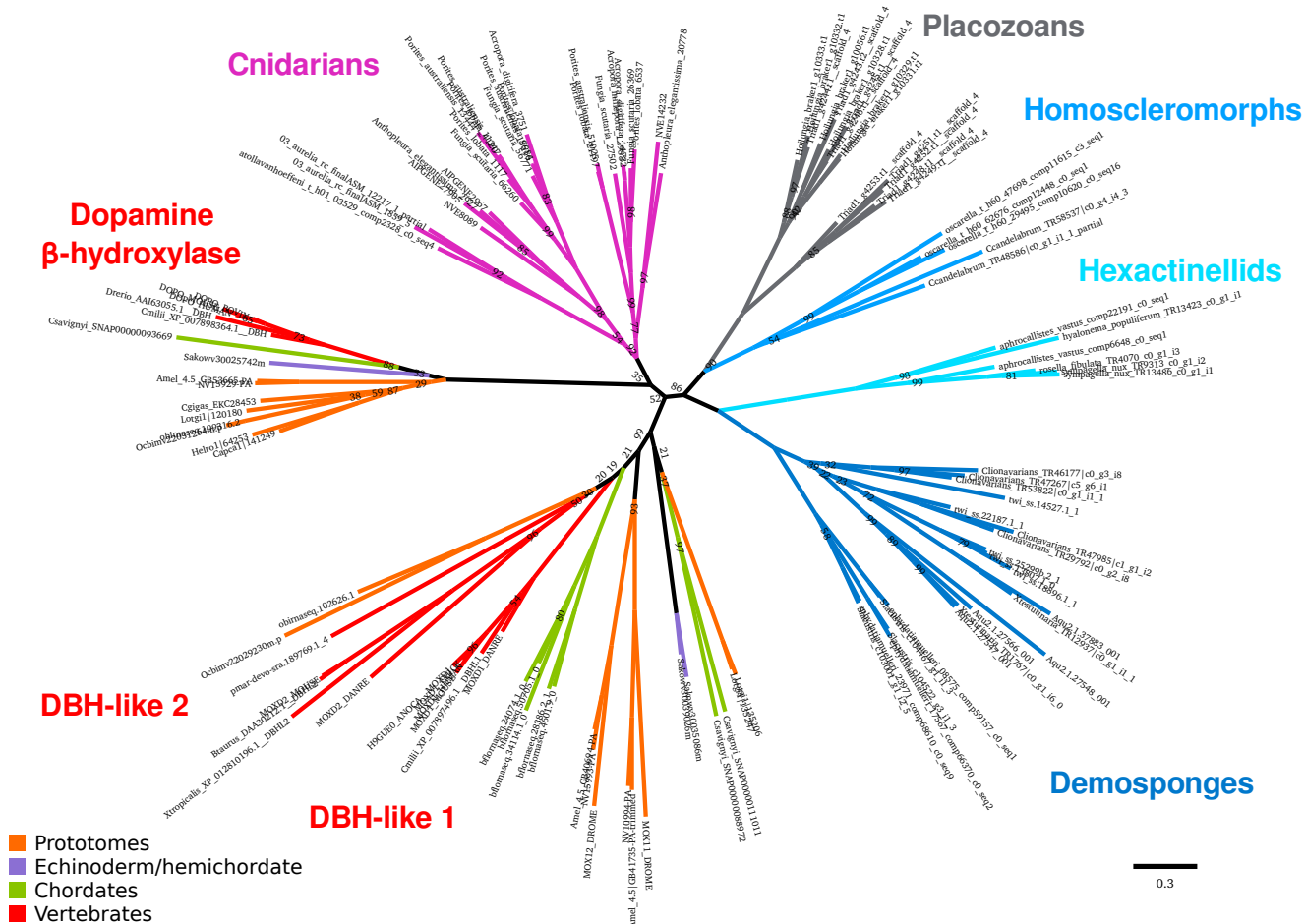
Calculated protein percent identity between 570 one-to-one orthologs between *T. wilhelma* and *A. queenslandica* (blue), *S. ciliatum* (green), the anemone *N. vectensis* (purple) and human (red). Average identity between *T. wilhelma* and *A. queenslandica* is shown as the dotted black line at 57%.



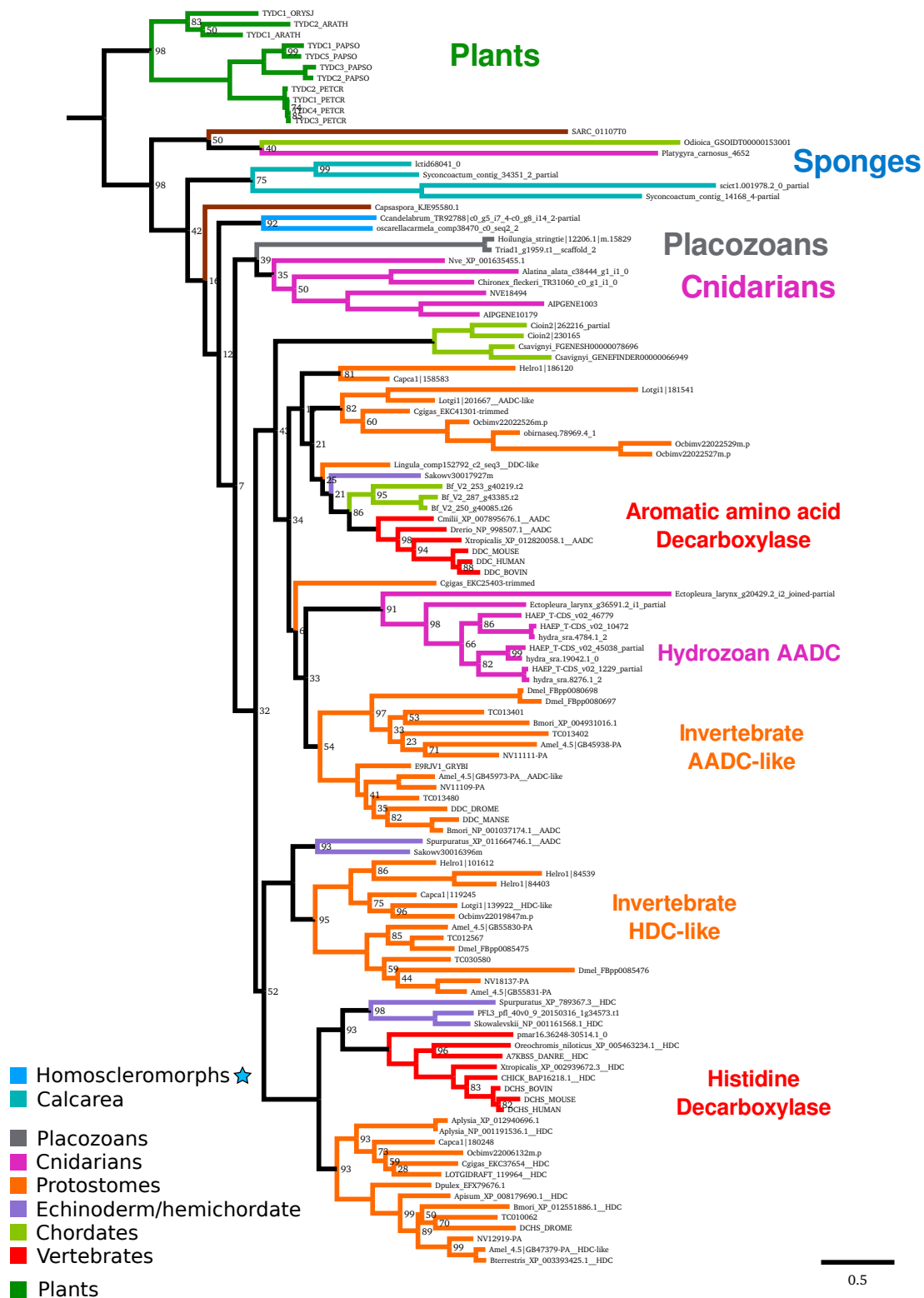
**Supplemental Figure 6: Length of microsyntenic blocks**

Histogram of number of genes in detected microsyntenic blocks between *T. wilhelma* and *A. queenslandica* v2.0 gene models.

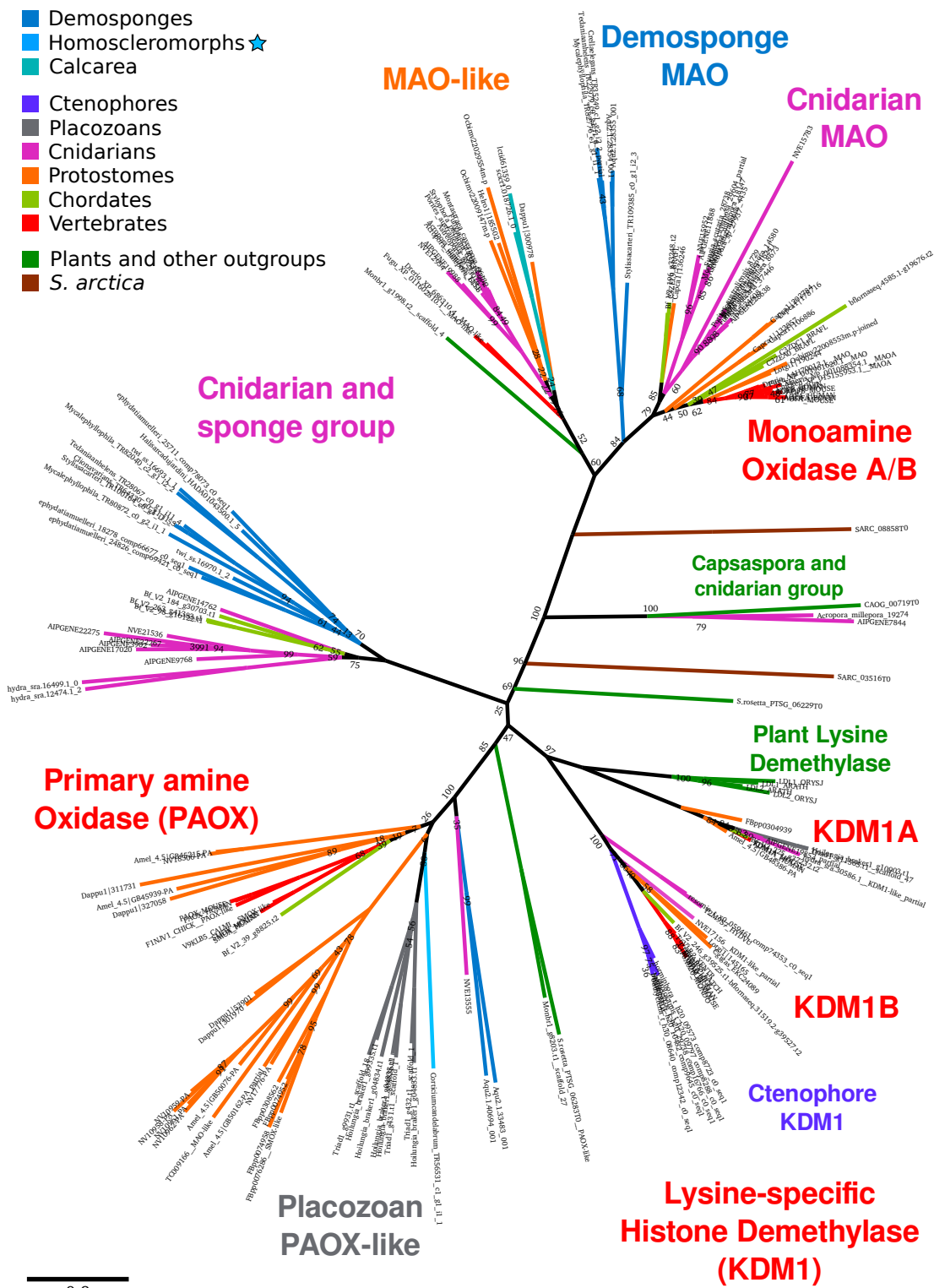




**Supplemental Figure 7: Dopamine  $\beta$ -hydroxylase homologs across metazoans**  
 Tree of DBH and DBH-like proteins across all metazoan groups, generated with RAXML using the PROTGAMMALG model. Bootstrap values are 100 unless otherwise shown.

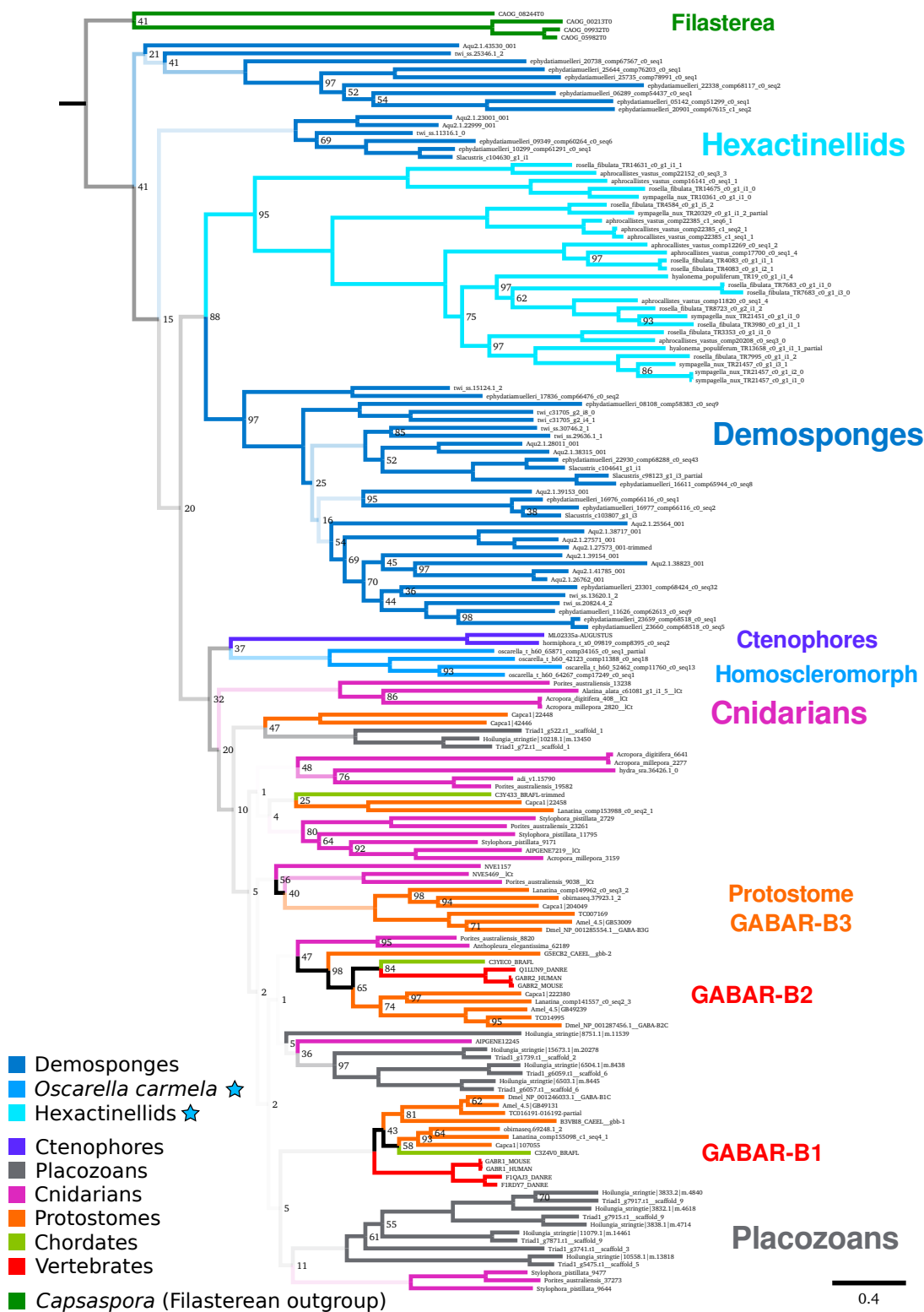


**Supplemental Figure 8: Aromatic amino acid decarboxylase homologs across metazoans**  
 Tree of AADC and histidine decarboxylase (HDC) proteins across all metazoan groups, generated with RAXML using the PROTGAMMALG model. Bootstrap values are 100 unless otherwise shown.



**Supplemental Figure 9: Monoamine oxidase homologs across metazoans**

Tree of monoamine oxidase (MAO) and related proteins across all metazoan groups, generated with RAxML using the PROTGAMMALG model. Searches for lysine demethylase (KDM1) and primary amine oxidase (PAOX) were not exhaustive, and were added to display the sole positions of ctenophores (only have KDM1) and placozoans (only have PAOX) in this protein family. Bootstrap values are 100 unless otherwise shown.



**Supplemental Figure 10: mGABA receptors across metazoans**

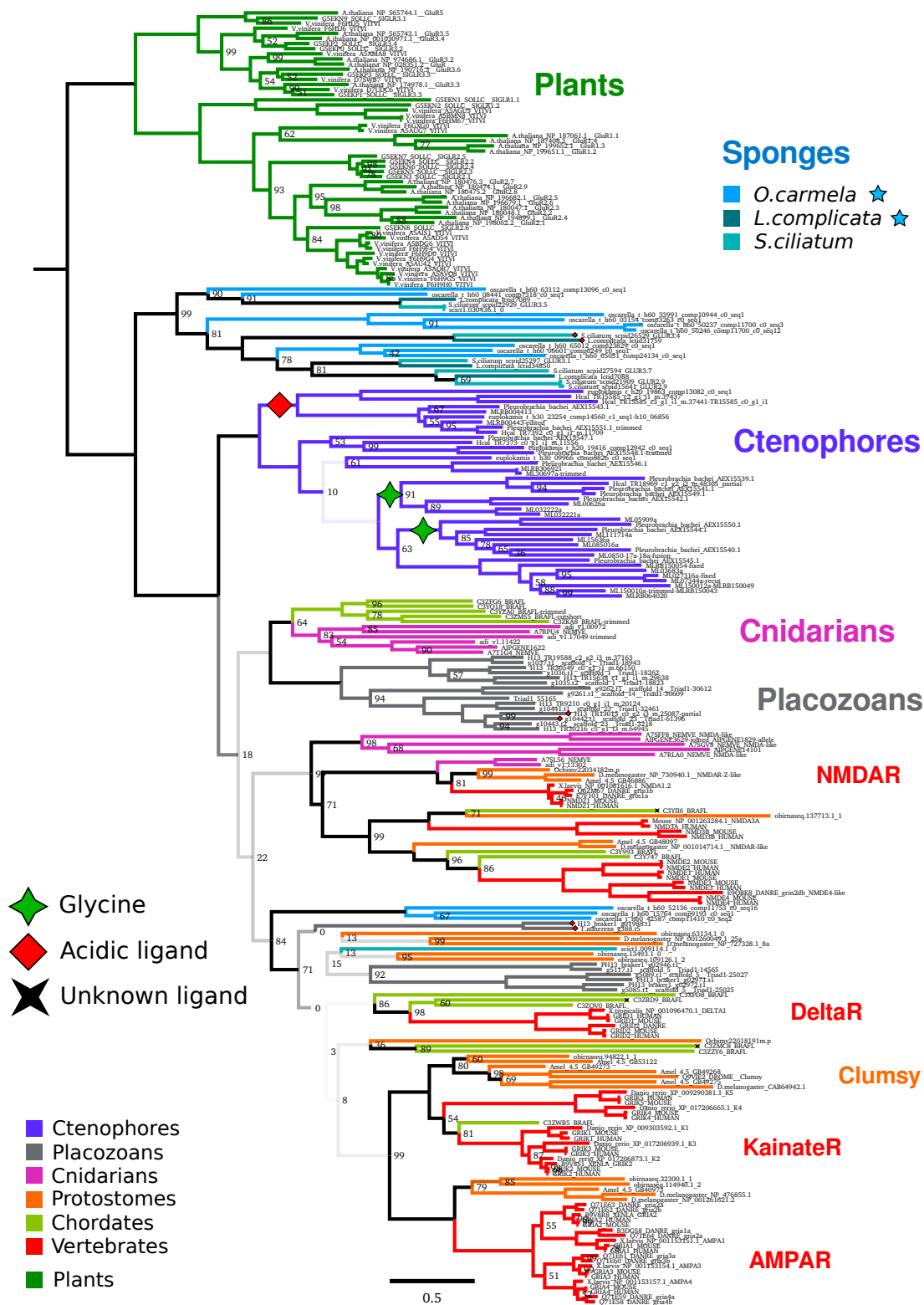
Complete version of Figure 4 metabotropic GABA receptor (GABA-B type) protein tree generated with RAxML. Bootstrap values are 100 unless otherwise shown. The majority of deeper nodes were poorly resolved; branch transparency corresponds to bootstrap support for values under 50, meaning half-transparent.

	246	268	287	367	395	446
<b>GABR1_HUMAN</b>	- PGC	- YGS	FRTP	GLFY	LIG	GGFO
<b>GABR1_MOUSE</b>	- PGC	- YGS	FRTP	GLFY	LIG	GGFO
<b>F1QA3_DANRE</b>	- PGC	- YGS	FRTP	GLFY	LIG	GGFO
<i>Dmel</i> _NP_001246033.1_GABA-B1C	- AGC	- YGA	FRTP	GLFY	FIG	EGYQ
<i>Amel</i> _4.5 GB49131	- AGC	- YGA	FRTP	GLFY	FIG	EGYQ
<i>C3Z4V0_BRAFL</i>	- TG	- YGS	FRTP	GVFY	FIG	PGPY
<i>B3VB18_CAEEI_gbb-1</i>	- TG	- YGS	FRTP	GLFY	FIG	GGFP
<b>GABR2_HUMAN</b>	GGVC	- FAA	FRTV	GOF	IPG	GPSK
<b>GABR2_MOUSE</b>	GGVC	- FAA	FRTV	GOF	IPG	GPSK
<i>O1LUN9_DANRE</i>	GGVC	- FAA	FRTV	GOF	IPG	ETSK
<i>C3VEC0_BRAFL</i>	GGT	- YG	FRTV	GLF	IMG	RTMP
<i>Dmel</i> _NP_001287456.1_GABA-B2C	GAA	- YAD	FRV	GNF	IMG	ETSR
<i>Amel</i> _4.5 G849239	GAA	- YAD	FRV	GNF	IMG	ETSR
<i>HoiIungia_stringtie</i>  15673.1 m.20278	GSGY	- YGA	LRT	ASF	FLG	DYSY
<i>Triad1_g1739.t1_scaffold_2</i>	GSGY	- YGA	LRT	ASF	FLG	ESYV
<i>HoiIungia_stringtie</i>  6503.1 m.8445	GAGY	- YGA	FRT	ASF	LIG	QOD
<i>Triad1_g6057.t1_scaffold_6</i>	GAGY	- YGA	FRT	ASF	LIG	KRD
<i>HoiIungia_stringtie</i>  6504.1 m.8438	GADF	- YGS	YRT	ANFY	LLG	QPLG
<i>G5ECB2_CAEEI_gbb-2</i>	GGC	- YAE	FRV	VDV	LPG	PNN
<i>AIPGENE12245</i>	GAGY	- YGS	YRT	GNF	LLG	RRDD
<i>HoiIungia_stringtie</i>  10558.1 m.13818	GAON	- HGS	FRT	AIG	LVG	TMT
<i>Triad1_g5475.t1_scaffold_5</i>	GAON	- HGS	FRT	AIG	LVG	TMT
<i>HoiIungia_stringtie</i>  11079.1 m.14461	GPV	- NAA	FRT	VYA	LLG	YLAS
<i>HoiIungia_stringtie</i>  3832.1 m.4618	GPV	- PLS	MRT	LFAY	LPG	PLLG
<i>HoiIungia_stringtie</i>  3833.2 m.4840	GPV	- HTA	FRT	LFAY	MPG	PLYG
<i>Triad1_g7917.t1_scaffold_9</i>	GPV	- HTA	FRT	LFAY	LPG	PLYG
<i>HoiIungia_stringtie</i>  3838.1 m.4714	GPL	- PSA	FRT	LNAY	LPG	HLSG
<i>Triad1_g3741.t1_scaffold_3</i>	GPL	- SGV	FRT	LFAY	ITF	PFKY
<i>Capca1 42446</i>	AAAN	- YAS	FRT	TIC	MPG	HATN
<i>Capca1 22448</i>	GGG	- YGC	FRT	TAS	FAN	TKDF
<i>HoiIungia_stringtie</i>  10218.1 m.13450	GAA	- YAS	FRT	AGF	IPG	GAID
<i>Triad1_g72.t1_scaffold_1</i>	GAA	- YAS	FRT	GGF	ITF	GAID
<i>Triad1_g522.t1_scaffold_1</i>	GSA	- YAS	FRT	TGF	WFT	GGKK
<i>Capca1 22458</i>	GGG	- YGA	FRT	YNS	IPG	GPNP
<i>Lanatina_comp153988_c0_seq2_1</i>	GGG	- YGA	WRN	FDG	YMG	GFND
<i>AIPGENE7219</i>	PGG	- MAN	VRT	GMF	LLM	ENPD
<i>C3Y433_BRAFL-trimmed</i>	PGG	- YAG	FRT	ALM	FMG	KSTV
<i>HoiIungia_stringtie</i>  8751.1 m.11539	PGG	- YAG	FRT	LSS	FMG	GGWK
<i>Dmel</i> _NP_001285554.1_GABA-B3G	GSA	- YGS	YRT	GSA	LHE	AISQ
<i>Amel</i> _4.5 G853009	GSA	- YGS	YRT	GSA	LHE	TYSK
<i>Lanatina_comp149962_c0_seq3_2</i>	GSS	- HAS	YRL	GSA	LVG	TRS
<i>NVE5469/23-200</i>	GPAC	- YAS	YRT	GMF	LLD	TANL
<i>twi</i> _ss.15124.1_2/34-214	GAD	- PLS	LRV	LVS	TFA	ALFG
<i>ephydatiamuelleri_17836_comp66476_c0_seq2</i>	GAD	- PLS	LRV	VNT	YFA	SLDD
<i>oscarella_t_h60_42123_comp11388_c0_seq18</i>	AEG	- FLN	YLL	GLFY	FTG	DFQY
<i>oscarella_t_h60_52462_comp11760_c0_seq13</i>	AAG	- FLN	HQI	ALL	WYH	DFTY
<i>oscarella_t_h60_64267_comp17249_c0_seq1</i>	AEG	- FLN	HQL	GLFY	WFO	DFHY
<i>Aqu2.1.25564_001</i>	GGG	- YYS	FTL	LNC	TLG	DDHS
<i>Aqu2.1.26762_001</i>	GGG	- YSS	FQI	LNM	WYS	QPTY
<i>Aqu2.1.41785_001</i>	GGG	- FSS	FQI	LNM	WYN	LSSY
<i>Aqu2.1.39154_001</i>	GAG	- YSS	FRT	LNM	WYQ	YPI
<i>twi</i> _ss.13620.1_2	GCG	- YGS	FRT	INT	WYQ	SQTY
<i>twi</i> _ss.20824.4_2	GCG	- YGT	FRAN	LNS	WYA	ILD
<i>ephydatiamuelleri_23659_comp68518_c0_seq1</i>	GCG	- YFS	FRT	LNS	WYQ	EASD
<i>ephydatiamuelleri_23660_comp68518_c0_seq8</i>	GCG	- YFS	FRT	VNS	WYQ	EASD
<i>ephydatiamuelleri_11626_comp62613_c0_seq9</i>	GCG	- YVQ	FRT	LNS	WYN	EKSE
<i>ephydatiamuelleri_23301_comp68424_c0_seq32</i>	GCG	- PEP	FRT	LNM	WYQ	TNG
<i>Aqu2.1.27571_001</i>	GCG	- AIA	FRT	INT	WYN	TELG
<i>Aqu2.1.27573_001-trimmed</i>	GCG	- YAT	FRT	INT	WYN	TELG
<i>Aqu2.1.38717_001</i>	GCG	- FAA	YRL	INAY	WYA	TELG
<i>Aqu2.1.28011_001</i>	GSG	- CAS	FOM	LAM	WYT	SYL
<i>Aqu2.1.38315_001</i>	GAG	- CAS	FOM	LAM	WYT	RYST
<i>twi</i> _ss.29636.1_1	GAG	- CVS	FOM	LAM	WYT	THYD
<i>ephydatiamuelleri_22930_comp68288_c0_seq43</i>	GGG	- CAS	FOM	LAM	WYT	SSSQ
<i>ephydatiamuelleri_16611_comp65944_c0_seq8</i>	GSG	- HSS	FOM	LAV	WYT	DPNR
<i>twi</i> _ss.30746.2_1	GSG	- HSS	FOM	LAV	WYT	SSSQ
<i>Aqu2.1.38823_001</i>	GGG	- YST	VQV	LNT	WYN	TSM
<i>Aqu2.1.39153_001</i>	GAG	- YAS	FRT	LNM	WYP	SESV
<i>ephydatiamuelleri_16977_comp66116_c0_seq2</i>	GCD	- YD	FRT	LSM	WYP	FQTS
<i>ephydatiamuelleri_16976_comp66116_c0_seq1</i>	GGG	- YD	IRT	LNM	WYP	FQTS
<i>twi</i> _c31705_g2_i4_1	GAG	- HSS	LRA	INA	WYT	LVDE
<i>twi</i> _c31705_g2_i8_0	GAG	- HSS	FRT	LNVC	WYT	AADE
<i>ephydatiamuelleri_08108_comp58383_c0_seq9</i>	APG	- YGF	FRT	INA	WNT	SITE
<i>Aqu2.1.23001_001/19-196</i>	GGG	- YAS	FRV	GFF	WYT	QDFL
<i>twi</i> _ss.11316.1_0/38-229	DGG	- LSA	FRT	GYF	WYT	EPHY
<i>ephydatiamuelleri_10299_comp61291_c0_seq1</i>	GGG	- FGS	FRT	IWM	WYS	AVSK
<i>ephydatiamuelleri_09349_comp60264_c0_seq6</i>	DGG	- FGS	YRT	AWM	WYT	NMHS
<i>hydra_sra.36426.1_0/39-213</i>	GPP	- YTE	FQI	ALF	WYT	KCDN
<i>hormiphora_t_x0_09819_comp8395_c0_seq2</i>	GPV	- FTA	LRL	ANF	SLP	KEDK
<i>ML02335a-AUGUSTUS</i>	GPV	- FTA	LRL	ANF	SLP	EIEK
<i>Aqu2.1.43530_001</i>	GPP	- YAS	FRT	VYS	WYT	NDL
<i>twi</i> _ss.25346.1_2	GPG	- YG	YQT	AF	WYS	GPPD
<i>ephydatiamuelleri_25735_comp78991_c0_seq1</i>	GPG	- VSA	LRT	GFL	WYS	IDLL
<i>ephydatiamuelleri_25644_comp76203_c0_seq1</i>	GLA	- VSA	PRP	AF	WYS	SDLY
<i>ephydatiamuelleri_06289_comp54437_c0_seq1</i>	GPL	- VHL	GIS	-	WYS	VDS
<i>ephydatiamuelleri_05142_comp51299_c0_seq1</i>	GPV	- VHM	GIS	VN	WYS	SINK
<i>ephydatiamuelleri_20901_comp67615_c1_seq2</i>	GLS	- AHS	VG	AF	WYS	TSRL
<i>ephydatiamuelleri_22338_comp68117_c0_seq2</i>	GPS	- LHM	FGM	LLA	WYS	ASPP
<i>ephydatiamuelleri_20738_comp67567_c0_seq1</i>	GLP	- ISS	YRV	G	WYS	NDM
<i>CAOG_0021370</i>	QAL	- IIV	FYN	YPK	WYT	SCRF
<i>CAOG_0598270</i>	QAL	- IIV	FAG	YPM	WYT	YCRF
<i>CAOG_0993270</i>	QAL	- IIV	FAG	YPM	WYT	YCRF

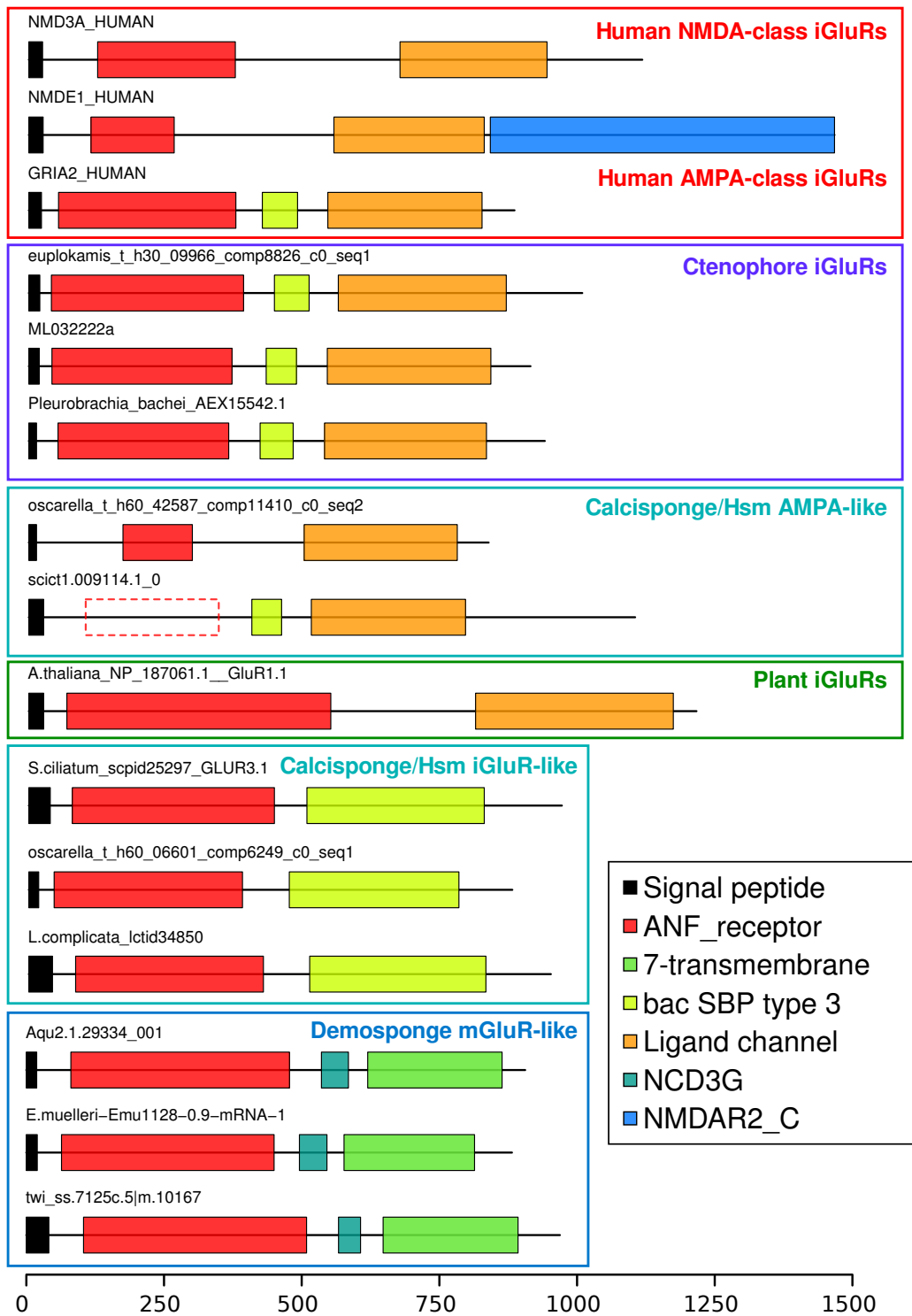
### Supplemental Figure 11: Binding pocket alignment of mGABARs across metazoans

Select residues involved in the binding of GABA are highlighted, and numbers correspond to the position in the human protein GABR1/GABAR-B1, based on the structure of GABR1 [Geng et al., 2013]. Highly conserved residues not thought to be involved in binding are highlighted in blue. The two human proteins are highlighted in pink. Placozoan proteins are highlighted in gray. Sponge proteins are highlighted in blue. The two ctenophore proteins are highlighted in violet.



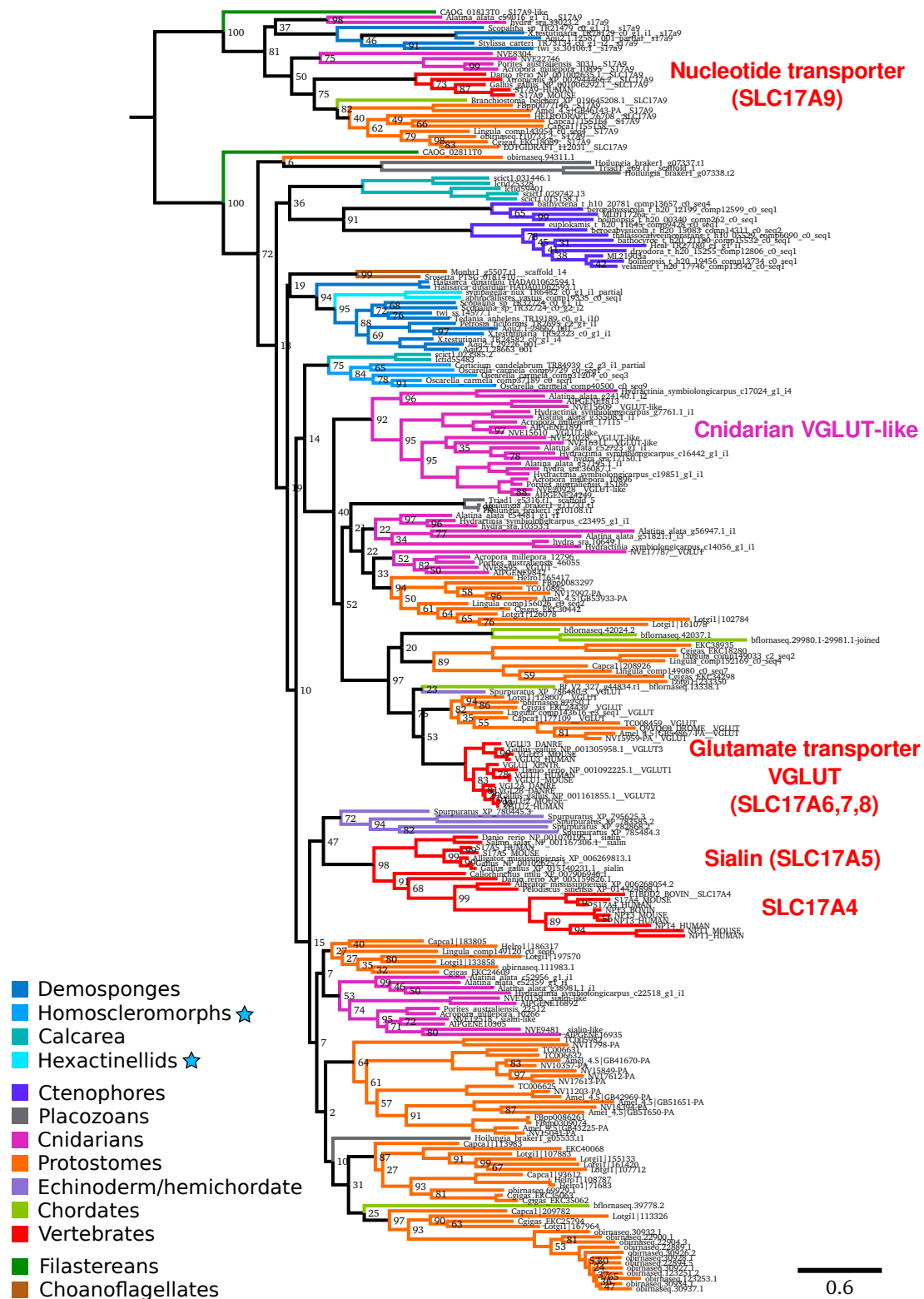


**Supplemental Figure 12: Phylogenetic tree of ionotropic glutamate receptors across metazoans**  
 Protein tree generated with RAxML using the PROTCATWAG model. Bootstrap values are 100 unless otherwise shown. These receptors are not found in the genomes or transcriptomes of demosponges or hexactinellids, so the Sponge clade refers to calcareous sponges and homoscleromorphs. For the three sponges, the blue star indicates sequences derived from a transcriptome. Based on [Alberstein et al., 2015], some receptors are predicted to bind ligands other than glutamate, shown with the green star, red diamond, and black star, for glycine, acidic ligands, and unknown, respectively. Four placozoan proteins have substitutions at the conserved acidic residue (D723 in human GluN1), as either GY in ctenophores, or GG/WY in placozoans; the carboxyl of the glutamic/aspartic acid is needed to coordinate the amino group of glutamate, suggesting that these proteins do not bind an  $\alpha$ -amino acid.

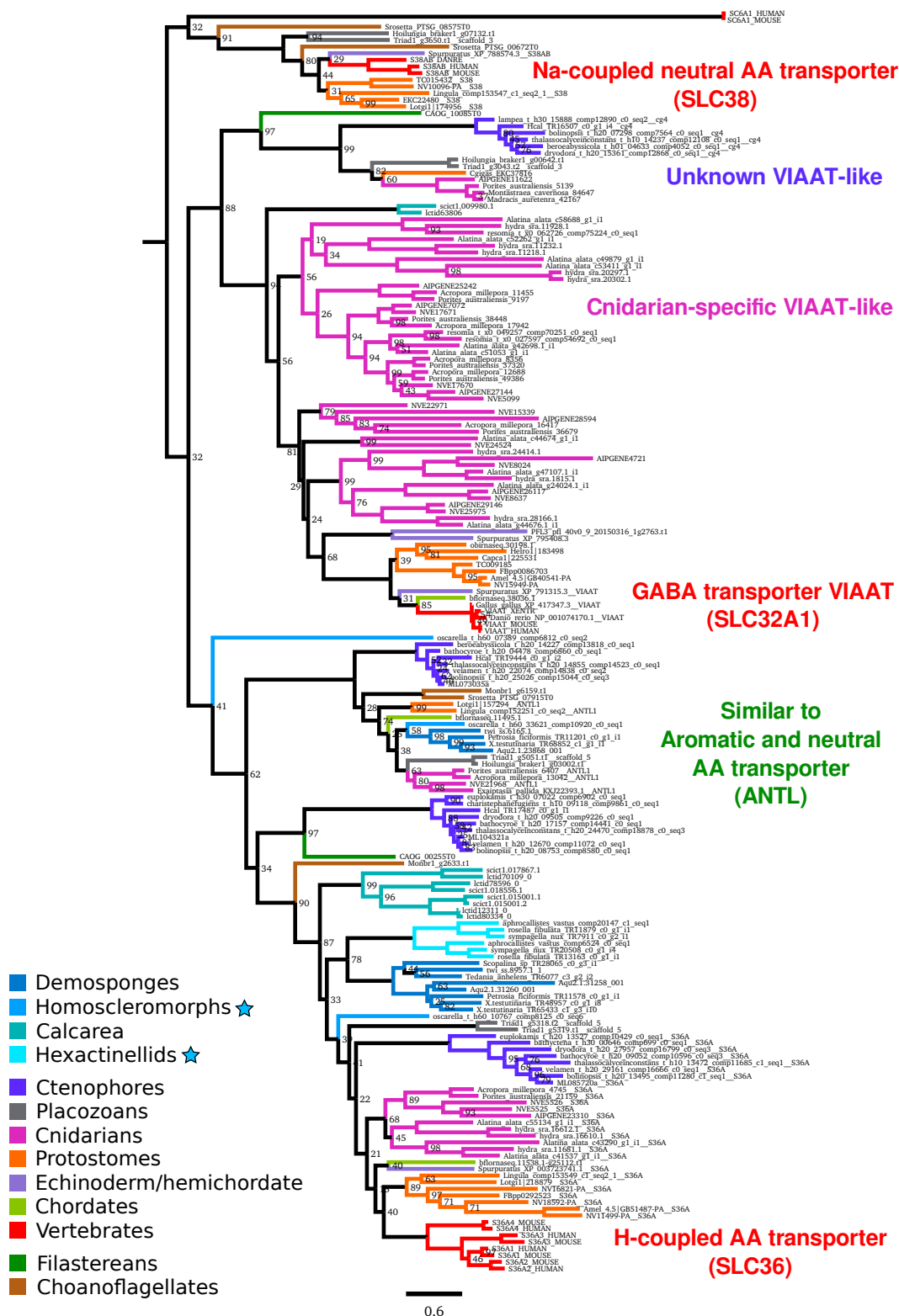


### Supplemental Figure 13: Domain organization of ionotropic glutamate receptors across meta-zoans

Scale bar displays number of amino acids. Top BLAST hits for human iGluRs in demosponges appear to be metabotropic, due to the presence of a 7-transmembrane domain instead of the ion channel, while the ligand-binding domain is conserved. Ctenophore iGluRs and some calcarea/homoscleromorph (Hsm) proteins have the vertebrate-type domain organization, though the other calcarea/homoscleromorph proteins (main sponge group in Supplemental Figure 12) have an SBP domain.

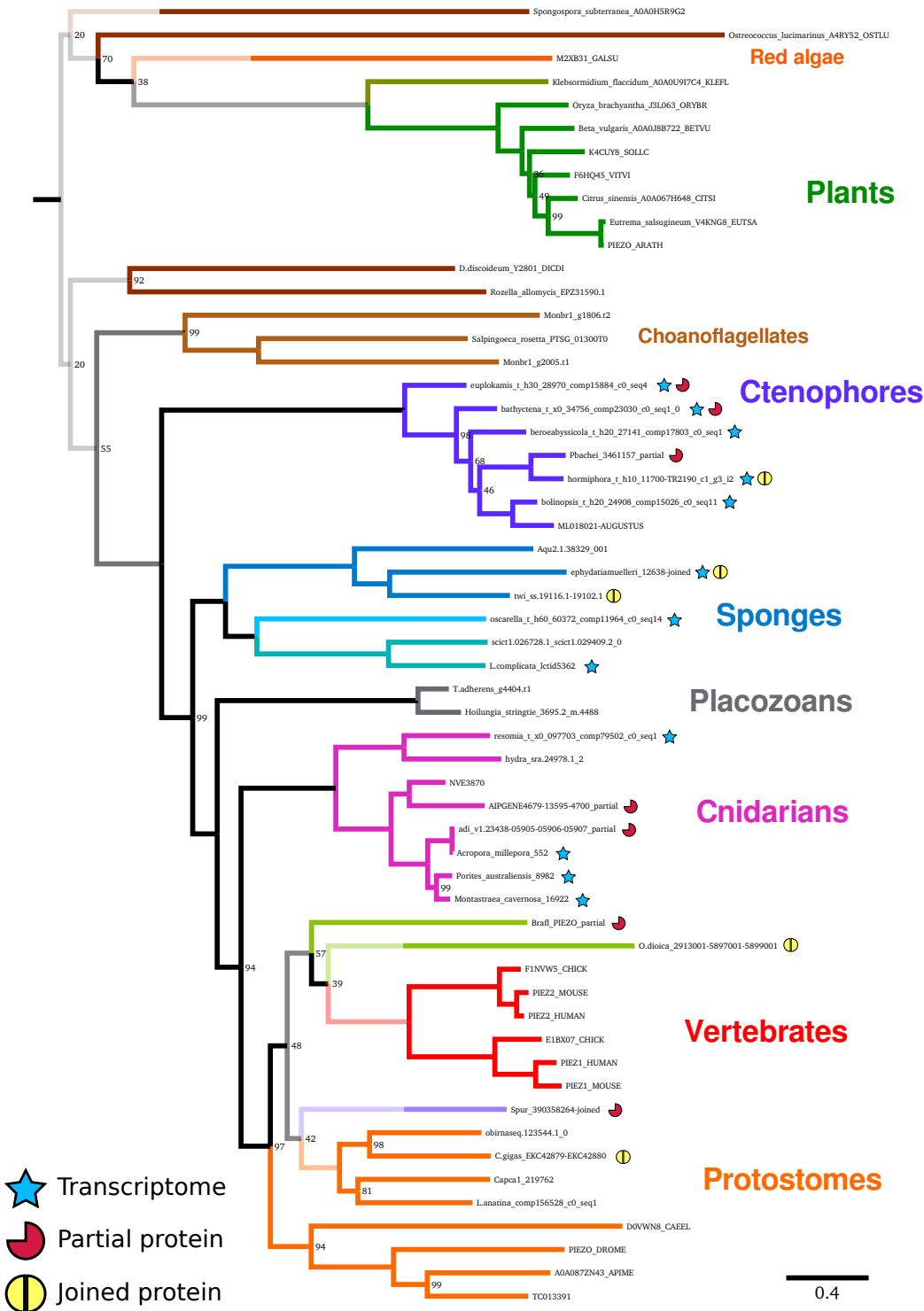


**Supplemental Figure 14: Vesicular glutamate transporter homologs across metazoans**  
 Tree of VGLUT (SLC17A6-8) proteins across all metazoan groups, generated with RAXML using the PROTGAMMALG model. Bootstrap values are 100 unless otherwise shown.



**Supplemental Figure 15: Vesicular inhibitory amino acid transporter homologs across metazoans**

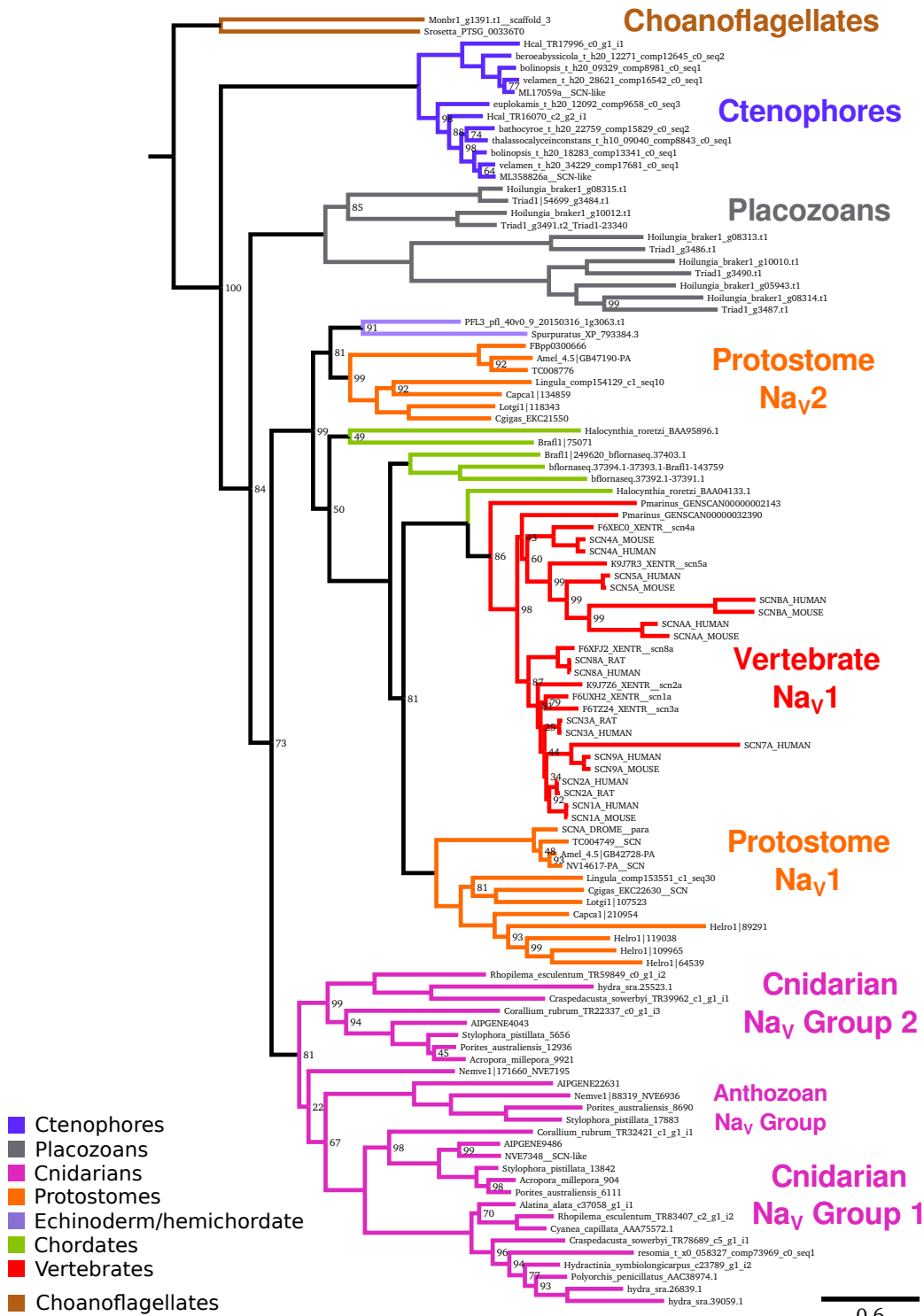
Tree of VIAAT (SLC32A1) proteins and related transporters across all metazoan groups, generated with RAXML using the PROTGAMMALG model. Bootstrap values are 100 unless otherwise shown.



**Supplemental Figure 16: Phylogenetic tree of Piezo homologs across metazoans**

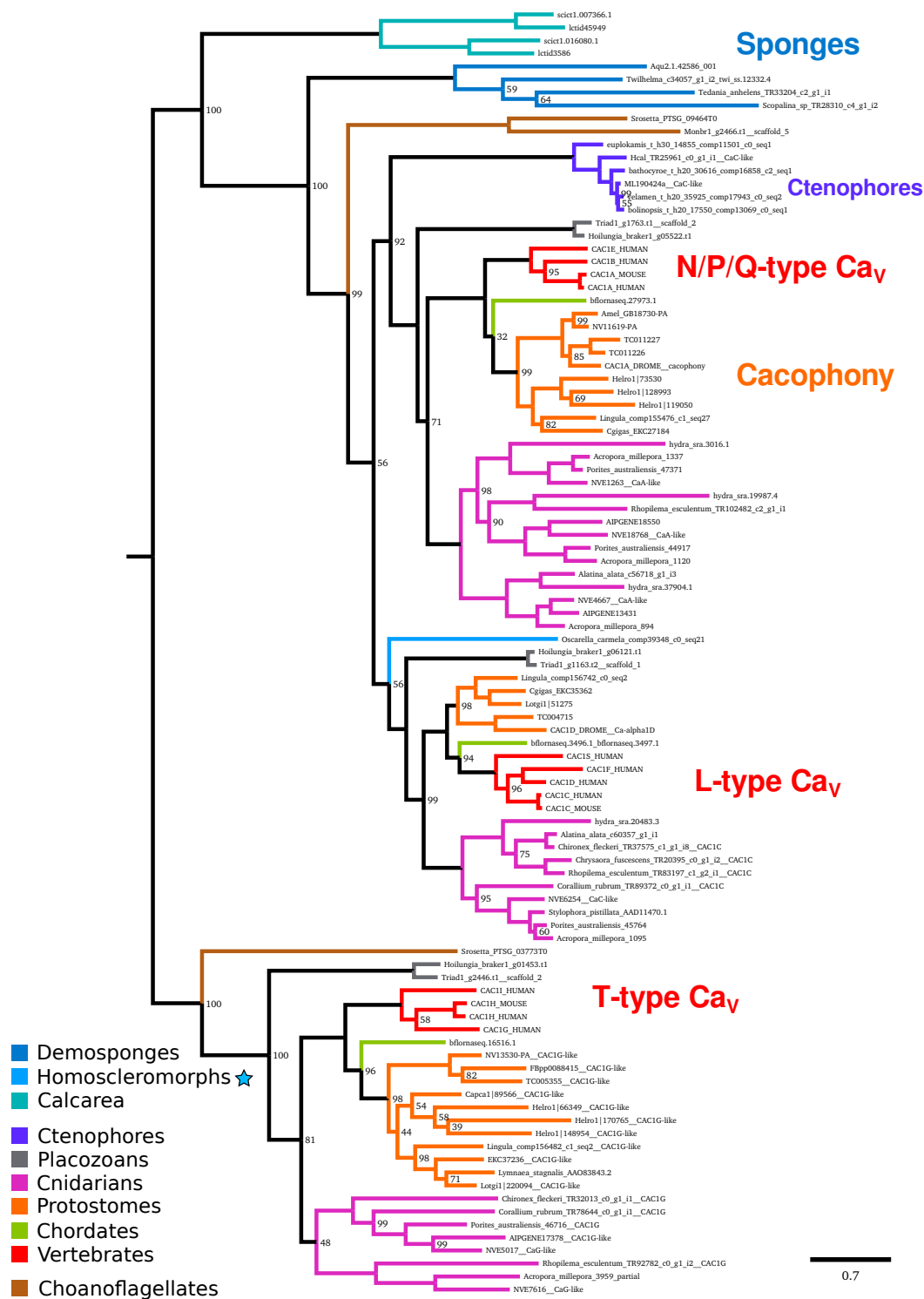
Protein tree generated with RAxML using the PROTCATWAG model and 100 bootstraps. Yellow circle indicates the sequence was complete when joined with other genes or partial sequences, red partial circle indicates the sequence is incomplete in the genome or transcriptome. A blue star indicates that the sequence derived from a transcriptome, so copy number cannot be determined with certainty. Bootstrap values are 100 unless otherwise shown.



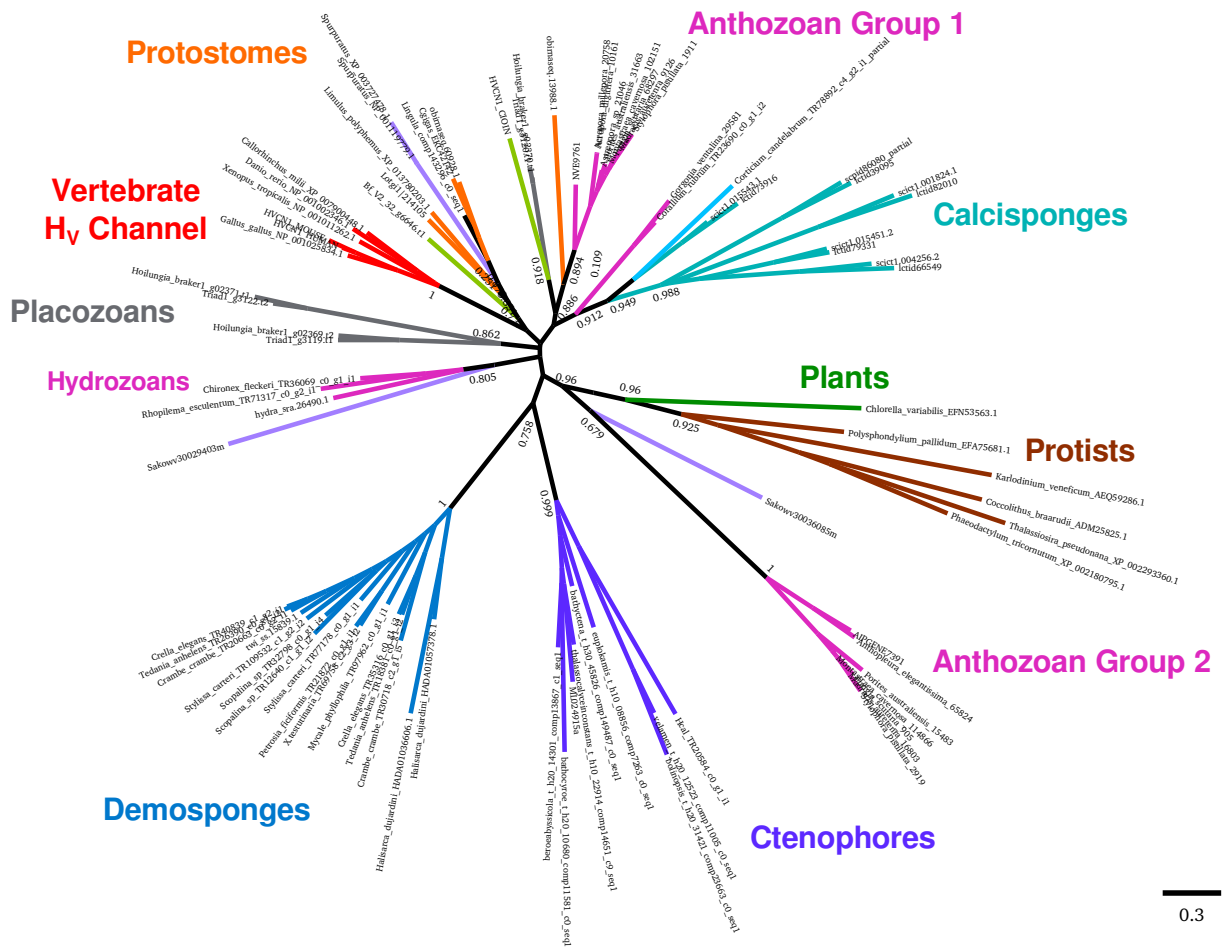


**Supplemental Figure 17: Phylogenetic tree of voltage-gated sodium channel alpha subunits across metazoans**

Protein tree generated with RAxML using the PROTGAMMALG model and 100 bootstraps. Bootstrap values are 100 unless otherwise shown.



**Supplemental Figure 18: Phylogenetic tree of voltage-gated calcium channels across metazoans**  
 Protein tree generated with RAXML using the PROTGAMMALG model and 100 bootstraps. Bootstrap values are 100 unless otherwise shown.



Supplemental Figure 19: Phylogenetic tree of voltage-gated proton channels across metazoans Protein tree generated with FastTree.

974 **2 Supplemental Tables**

**Supplemental Table 1: Summary statistics of the *Tethya wilhelma* genome. Holobiont genomic includes the sponge and all associated bacteria.**

Feature	Type	Count
Assembly size (Mb)	Sponge only	126.0
Total gaps (Mb)	Sponge only	1.348
Estimated kmer coverage	Sponge only	131x
Estimated mapping coverage	Sponge only	159.3x
Number of contigs	All	6,907
Number of contigs	Sponge only	6,109
Number of contigs	Bacterial	789
GC %	Sponge only	39.98%
Contig N50 (kb)	All	70.7
Contig N50 (kb)	Sponge only	73.4
Contig N50 (kb)	Bacterial	48.2
Genome size (bp)	Mitochondrion	19,754
Estimated mapping coverage	Mitochondrion	669.6x
GC %	Mitochondrion	34.43%
Paired-end reads	Holobiont genomic 100bp	259,518,468
Total paired-end bases (Gb)	Holobiont genomic	25.951
Reads aligning back to genome	All contigs	214,103,768
Mate-pair reads	Holobiont genomic 125bp	280,837,536
Total mate-pair bases (Gb)	Holobiont genomic	35.104
Moleculo (TruSeq) long reads	Holobiont genomic	125,150
Total Moleculo bases (Mb)	Holobiont genomic	436.7
Paired-end RNA-seq reads	dUTP Stranded	201,451,574
Paired-end RNA-seq bases (Gb)	dUTP Stranded	25.181
Trinity	De novo transcripts	127,012
RNA-seq mapping fraction	All contigs	68.3%
StringTie	Genome guided transcripts	46,572

**Supplemental Table 2: Summary of splice variation for the genome-guided transcriptome for *T. wilhelma* (Twi) and transcript set v2.0 for *A. queenslandica* (Aqu).**

Splice Type	Twi transcripts	Twi events/exons	Aqu transcripts	Aqu events/exons
Cassette exons	4779	9089	3591	5535
Canonical splicing	5049	-	2602	-
Skipped exons	3868	8329	1022	721
Alternative splice acceptor	-	3747	-	638
Intron retention	3295	3565	3437	3400
Alternative splice donor	-	3264	-	521
Alternative N-terminus	1964	-	1965	-
Alternative C-terminus	1788	-	1968	-
Intronic start	471	-	571	-
Intronic end	285	-	592	-
Non-canonical	73	-	135	-
Single exon with variants	246	-	13	-
No variants	12088	-	24027	-
Single exon and no variants	15421	-	12400	-

**Supplemental Table 3: Skipped exon and retained intron frame, for *T. wilhelma* (Twi) and *A. queenslandica* (Aqu). Skipped exons tend to have lengths as multiples of three.**

Feature	Position 1	Position 2	Position 3
Twi Skipped exons	6904	5179	5331
Twi Retained introns	1202	1204	1159
Aqu Skipped exons	2671	1914	1972
Aqu Retained introns	1244	1092	1101