Additional molecular support for the new chordate phylogeny.

Frédéric Delsuc, Georgia Tsagkogeorga, Nicolas Lartillot, Hervé Philippe

Additional Molecular Support for the New Chordate Phylogeny

Frédéric Delsuc[1,2*], Georgia Tsagkogeorga[1,2], Nicolas Lartillot[3] and Hervé Philippe[4]

[1]Université Montpellier II, CC064, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France

[2]CNRS, Institut des Sciences de l'Evolution (UMR5554), CC064, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France

[3]Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier, CNRS-Université Montpellier II, 161 Rue Ada, 34392 Montpellier Cedex 5, France

[4]Département de Biochimie, Université de Montréal, Succursale Centre-Ville, Montréal, Québec, Canada H3C 3J7

*Correspondence to: Frédéric Delsuc, CC064, Institut des Sciences de l'Evolution, UMR5554-CNRS, Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France. Email: Frederic.Delsuc@univ-montp2.fr.

Contract grant sponsor: Research Networks Program in Bioinformatics from the High Council for Scientific and Technological Cooperation between France and Israel.

SUMMARY

Recent phylogenomic analyses have suggested tunicates instead of cephalochordates as the closest living relatives of vertebrates. In direct contradiction with the long accepted view of Euchordates, this new phylogenetic hypothesis for chordate evolution has been the object of some scepticism. We assembled an expanded phylogenomic dataset focused on deuterostomes. Maximum-likelihood using standard models and Bayesian phylogenetic analyses using the CAT site-heterogeneous mixture model of amino-acid replacement both provided unequivocal support for the sister-group relationship between tunicates and vertebrates (Olfactores). Chordates were recovered as monophyletic with cephalochordates as the most basal lineage. These results were robust to both gene sampling and missing data. New analyses of ribosomal rRNA also recovered Olfactores when compositional bias was alleviated. Despite the inclusion of 25 taxa representing all major lineages, the monophyly of deuterostomes remained poorly supported. The implications of these phylogenetic results for interpreting chordate evolution are discussed in light of recent advances from evolutionary developmental biology and genomics.

Key words: Phylogenomics – Deuterostomes – Chordates – Tunicates – Cephalochordates – Olfactores – Ribosomal RNA – Jackknife – Evolution.

INTRODUCTION

Besides its fundamental role in systematics, phylogenetic reconstruction is a prerequisite for understanding the evolution of organisms. The essential contribution of phylogenetics for understanding morphological diversity has perhaps been best exemplified in the case of animal evolution (Telford and Budd, 2003). The Cambrian explosion has produced a bewildering diversity of body plans whose origins and evolution can only be apprehended by undertaking an integrative approach through evolutionary developmental biology (Evo-Devo) (Conway-Morris, 2003) . The knowledge of phylogenetic relationships, by allowing the polarisation of character transformations, sheds light on the extent of morphological convergence and reversal. A phylogenetic framework is therefore required for distinguishing ancestral characters from those representing morphological innovations. Comparative genomics is now providing the opportunity to track these morphological innovations back to the molecular level by revealing the patterns of gene acquisition/loss, and giving clues to the molecular adaptations that underline the evolution of body plans (Cañestro *et al.*, 2007).

Animal taxonomy has deep roots. The study of morphological and embryological characters has allowed the definition of the major phyla, but left their interrelationships almost unresolved (Nielsen, 2001). The advent of molecular data during the 1990s has revolutionized the traditional classification through a series of phylogenetic analyses of the 18S ribosomal RNA (rRNA) gene for an ever increasing number of key taxa (Aguinaldo *et al.*, 1997; Halanych *et al.*, 1995). This period culminated with the proposition of a new view of animal phylogeny at odds with the traditional paradigm of a steady increase towards morphological complexity, and revealing instead the major role played by secondary simplification from complex ancestors (Adoutte *et al.*, 2000; Lwoff, 1944). Despite these undeniable achievements, the resolving power provided by 18S rRNA and other single genes is nevertheless limited and a number of open questions in animal phylogeny remained to be answered (Halanych, 2004).

The most recent advances in animal phylogeny have come from phylogenomics (Delsuc *et al.*, 2005) which considerably increases the resolving power by considering numerous concatenated genes from Expressed Sequence Tags (ESTs) and complete genome projects (Philippe and Telford, 2006). Despite some troubled beginnings due to the shortcomings of using only a restricted set of taxa (Philippe *et al.*, 2005a), phylogenomics has provided strong corroborating support for the new animal phylogeny, essentially confirming the monophyly of Protostomia, Ecdysozoa and Lophotrochozoa (Baurain *et al.*, 2007; Dunn *et al.*, 2008; Lartillot and Philippe, 2008; Philippe *et al.*, 2005b). Phylogenomics has also helped solving some longstanding mysteries such as the position of chaetognaths which finally appear to

belong to Protostomia (Marlétaz *et al.*, 2006b; Matus *et al.*, 2006) and also proposed unexpected phylogenetic affinities for enigmatic taxa such as *Buddenbrockia plumatellae* recently unmasked as a cnidarian worm (Jimenez-Guri *et al.*, 2007), or *Xenoturbella bocki*, representing a fourth deuterostome phylum on its own (Bourlat *et al.*, 2006).

Among the most groundbreaking results from recent phylogenomic studies was the identification of tunicates (or urochordates) as the closest living relatives of vertebrates, instead of cephalochordates as traditionally accepted (Delsuc *et al.*, 2006). Some hints of this unexpected result had been observed in previous large-scale phylogenetic studies including a single tunicate representative (Blair and Hedges, 2005; Philippe *et al.*, 2005b; Vienne and Pontarotti, 2006). However, a substantial increase in taxon sampling turned out to be required for recovering convincing support in favour of such an unorthodox relationship. In particular, the fact that the inclusion of the divergent appendicularian tunicate *Oikopleura dioica* did not disrupt the sister-group relationship between tunicate and vertebrates gave a good indication about the strength of the phylogenetic signal in its favour (Delsuc *et al.*, 2006). The grouping of tunicates and vertebrates had already been proposed on morphological grounds by Richard P.S. Jefferies who coined the name Olfactores after the presence a putatively homologous olfactory apparatus in fossils that were proposed to be precursors of tunicates and vertebrates (Jefferies, 1991). This phylogenetic result has nevertheless been the object of some scepticism. One reason for this maybe that it further invalidates the traditional textbook view of chordate evolution as a steady increase towards morphological complexity culminating with vertebrates, as betrayed by the use of the term Euchordates (literally “true chordates”) for denoting the grouping of cephalochordates and vertebrates (Gee, 2001). The lack of obvious morphological synapomorphies for Olfactores, apart from the presence of migratory neural crest-like cells (Jeffery, 2007; Jeffery *et al.*, 2004), and the apparent conflict with analyses of rRNA data which tend to favour Euchordates (Cameron *et al.*, 2000; Mallatt and Winchell, 2007; Winchell *et al.*, 2002) might also partly explain the caution with which this result has been considered at first.

Phylogenomics, despite being a powerful approach, is not immune to potential reconstruction artefacts however. The possible pitfalls associated with phylogenomic studies include systematic errors that can be traced back to some kind of model misspecifications (Philippe *et al.*, 2005a) and caused mainly by heterogeneity of evolutionary rates among taxa (Lartillot *et al.*, 2007; Philippe *et al.*, 2005b) and compositional bias (Blanquart and Lartillot, 2008; Jeffroy *et al.*, 2006; Lartillot and Philippe, 2008; Phillips *et al.*, 2004). Empirical protocols have been designed to detect and reduce the impact of systematic error in genome-scale studies (Rodríguez-Ezpeleta *et al.*, 2007) but the ultimate solution lies in the

development of improved models of sequence evolution (Felsenstein, 2004; Philippe *et al.*, 2005a; Steel, 2005). The reliance of current phylogenomic studies on a relatively limited number of highly expressed genes (Philippe and Telford, 2006) and the potential impact of missing data on phylogenomic inference (Hartmann and Vision, 2008; Philippe *et al.*, 2004) are also regularly cited as limitations of the phylogenomic approach.

The aim of this paper is to evaluate the current evidence for the new chordate phylogeny by: (1) reanalyzing previous phylogenomic data using improved models of amino-acid replacement, (2) assembling and analyzing an updated phylogenomic dataset with more genes and more taxa, (3) assessing the impact of missing data and gene sampling on phylogenomic results, and (4) performing new analyses of rRNA data taking compositional bias into account.

MATERIALS AND METHODS

Phylogenomic Dataset Assembly

We built upon previous phylogenomic datasets assembled in the Philippe lab (Delsuc *et al.*, 2006; Jimenez-Guri *et al.*, 2007; Lartillot and Philippe, 2008; Philippe *et al.*, 2005b; Philippe *et al.*, 2004) to select a set of 179 orthologous markers showing sufficient conservation across metazoans to be useful for inferring the phylogeny of metazoans. Alignments were built and updated with available sequences downloaded from the Trace Archive (http://www.ncbi.nlm.nih.gov/Traces/) and the EST Database (http://www.ncbi.nlm.nih.gov/dbEST/) of GenBank at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) using the program ED from the MUST package (Philippe, 1993). Unambiguously aligned regions were identified and excluded for each individual gene using the program GBLOCKS (Castresana, 2000) with a few manual refinements using NET from the MUST package. The complete list of genes with corresponding final numbers of amino-acid sites is available as Supplementary Material.

The concatenation of the 179 genes was constructed with the program SCAFOS (Roure *et al.*, 2007) by defining 51 metazoan operational taxonomic units (OTUs) including 25 deuterostomes representing all major lineages. When several sequences were available for a given OTU, the slowest evolving one was selected according to their degree of divergence using ML distances computed by TREE-PUZZLE (Schmidt *et al.*, 2002) under a WAG+F model (Whelan and Goldman, 2001) within SCAFOS. The percentage of missing data per taxon was reduced by creating some chimerical sequences for species belonging to the same OTU. The complete alignment consists of 179 genes and 51 taxa for 53,799 unambiguously aligned

amino-acid sites with 32% missing data. In order to study the potential impact of missing data on phylogenetic inference (Hartmann and Vision, 2008; Philippe *et al.*, 2004; Wiens, 2006), a concatenation of the 106 genes with sequences available for at least 41 of the 51 OTUs was also constructed with SCAFOS. This reduced alignment consists of 106 genes and 51 taxa for 25,321 amino-acid sites and contains only 20% of missing data. The list of defined OTUs, chimerical sequences and percentages of missing data are available as Supplementary Material. Individual gene alignments and their concatenations are available upon request.

Phylogenomic Analyses

Bayesian phylogenetic analyses of the two phylogenomic datasets were conducted using the program PHYLOBAYES 2.3c (http://www.atgc-montpellier.fr/phylobayes/) under the CAT+Γ_4 site-heterogeneous mixture model (Lartillot and Philippe, 2004). For each dataset, four independent Monte Carlo Markov Chains (MCMCs) starting from a random topology were run in parallel for 20,000 cycles (1,500,000 generations), saving a point every cycle, and discarding the first 2000 points as the burnin. Bayesian posterior probabilities (PP) were obtained from the 50% majority-rule consensus tree of the 18,000 MCMC sampled trees using the program READPB of PHYLOBAYES.

Maximum likelihood (ML) reconstruction on the new phylogenomic dataset was also performed using the program TREEFINDER version of March 2008 (Jobb *et al.*, 2004) under the empirical WAG+F+Γ_8 model of amino-acid substitution. The α shape parameter of the Γ distribution was estimated along with the topology and the branch lengths. Reliability of nodes was estimated by bootstrap resampling with 100 pseudo-replicate datasets generated by the program SEQBOOT of the PHYLIP package (Felsenstein, 2001). The 100 corresponding ML heuristic searches were run in parallel on a computing cluster and the majority-rule consensus of the 100 resulting trees was computed using TREEFINDER.

Jackknife Procedure

The robustness of our phylogenomic inference with respect to gene sampling was assessed by applying a jackknife procedure. Fifty jackknife replicates of 50 genes drawn randomly from the full pool of 179 genes were generated. The only condition we imposed to this jackknife procedure was to require that each taxon is represented by at least one gene in each replicate. The 50 jackknife supermatrices ranging from 11,163 to 17,181 amino-acid sites with 27 to 35% missing data were then analyzed using PHYLOBAYES under the CAT+Γ_4 model. To ensure correct convergence of the MCMC on each replicate, an automated stopping rule was used. Specifically, for each jackknife replicate, two independent parallel (and

synchronous) MCMC were run, until the posterior probability discrepancy between the two chains was less than 0.15 (maximum discrepancy over all bipartitions), and after removing the first 1000 sampled trees of each chain as the burnin. A global majority-rule consensus tree was obtained from the 50 replicates as follows: for each jackknife replicate (D_r) taken in turn, we computed the frequency-table of all bipartitions (splits) observed in the sample collected from the posterior distribution p(T|D_r). The frequencies associated to each bipartition were then averaged over the 50 replicates, and the resulting frequency table was used to build a consensus tree. The support values displayed by this Bayesian consensus tree are thus jackknife-resampled posterior probabilities (PP_{JK}). High PP_{JK} values indicate nodes that have high probability support in most jackknife replicates.

Phylogenetic Analyses of Ribosomal RNA

The 46-taxa dataset of combined 18S+28S rRNA genes assembled by Mallatt and Winchell (Mallatt and Winchell, 2007) for studying deuterostome phylogeny was reanalyzed. This alignment contains a total of 3,925 unambiguously aligned nucleotide sites. A principal component analysis (PCA) of nucleotide composition was realized using the R statistical package (Team, 2007). The best fitting model of nucleotide sequence evolution was evaluated using MODELTEST 3.7 (Posada and Crandall, 1998). The TIM+Γ_4+I transitional model (Posada and Crandall, 2001) was selected according to the Akaïke information criterion. ML phylogenetic analysis of this nucleotide dataset was conducted with PAUP* 4.0b10 (Swofford, 2002) using a heuristic search with Tree Bisection Reconnection (TBR) branch swapping starting from a Neighbor-Joining (NJ) tree.

The nucleotide dataset was RY-coded by pooling puRines (AG => R) and pYrimidines (CT => Y) in an attempt to alleviate both compositional heterogeneity and substitutional saturation of transition events. This RY-coded dataset was then analyzed by conducting a ML heuristic search with TBR branch swapping on a NJ starting tree using PAUP* under the CF+Γ_8 model for discrete characters (Cavender and Felsenstein, 1987). The α shape parameter of the Γ distribution was previously estimated during a ML heuristic search on the nucleotide dataset conducted with TREEFINDER under the GTR2+Γ_8 two-state model.

Reliability of nodes was estimated for each dataset by non-parametric bootstrap resampling using 100 pseudo-replicates generated by SEQBOOT. The 100 corresponding ML heuristic searches using PAUP* with the previously estimated ML parameters, NJ starting trees, and TBR branch swapping were parallelized on a computing cluster. ML bootstrap percentages were obtained from the 50% majority-rule consensus tree of the 100 bootstrap ML trees using TREEFINDER.

RESULTS AND DISCUSSION

Effect of an Improved Model of Sequence Evolution

Our initial assessment of deuterostome phylogenetic relationships was based on a phylogenomic dataset encompassing 146 nuclear genes (33,800 amino-acids) from 38 taxa including 14 deuterostomes (Delsuc *et al.*, 2006). ML and Bayesian phylogenetic analyses conducted under the standard WAG+F+Γ_4 model provided strong support for grouping tunicates with vertebrates (including cyclostomes), but also disrupted chordate monophyly because cephalochordates grouped with echinoderms, albeit with non-significant statistical support (Delsuc *et al.*, 2006). The limited taxon sampling available at the time for Ambulacraria (echinoderms and hemichordates), i.e. a single echinoderm, prompted us to be cautious about this result and to call for the inclusion of xenoturbellidans, hemichordates, and more echinoderms before drawing definitive conclusions. In fact, a subsequent phylogenomic study did exactly what we pleaded for by adding a representative species for each of these three groups (Bourlat *et al.*, 2006). The inclusion of these taxa allowed retrieving the monophyly of chordates in Bayesian analyses using standard amino-acid models, although the alternative hypothesis of chordate paraphyly was still not statistically rejected by ML non-parametric tests (Bourlat *et al.*, 2006). Importantly, the strong statistical support for the monophyly of Olfactores was unaffected by taxon addition (Bourlat *et al.*, 2006).

Models accounting for site-specific modulations of the amino-acid replacement process, such as the CAT mixture model (Lartillot and Philippe, 2004), seem to offer a significantly better fit to real data than empirical substitution matrices currently used in standard models of amino-acid sequence evolution. Accounting for site-specific amino-acid propensities has also been shown to induce a significant improvement of phylogenetic reconstruction in difficult cases such as long-branch attraction (Baurain *et al.*, 2007; Lartillot *et al.*, 2007; Lartillot and Philippe, 2008). This improvement essentially lays in the ability of the CAT model to detect multiple conservative substitutions more efficiently than standard amino-acid models (Lartillot *et al.*, 2007).

In order to test for an eventual effect of model misspecification on previous phylogenomic analyses, we reanalyzed our previous dataset (Delsuc *et al.*, 2006) under the CAT+Γ_4 model. This analysis provides strong corroborating support for the grouping of tunicates and vertebrates (Fig. 1). However, in contrast with previous analyses using empirical amino-acid replacement matrices, which favoured a sister-group relationship between cephalochordates and echinoderms, the use of the CAT+Γ_4 mixture model strongly supports the classical view of monophyletic chordates and deuterostomes (Fig. 1). The fact that chordate polyphyly is

disrupted both by a richer taxon sampling (Bourlat *et al.*, 2006), or upon the use of a more elaborate model, suggests that the previously observed grouping of cephalochordates and echinoderms (Delsuc *et al.*, 2006) was probably a phylogenetic reconstruction artefact. On the other hand, the fact that the grouping of tunicates and vertebrates is insensitive to the model used, adds further credence to the Olfactores hypothesis.

An Updated Phylogenomic Dataset

The continuously growing genomic databases allowed us to build an updated phylogenomic dataset that includes both more genes and more taxa than previously considered to address the question of deuterostome phylogeny. This new dataset of 179 genes for 51 taxa includes 25 deuterostomes representing all major lineages: Xenoturbellida (1 taxon), Hemichordata (1), Echinodermata (5), Cephalochordata (1), Tunicata (6), Cyclostomata (2) and Vertebrata (9), plus 26 selected slow evolving metazoan taxa including Cnidarians and Poriferans as the most distant outgroups. Chordates are particularly well sampled with the inclusion, for the first time, of six tunicate species covering the four major clades evidenced by 18S rRNA studies (Swalla *et al.*, 2000). This diverse taxon sampling is essential to further test the new chordate phylogeny recently revealed by phylogenomics (Bourlat *et al.*, 2006; Delsuc *et al.*, 2006).

Bayesian (CAT+Γ_4) and ML (WAG+F+Γ_8) phylogenetic reconstructions conducted on this updated dataset (179 genes, 53,799 amino-acid sites, 51 taxa) resulted in a highly resolved tree (Fig. 2a). These analyses provided strong support for Ambulacraria ($PP_{CAT+\Gamma4}$ = 1.0 / $BP_{WAG+F+\Gamma8}$ = 97), chordates (1.0 / 69) and olfactores (1.0 / 100). Xenambulacraria (*Xenoturbella* + Ambulacraria) and a basal position for the chaetognath *Spadella* among Protostomia were also moderately supported by our analyses (Fig. 2a). These results are compatible with a recent phylogenomic analysis which also found strong support for Ambulacraria, chordates and Olfactores when using the CAT mixture model (Dunn *et al.*, 2008). However, the monophyly of Deuterostomes is unresolved in both Bayesian and ML phylogenetic reconstructions (Fig. 2a).

The complete dataset obtained by concatenating all 179 genes contains 32% missing data. Previous studies of the impact of missing data on the accuracy of phylogenetic inference have concluded that probabilistic methods are relatively tolerant to missing data (Hartmann and Vision, 2008; Philippe *et al.*, 2004; Wiens, 2003, , 2005), the most important factor being the absolute amount of available data for a given taxon. In phylogenomics, even incomplete taxa are usually represented by thousand of sites and the impact of missing data on accuracy is therefore relatively limited (Philippe *et al.*, 2004). Nonetheless, incomplete taxa might still be

difficult to place with confidence especially when they represent isolated lineages such as *Xenoturbella* (65% missing data) and *Spadella* (75%) in our dataset. In order to control for a potential effect of missing data on our phylogenomic results, we restricted our dataset to the 106 genes with sequences available for at least 41 of the 51 taxa. The concatenation of these 106 genes produces a matrix with 25,321 amino-acid sites that contains only 20% of missing data.

Bayesian and ML phylogenetic inference on this reduced dataset produced a tree fully congruent with the phylogenetic picture given by the complete dataset (Fig. 2b). In particular, the support for the monophyly of chordates was still maximal in terms of $PP_{CAT+\Gamma4}$, but $BP_{WAG+F+\Gamma8}$ increased from 69 to 88%. The monophyly of Olfactores received maximal support in both cases and appeared not affected by missing data. Statistical support in terms of $PP_{CAT+\Gamma4}$ and $BP_{WAG+F+\Gamma8}$ was generally increased especially for locating incomplete taxa such as the *Xenoturbella* as the sister-group to Ambulacraria (from 0.98 / 69 to 1.0 / 88) and *Spadella* at the base of Protostomia (from 0.80 / 56 to 1.0 / 80). Altogether, reducing the amount of missing data, despite also reducing the total number of available sites, seems to result in a slight increase in bootstrap proportions. The only real noticeable difference between the two models concerns the monophyly of deuterostomes. The Bayesian inference under the CAT mixture model suggests deuterostome paraphyly by supporting a basal position of chordates within Bilateria (Fig. 2b) as previously reported (Lartillot and Philippe, 2008). However, ML retrieved the monophyly of deuterostomes, but with $BP_{WAG+F+\Gamma8}$ of only 50%, leaving the monophyly of deuterostomes unresolved by our data.

Robustness of Phylogenomics to Gene Sampling

A legitimate question that can be directed to the phylogenomic approach is the degree to which the results are robust to the sample of genes used to infer phylogenetic trees. This potential concern was addressed by applying a jackknife statistical resampling protocol: fifty datasets were assembled by randomly drawing 50 genes from the total 179 genes, and subjected to Bayesian phylogenetic reconstruction using the CAT+Γ_4 mixture model (see Methods). The resulting majority-rule consensus tree shows that the vast majority of inferred phylogenetic relationships are highly repeatable across the 50 jackknife replicates (Fig. 3). Olfactores, Chordata, and Ambulacraria all received PP_{JK} of more than 90% indicating that phylogenetic support is not dependent upon a particular gene combination. Xenambulacraria appears slightly more affected by gene sampling (PP_{JK} = 80%), but this relative instability

might be explained by the poor gene representation available for *Xenoturbella* with only 98 genes over 179 (55%). The same kind of reasoning could apply to the relatively unstable positions of the chaetognath *Spadella* within protostomes (PP_{JK} = 62%) and of *Holothuria* within echinoderms (PP_{JK} = 51%) (Fig. 3).

In fact, the only major clade whose monophyly appears to be influenced by gene sampling is deuterostomes for which PP_{JK} was less than 50% (Fig. 3). In practise, this means that depending on the particular combination of 50 genes considered, deuterostomes might appear either monophyletic or paraphyletic, with the three possible topological alternatives retrieved in almost similar proportions: Deuterostomes (38%), basal chordates (28%), and basal Xenambulacraria (22%). Despite the inclusion of 25 taxa representing all major lineages in our dataset, these results confirm deuterostomes as one of the most difficult groups to resolve in the animal phylogeny despite its wide acceptance (see (Lartillot and Philippe, 2008)).

New Analyses of Ribosomal RNA Genes

The sister-group relationship between tunicates and vertebrates (Olfactores) observed in phylogenomics is in conflict with most (if not all) analyses of rRNA which favour cephalochordates as the closest relatives of vertebrates (Euchordates) (Cameron *et al.*, 2000; Mallatt and Winchell, 2007; Swalla *et al.*, 2000; Wada and Satoh, 1994; Winchell *et al.*, 2002). However, the statistical support for Euchordates in rRNA-based phylogenetic studies is moderate. Indeed, the first 18S rRNA study, based on a limited taxon sampling of deuterostomes, reported a bootstrap value of only 45% for Euchordates (Wada and Satoh, 1994). A subsequent 18S rRNA study considering only slowly evolving sequences for 16 deuterostomes found only a moderate bootstrap support of 71% for grouping cephalochordates and vertebrates (Cameron *et al.*, 2000). A study focused on tunicates also obtained moderate support for Euchordates (58 to 85% depending on the dataset and reconstruction method) but failed to support chordate monophyly likely because tunicate 18S rRNA sequences are rapidly evolving (Swalla *et al.*, 2000).

The next studies used the combination of 18S and 28S rRNAs. An investigation using 28 taxa for the two rRNA subunits found strong boostrap support (89 to 97% depending on the method) for Euchordates (Winchell *et al.*, 2002). However, this study again failed to support chordate monophyly. Detailed analyses confirmed that tunicate genes have evolved rapidly and showed that they are compositionally biased towards AT, rendering tunicates virtually impossible to locate convincingly in the tree on the basis of rRNA data (Winchell *et al.*, 2002). Finally, increasing the sampling to 46 taxa for this 18S+28S rRNA data did not helped in further resolving the relationships among the major groups of deuterostomes and even

decreased the ML bootstrap support for Euchordates from 97% in the previous study to 50% (Mallatt and Winchell, 2007).

In order to gauge the extent to which the rRNA data conflicts with our phylogenomic results, we reanalyzed the 46-taxa dataset of Mallatt and Winchell (Mallatt and Winchell, 2007). The heterogeneity of base composition in this dataset is well illustrated by the PCA presented in Figure 4a. At one extreme, tunicates (especially *Oikopleura*) are particularly AT-rich, and at the other extreme, Myxinidae (*Myxine* and *Eptatretus*) and the pterobranch hemichordate *Cephalodiscus* are highly GC-rich. We therefore compared phylogenetic reconstructions conducted on nucleotides and on RY-coded data, a coding scheme allowing reducing both substitutional saturation and nucleotide compositional bias (Fig. 4b). The two inferred ML trees appear mostly congruent except for two major topological shifts.

The strongest topological change occurred within hemichordates (Fig. 2b). Whereas the use of a standard DNA model strongly supports the paraphyly of enteropneusts by grouping the pterobranch *Cephalodiscus* with *Saccoglossus* and *Harrimania* (BP = 95), RY-coding allows recovering the monophyly of enteropneusts with high bootstrap support (BP = 90). This helps in understanding the conflict between 18S rRNA that supports enteropneust paraphyly (Cameron *et al.*, 2000; Halanych, 1995) and 28S rRNA that rather favours their monophyly (Mallatt and Winchell, 2007; Winchell *et al.*, 2002). This result is of particular importance because it potentially invalidates the controversial hypothesis that pterobranchs evolved from an enteropneust (Cameron *et al.*, 2000; Halanych *et al.*, 1995) by suggesting that it is likely an artefact of 18S rRNA-based phylogenetic reconstructions due to shared nucleotide compositional bias between pterobranchs and Harrimaniidae (Fig. 4a).

Second, the support for the monophyly of Euchordates observed with nucleotides (BP = 84) disappeared in favour of the monophyly of Olfactores in the RY-coding dataset, albeit with no statistical support (BP = 44). This nevertheless strongly suggests that the high composition bias of tunicate sequences has blurred the phylogenetic signal for Olfactores in previous analyses. Thus, according to our interpretation, reducing compositional bias and substitutional saturation by RY-recoding allows recovering a limited signal for Olfactores in agreement with our phylogenomic analysis of amino-acid data. It is worth noting however, that rRNA does not statistically support chordate monophyly in both cases (Fig. 2b).

Molecular Phylogenetic Conclusions

Our aim was to reanalyse the phylogenetic relationships among chordates. The revision of the position of tunicates proposed by recent phylogenomic studies (Bourlat *et al.*, 2006; Delsuc *et al.*, 2006; Dunn *et al.*, 2008) by concluding in favour of the monophyly of

Olfactores, has not yet been considered as totally convincing, essentially because it is at odds with both the traditional view based on embryological and morphological characters (Rowe, 2004; Schaeffer, 1987), and with earlier molecular phylogenetic analyses based on rRNA (Cameron *et al.*, 2000; Mallatt and Winchell, 2007; Swalla *et al.*, 2000; Wada and Satoh, 1994; Winchell *et al.*, 2002). The unexpected sister-group relationship between echinoderms and cephalochordates observed in one of these studies (Delsuc *et al.*, 2006) may also have suggested the possibility that the monophyly of Olfactores was due to an artefactual attraction of cephalochordates with echinoderms (Bourlat *et al.*, 2006).

In the present analysis, we have tried to address these points, essentially by reanalyzing both phylogenomic and rRNA data, under better taxonomic sampling and using more elaborate methods and probabilistic models. First, we demonstrate that, although the grouping of echinoderms and cephalochordates was indeed a probable artefact, disappearing upon the addition of several taxa or using an improved model of sequence evolution, the monophyly of Olfactores appears to be robust with respect to taxon sampling and model choice. Second, our reanalysis of rRNA data using RY-recoding also reveals a weak signal in favour of Olfactores, and suggests that the grouping of vertebrates and cephalochordates in former studies may have been an artefact driven by compositional biases. Altogether, our analyses allows a coherent interpretation of all empirical results observed thus far concerning chordate phylogeny, yielding further evidence in favour of the monophyly of both chordates and Olfactores.

At larger scale, however, we observe an overall lack of support for the monophyly of deuterostomes. Deuterostomes have nearly unanimously been considered as an unquestionable monophyletic group, a hypothesis backed up by traditional comparative analyses of embryological characters such as the fate of the blastopore (Nielsen, 2001), and morphological traits such as gill slits (Schaeffer, 1987). However, in our analyses, the status of deuterostomes seems to be sensitive to the model used, with CAT slightly favouring the paraphyletic configuration, and WAG the more traditional monophyly. In either case, the support measured by non-parametric resampling procedures (site-wise bootstrap or gene-wise jackknife) is weak.

Other phylogenomic studies (Dunn *et al.*, 2008; Lartillot and Philippe, 2008) also failed to obtain strong support for the relative phylogenetic positions of chordates and Ambulacraria. Moderate support for the monophyly of Deuterostomes was only obtained under empirical matrix models, the support disappearing when the CAT model was used instead (Dunn *et al.*, 2008; Lartillot and Philippe, 2008). Although profile mixture models such as CAT, whereas having a better fit than empirical matrices such as WAG, may have some inherent weaknesses

as to their phylogenetic accuracy, the WAG empirical matrix fails in many cases, particularly when confronted to a high level of saturation (Lartillot *et al.*, 2007). This casts doubts on results that seem to receive support exclusively under this model, as is the case for deuterostome monophyly. More observations are needed to better gauge the relative merits of either type of model. Overall, although deuterostome monophyly still remains a reasonable working hypothesis to date, more work is needed before the question can be settled.

Corroborative Evidence for Olfactores Monophyly

The monophyly of Olfactores receives strong support from sequence-based phylogenomic inference. Rare genomic changes has also provided some evidence in its favour: the domain structure of cadherins (Oda *et al.*, 2002), a unique amino-acid insertion in fibrillar collagen (Wada *et al.*, 2006), and the distribution of micro RNAs (miRNAs) (Heimberg *et al.*, 2008). Finally, the *Branchiostoma floridae* genome helps confirming the sister-group relationship between tunicates and vertebrates in offering additional evidence from analyses of intron dynamics (Putnam *et al.*, 2008).

Cadherins are a superfamily of highly conserved adhesion molecules mediating cell communication and signalling that are pivotal for developmental processes of multicellular organisms. Their recent detection in the closest unicellular relatives of metazoans, the choanoflagellates, has highlighted their potential role in the origin of multicellularity (Abedin and King, 2008). Comparative studies on the classic cadherin subfamily has revealed that the structural element called Primitive Classic Cadherin Domain (PCCD) complex, otherwise termed non-chordate classic cadherin domain, is also present in cephalochordates, but has been lost in both tunicates and vertebrates (Oda *et al.*, 2002). The most parsimonious scenario is that this particular protein domain complex has been lost in the common ancestor of tunicates and vertebrates and constitutes a synapomorphy of Olfactores. However, cephalochordates possess two classic cadherin genes which originated by lineage-specific tandem duplication and that have a particular structure in lacking extracellular repeats found in all other investigated metazoans (Oda *et al.*, 2004). This derived state renders difficult to ascertain domain homology among chordate classic cadherin genes and casts doubt on its phylogenetic significance.

Further potential evidence for the clade Olfactores has been inferred from the evolution of fibrillar collagen genes within chordates. These genes represent important components of the notochord, the cartilage and mineralized bones in vertebrates. Phylogenetic analyses suggested that three ancestral fibrillar collagens gave rise to the gene diversity observed in living deuterostomes (Wada *et al.*, 2006). Comparative sequence analyses showed that

tunicates and vertebrates share a molecular signature in the form of a six to seven amino-acid insertion in the C-terminus noncollagenous domain of one type of fibrillar collagens, that is absent in cephalochordates and echinoderms (Wada *et al.*, 2006). This insertion was interpreted as supporting the idea that vertebrates are more closely related to tunicates than to cephalochordates (Wada *et al.*, 2006). The homology of the insertion appears nevertheless difficult to assert with certainty given the high degree of sequence divergence observed in this region of the molecule. More tunicate fibrillar collagen sequences might help in better understanding the dynamics of this peculiar amino-acid insertion and the phylogenetic signal it conveys.

The comparison of miRNA repertoires in metazoans has also recently unearthed some potential signatures for the sister-group relationship of tunicates and vertebrates (Heimberg *et al.*, 2008). miRNAs are small non-coding RNAs involved in regulation of gene expression in eukaryotes and playing an important role in the development of metazoans. Comparative genomic studies of miRNAs underlined that, during the evolution of metazoans, major body-plan innovations seemed to coincide with dramatic expansions of miRNA repertoires, suggesting a potential role in the increase of morphological complexity (Hertel *et al.*, 2006; Sempere *et al.*, 2006). The most recent study unveiled that three miRNA families (*mir-126*, *mir-135* and *mir-155*) were likely acquired in the common ancestor of tunicates and vertebrates (Heimberg *et al.*, 2008). Taking into consideration that miRNAs might be only rarely secondarily lost once they have been recruited, this finding provides corroborative evidence for the clade Olfactores. It should be noted however that, of these three miRNA families, only *mir-126* constitutes an exclusive synapomorphy for Olfactores without subsequent secondary lost in descendant taxa confirmed by Northern analysis. Moreover, the profound reorganization of miRNA repertoire undergone by tunicates requires being cautious when interpreting acquisition of miRNAs as potential signatures for reconstructing their phylogenetic relationships (Fu *et al.*, 2008).

Additional sequence-based phylogenomic reconstructions and analyses of rare genomic changes have been issued along with the recently published draft sequence of a cephalochordate (*Branchiostoma floridae*) genome (Putnam *et al.*, 2008). The phylogenetic analysis of a concatenation of 1,090 orthologs from 12 complete genomes retrieved maximal Bayesian support for Olfactores and chordates, whereas the corresponding bootstrap support was maximal for Olfactores but of only 78% for chordate monophyly (Putnam *et al.*, 2008). Moreover, the analysis of individual gene phylogenies revealed twice more cases where Olfactores was favoured over Euchordates than the reverse (Putnam *et al.*, 2008). Further evidence was obtained by analysing the phylogenetic signal deduced from the dynamics of

intron gain and loss among chordate genomes. Despite extensive intron losses along the tunicate lineage, a number of shared intron gain/loss events can be identified as a signature of tunicates and vertebrates common ancestry (Putnam *et al.*, 2008). Overall, the new evidence brought by the analysis of the *Branchiostoma floridae* genome essentially corroborates our present phylogenetic results.

Implications for Chordate Evo-Devo

The additional evidence presented for the new chordate phylogeny provides a robust phylogenetic framework for (re)interpreting the evolution of morphological characters and developmental features. Inverting the phylogenetic position of tunicates and cephalochordates within monophyletic chordates highlights the prevalence of morphological simplification with characters that are likely ancestral for chordates, such as metameric segmentation, being lost secondarily in the tunicate lineage. On the other hand, the loss of pre-oral kidney and the presence of multiciliated epithelial cells might in fact constitute morphological synapomorphies for olfactores (Ruppert, 2005). The new chordate phylogeny further portrays tunicates as highly derived chordates with specialized lifestyles and developmental modes, whereas cephalochordates might have retained more ancestral chordate characteristics. We will use two examples to illustrate the importance of considering the new phylogenetic status of tunicates as the sister-group of vertebrates in the context of evolutionary developmental biology.

The first illustration concerns evolutionary origin of such fundamental structures as the neural crest and olfactory placodes. Migratory neural crest cells and sensory placodes have long been considered as vertebrate innovations. Implicated respectively in the development of major tissues and sensory organs, their origin is generally correlated with the increase in morphological complexity of vertebrates. However, recent molecular developmental studies have revealed the presence in tunicates of migratory neural crest-like cells (Jeffery, 2006; Jeffery *et al.*, 2004) and olfactory placodes (Bassham and Postlethwait, 2005; Mazet and Shimeld, 2005). When reinterpreted in light of the new chordate phylogeny, these results implied that both of these features did not evolve *de novo* in the vertebrate lineage, but rather evolved from specialized pre-existing structures in the common ancestor of vertebrates and tunicates.

The second example illustrates how the new phylogenetic context helps in understanding the genomic and developmental peculiarities of tunicates within chordates. The new phylogenetic picture implied that tunicate genomes have undergone significant genome reduction from the ancestral chordate genome (Holland, 2007). This genome compaction is

also associated with a high rate of genomic evolution at the levels of both primary sequences (Delsuc *et al.*, 2006; Edvardsen *et al.*, 2004) and genome organisation (Holland and Gibson-Brown, 2003). One of the most spectacular rearrangements of tunicate genomes is the lost of several Hox genes, the disintegration of the Hox cluster, and the lost of temporal colinearity in Hox gene expression during development (Ikuta *et al.*, 2004; Seo *et al.*, 2004). These observations raise the question of how tunicates, with their altered Hox clusters, are still able to develop a chordate body plan. In chordates, and deuterostomes more generally, temporal colinearity is regulated by the Retinoic-Acid (RA) signalling pathway which controls the antero-posterior patterning of the embryo (Cañestro *et al.*, 2006; Marlétaz *et al.*, 2006a). However, axial patterning in tunicates seems to have become independent of RA-signaling, with the genes of the RA machinery even being lost in *Oikopleura* (Cañestro and Postlethwait, 2007). Functional studies have shown that if "*Oikopleura* can be considered as a classical RA-signaling knock-down mutant naturally produced by evolution", it is still capable of developing a typical chordate body plan (Cañestro and Postlethwait, 2007). With cephalochordates, which possess the RA genomic toolkit, being basal among chordates, RA-signalling must have been present in the tunicate ancestor and secondarily lost in *Oikopleura* suggesting that appendicularians use alternative mechanisms for the development of chordate features (Cañestro *et al.*, 2007; Holland, 2007).

The new chordate phylogeny strengthens the view that tunicates and cephalochordates represent complementary models for studying vertebrate Evo-Devo (Schubert *et al.*, 2006). Tunicates are phylogenetically closer to vertebrates but appear both morphologically and molecularly highly derived. The diversity of their developmental modes offers the opportunity to study the evolution of alternative adaptive solutions to the typical chordate development. In having retained more ancestral features, cephalochordates provide an ideal outgroup for polarizing evolutionary changes that occurred in tunicates and vertebrates. With the cephalochordate *Branchiostoma floridae* genome (Putnam *et al.*, 2008) and the upcoming genome sequence of the appendicularian *Oikopleura dioica*, the newly established phylogenetic framework makes chordate comparative genomics appearing full of promises for the Evo-Devo community as exemplified in a recent work on the origin and evolution of the Pax gene family (Bassham *et al.*, 2008).

ACKNOWLEDGMENTS

We thank the associate editors Billie Swalla and José Xavier-Neto for the opportunity to write this paper. We also thank John Mallatt for kindly providing his 18S-28S rRNA alignment, Julien Claude for help in using R, and two anonymous reviewers for comments.

The extensive phylogenomic calculations benefited from the ISEM computing cluster. This is the contribution ISEM 2008-062 of the Institut des Sciences de l'Evolution.

LITERATURE CITED

Abedin M, King N. 2008. The premetazoan ancestry of cadherins. Science 319:946-948.

Adoutte A, Balavoine G, Lartillot N, Lespinet O, Prud'homme B, de Rosa R. 2000. The new animal phylogeny: reliability and implications. Proc Natl Acad Sci USA 97:4453-4456.

Aguinaldo AM, Turbeville JM, Linford LS, Rivera MC, Garey JR, Raff RA, Lake JA. 1997. Evidence for a clade of nematodes, arthropods and other moulting animals. Nature 387:489-493.

Bassham S, Cañestro C, Postlethwait JH. 2008. Evolution of developmental roles of Pax2/5/8 paralogs after independent duplication in urochordate and vertebrate lineages. BMC Biol 6:35.

Bassham S, Postlethwait JH. 2005. The evolutionary history of placodes: a molecular genetic investigation of the larvacean urochordate Oikopleura dioica. Development 132:4259-4272.

Baurain D, Brinkmann H, Philippe H. 2007. Lack of resolution in the animal phylogeny: closely spaced cladogeneses or undetected systematic errors? Mol Biol Evol 24:6-9.

Blair JE, Hedges SB. 2005. Molecular phylogeny and divergence times of deuterostome animals. Mol Biol Evol 22:2275-2284.

Blanquart S, Lartillot N. 2008. A site- and time-heterogeneous model of amino acid replacement. Mol Biol Evol 25:842-858.

Bourlat SJ, Juliusdottir T, Lowe CJ, Freeman R, Aronowicz J, Kirschner M, Lander ES, Thorndyke M, Nakano H, Kohn AB, Heyland A, Moroz LL, Copley RR, Telford MJ. 2006. Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. Nature 444:85-88.

Cameron CB, Garey JR, Swalla BJ. 2000. Evolution of the chordate body plan: new insights from phylogenetic analyses of deuterostome phyla. Proc Natl Acad Sci USA 97:4469-4474.

Cañestro C, Postlethwait JH. 2007. Development of a chordate anterior-posterior axis without classical retinoic acid signaling. Dev Biol 305:522-538.

Cañestro C, Postlethwait JH, Gonzalez-Duarte R, Albalat R. 2006. Is retinoic acid genetic machinery a chordate innovation? Evol Dev 8:394-406.

Cañestro C, Yokoi H, Postlethwait JH. 2007. Evolutionary developmental biology and genomics. Nat Rev Genet 8:932-942.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 17:540-552.

Cavender JA, Felsenstein J. 1987. Invariants of phylogenies in a simple case with discrete states. J Classif 4:57-71.

Conway-Morris S. 2003. The Cambrian "explosion" of metazoans and molecular biology: would Darwin be satisfied? Int J Dev Biol 47:505-515.

Delsuc F, Brinkmann H, Chourrout D, Philippe H. 2006. Tunicates and not cephalochordates are the closest living relatives of vertebrates. Nature 439:965-968.

Delsuc F, Brinkmann H, Philippe H. 2005. Phylogenomics and the reconstruction of the tree of life. Nat Rev Genet 6:361-375.

Dunn CW, Hejnol A, Matus DQ, Pang K, Browne WE, Smith SA, Seaver E, Rouse GW, Obst M, Edgecombe GD, Sorensen MV, Haddock SH, Schmidt-Rhaesa A, Okusu A, Kristensen RM, Wheeler WC, Martindale MQ, Giribet G. 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. Nature 452:745-749.

Edvardsen RB, Lerat E, Maeland AD, Flat M, Tewari R, Jensen MF, Lehrach H, Reinhardt R, Seo HC, Chourrout D. 2004. Hypervariable and highly divergent intron-exon organizations in the chordate *Oikopleura dioica*. J Mol Evol 59:448-457.

Felsenstein J. 2001. PHYLIP (Phylogenetic Inference Package) version 3.6. In: Distributed by the author, Department of Genetics, University of Washington, Seattle.

Felsenstein J. 2004. Inferring phylogenies. Sunderland, MA, USA: Sinauer Associates, Inc. 645 p.

Fu X, Adamski M, Thompson EM. 2008. Altered miRNA repertoire in the simplified chordate, Oikopleura dioica. Mol Biol Evol 25:1067-1080.

Gee H. 2001. Deuterostome phylogeny: the context for the origin and evolution of the vertebrates. In: Ahlberg PE, editor. Major events in early vertebrate evolution: palaeontology, phylogeny, genetics, and development. London: Taylor and Francis. p 1-14.

Halanych KM. 1995. The phylogenetic position of the pterobranch hemichordates based on 18S rDNA sequence data. Mol Phylogenet Evol 4:72-76.

Halanych KM. 2004. The new view of animal phylogeny. Annu Rev Ecol Evol Syst 35:229-256.

Halanych KM, Bacheller JD, Aguinaldo AM, Liva SM, Hillis DM, Lake JA. 1995. Evidence from 18S ribosomal DNA that the lophophorates are protostome animals. Science 267:1641-1643.

Hartmann S, Vision TJ. 2008. Using ESTs for phylogenomics: can one accurately infer a phylogenetic tree from a gappy alignment? BMC Evol Biol 8:95.

Heimberg AM, Sempere LF, Moy VN, Donoghue PC, Peterson KJ. 2008. MicroRNAs and the advent of vertebrate morphological complexity. Proc Natl Acad Sci USA 105:2946-2950.

Hertel J, Lindemeyer M, Missal K, Fried C, Tanzer A, Flamm C, Hofacker IL, Stadler PF. 2006. The expansion of the metazoan microRNA repertoire. BMC Genomics 7:25.

Holland LZ. 2007. Developmental biology: a chordate with a difference. Nature 447:153-155.

Holland LZ, Gibson-Brown JJ. 2003. The *Ciona intestinalis genome*: when the constraints are off. Bioessays 25:529-532.

Ikuta T, Yoshida N, Satoh N, Saiga H. 2004. *Ciona intestinalis* Hox gene cluster: Its dispersed structure and residual colinear expression in development. Proc Natl Acad Sci USA 101:15118-15123.

Jefferies RPS. 1991. Two types of bilateral symmetry in the Metazoa: chordate and bilaterian. In: Bock GR, Marsh J, editors. Biological Asymmetry and Handedness. Chichester: Wiley. p 94-127.

Jeffery WR. 2006. Ascidian neural crest-like cells: phylogenetic distribution, relationship to larval complexity, and pigment cell fate. J Exp Zoolog B Mol Dev Evol 306:470-480.

Jeffery WR. 2007. Chordate ancestry of the neural crest: new insights from ascidians. Semin Cell Dev Biol 18:481-491.

Jeffery WR, Strickler AG, Yamamoto Y. 2004. Migratory neural crest-like cells form body pigmentation in a urochordate embryo. Nature 431:696-699.

Jeffroy O, Brinkmann H, Delsuc F, Philippe H. 2006. Phylogenomics: the beginning of incongruence? Trends Genet 22:225-231.

Jimenez-Guri E, Philippe H, Okamura B, Holland PWH. 2007. *Buddenbrockia* is a Cnidarian worm. Science 317:116-118.

Jobb G, von Haeseler A, Strimmer K. 2004. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. BMC Evol Biol 4:18.

Lartillot N, Brinkmann H, Philippe H. 2007. Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. BMC Evol Biol 7 Suppl 1:S4.

Lartillot N, Philippe H. 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol Biol Evol 21:1095-1109.

Lartillot N, Philippe H. 2008. Improvement of molecular phylogenetic inference and the phylogeny of Bilateria. Philos Trans R Soc Lond B Biol Sci 363:1463-1472.

Lwoff A. 1944. L'évolution physiologique : étude des pertes de fonctions chez les microorganismes. Paris: Hermann. 308 p.

Mallatt J, Winchell CJ. 2007. Ribosomal RNA genes and deuterostome phylogeny revisited: more cyclostomes, elasmobranchs, reptiles, and a brittle star. Mol Phylogenet Evol 43:1005-1022.

Marlétaz F, Holland LZ, Laudet V, Schubert M. 2006a. Retinoic acid signaling and the evolution of chordates. Int J Biol Sci 2:38-47.

Marlétaz F, Martin E, Perez Y, Papillon D, Caubit X, Lowe CJ, Freeman B, Fasano L, Dossat C, Wincker P, Weissenbach J, Le Parco Y. 2006b. Chaetognath phylogenomics: a protostome with deuterostome-like development. Curr Biol 16:R577-578.

Matus DQ, Copley RR, Dunn CW, Hejnol A, Eccleston H, Halanych KM, Martindale MQ, Telford MJ. 2006. Broad taxon and gene sampling indicate that chaetognaths are protostomes. Curr Biol 16:R575-576.

Mazet F, Shimeld SM. 2005. Molecular evidence from ascidians for the evolutionary origin of vertebrate cranial sensory placodes. J Exp Zoolog B Mol Dev Evol 304:340-346.

Nielsen C. 2001. Animal evolution, interrelationships of the living phyla. Oxford, UK: Oxford University Press.

Oda H, Akiyama-Oda Y, Zhang S. 2004. Two classic cadherin-related molecules with no cadherin extracellular repeats in the cephalochordate amphioxus: distinct adhesive specificities and possible involvement in the development of multicell-layered structures. J Cell Sci 117:2757-2767.

Oda H, Wada H, Tagawa K, Akiyama-Oda Y, Satoh N, Humphreys T, Zhang S, Tsukita S. 2002. A novel amphioxus cadherin that localizes to epithelial adherens junctions has an unusual domain organization with implications for chordate phylogeny. Evol Dev 4:426-434.

Philippe H. 1993. MUST, a computer package of Management Utilities for Sequences and Trees. Nucleic Acids Res 21:5264-5272.

Philippe H, Delsuc F, Brinkmann H, Lartillot N. 2005a. Phylogenomics. Annu Rev Ecol Evol Syst 36:541-562.

Philippe H, Lartillot N, Brinkmann H. 2005b. Multigene analyses of bilaterian animals corroborate the monophyly of ecdysozoa, lophotrochozoa, and protostomia. Mol Biol Evol 22:1246-1253.

Philippe H, Snell EA, Bapteste E, Lopez P, Holland PW, Casane D. 2004. Phylogenomics of eukaryotes: impact of missing data on large alignments. Mol Biol Evol 21:1740-1752.

Philippe H, Telford MJ. 2006. Large-scale sequencing and the new animal phylogeny. Trends Ecol Evol 21:614-620.

Phillips MJ, Delsuc F, Penny D. 2004. Genome-scale phylogeny and the detection of systematic biases. Mol Biol Evol 21:1455-1458.

Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-818.

Posada D, Crandall KA. 2001. Selecting the best-fit model of nucleotide substitution. Syst Biol 50:580-601.

Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu JK, Benito-Gutierrez EL, Dubchak I, Garcia-Fernandez J, Gibson-Brown JJ, Grigoriev IV, Horton AC, de Jong PJ, Jurka J, Kapitonov VV, Kohara Y, Kuroki Y, Lindquist E, Lucas S, Osoegawa K, Pennacchio LA, Salamov AA, Satou Y, Sauka-Spengler T, Schmutz J, Shin IT, Toyoda A, Bronner-Fraser M, Fujiyama A, Holland LZ, Holland PW, Satoh N, Rokhsar DS. 2008. The amphioxus genome and the evolution of the chordate karyotype. Nature 453:1064-1071.

R Development Core Team. 2007. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rodríguez-Ezpeleta N, Brinkmann H, Roure B, Lartillot N, Lang BF, Philippe H. 2007. Detecting and overcoming systematic errors in genome-scale phylogenies. Syst Biol 56:389-399.

Roure B, Rodriguez-Ezpeleta N, Philippe H. 2007. SCaFoS: a tool for selection, concatenation and fusion of sequences for phylogenomics. BMC Evol Biol 7 Suppl 1:S2.

Rowe T. 2004. Chordate phylogeny and development. In: Cracraft J, Donoghue MJ, editors. Assembling the Tree of Life. Oxford: Oxford University Press. p 384-409.

Ruppert EE. 2005. Key characters uniting hemichordates and chordates: homologies or homoplasies? Can. J. Zool. 83:8-23.

Schaeffer B. 1987. Deuterostome monophyly and phylogeny. Evol Biol 21:179-235.

Schmidt HA, Strimmer K, Vingron M, von Haeseler A. 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18:502-504.

Schubert M, Escriva H, Xavier-Neto J, Laudet V. 2006. Amphioxus and tunicates as evolutionary model systems. Trends Ecol Evol 21:269-277.

Sempere LF, Cole CN, McPeek MA, Peterson KJ. 2006. The phylogenetic distribution of metazoan microRNAs: insights into evolutionary complexity and constraint. J Exp Zoolog B Mol Dev Evol 306:575-588.

Seo HC, Edvardsen RB, Maeland AD, Bjordal M, Jensen MF, Hansen A, Flaat M, Weissenbach J, Lehrach H, Wincker P, Reinhardt R, Chourrout D. 2004. Hox cluster disintegration with persistent anteroposterior order of expression in *Oikopleura dioica*. Nature 431:67-71.

Steel M. 2005. Should phylogenetic models be trying to "fit an elephant"? Trends Genet 21:307-309.

Swalla BJ, Cameron CB, Corley LS, Garey JR. 2000. Urochordates are monophyletic within the deuterostomes. Syst Biol 49:52-64.

Swofford DL. 2002. PAUP*: Phylogenetic Analysis Using Parsimony and other methods version 4.0b10. In: Sinauer, Sunderland, MA.

Telford MJ, Budd GE. 2003. The place of phylogeny and cladistics in Evo-Devo research. Int J Dev Biol 47:479-490.

Vienne A, Pontarotti P. 2006. Metaphylogeny of 82 gene families sheds a new light on chordate evolution. Int J Biol Sci 2:32-37.

Wada H, Okuyama M, Satoh N, Zhang S. 2006. Molecular evolution of fibrillar collagen in chordates, with implications for the evolution of vertebrate skeletons and chordate phylogeny. Evol Dev 8:370-377.

Wada H, Satoh N. 1994. Details of the evolutionary history from invertebrates to vertebrates, as deduced from the sequences of 18S rDNA. Proc Natl Acad Sci USA 91:1801-1804.

Whelan S, Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18:691-699.

Wiens JJ. 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Syst Biol 52:528-538.

Wiens JJ. 2005. Can incomplete taxa rescue phylogenetic analyses from long-branch attraction? Syst Biol 54:731-742.

Wiens JJ. 2006. Missing data and the design of phylogenetic analyses. J Biomed Inform 39:34-42.

Winchell CJ, Sullivan J, Cameron CB, Swalla BJ, Mallatt J. 2002. Evaluating hypotheses of deuterostome phylogeny and chordate evolution with new LSU and SSU ribosomal DNA data. Mol Biol Evol 19:762-776.

FIGURE LEGENDS

FIG. 1: Reanalysis of previous phylogenomic data using an improved model of sequence evolution. The Delsuc *et al.* (2006) phylogenomic dataset of 146 genes (38 taxa and 33,800 sites) was analyzed under the CAT+Γ_4 site-heterogeneous mixture model of amino-acid replacement. Values at nodes represent Bayesian posterior probabilities (PP). Circles indicate nodes with maximal support PP =1.0. The scale bar represents the estimated number of substitutions per site.

FIG. 2: Phylogenetic analyses of an updated phylogenomic dataset. (**a**) Bayesian consensus tree obtained using the CAT+Γ_4 mixture model on the complete dataset based on the concatenation of 179 genes (51 taxa and 53,799 amino-acid sites) containing 32% missing data. (**b**) Bayesian inference using the CAT+Γ_4 mixture model on the dataset reduced to the concatenation of the 106 genes for which sequences were available for at least 41 of the 51 taxa (25,321 amino-acid sites) containing only 20% of missing data. Values at nodes indicate Bayesian posterior probabilities (PP) / Maximum-likelihood bootstrap percentages (BP; 100 replicates) obtained under the WAG+Γ_8. Circles indicate strongly supported nodes with PP $\geq$ 0.95 and BP $\geq$ 95. The scale bar represents the estimated number of substitutions per site.

FIG. 3: Assessing the robustness of phylogenetic results to gene sampling using a jackknife procedure. The Bayesian phylogenetic inference was conducted under the CAT+Γ_4 mixture model on 50 jackknife replicates of 50 genes over a total of 179. The tree presented is the weighted majority-rule consensus of all trees sampled every 10 cycles across the 50 replicates after removing the first 1000 trees in each MCMC as the burnin. Values at nodes represent corresponding jackknife-resampled posterior probabilities indices (PP_{JK}). Circles indicate highly repeatable nodes with $PP_{JK} \geq 95\%$. The scale bar represents the number of substitutions per site.

FIG. 4: New phylogenetic analyses of ribosomal RNA genes. (**a**) Principal component analysis of nucleotide composition of the combined 18S+28S rRNA dataset. The graph represents the projection of individuals on the first two axes, which explain more than 98% of the total variance. (**b**) Maximum-likelihood analyses of the 18S+28S dataset using the best-fitting standard DNA model (TIM+Γ_4+I) on nucleotides (left) and a two-state model (CF+Γ_8) after RY-coding of nucleotides (right). ML bootstrap percentages are given at nodes when greater than 70 except within vertebrates. Circles indicate strongly supported nodes with BP $\geq$

95. Squares points to shifting nodes of interest between the two ML trees. Scale bars represent the number of substitutions per site.

Figure 1

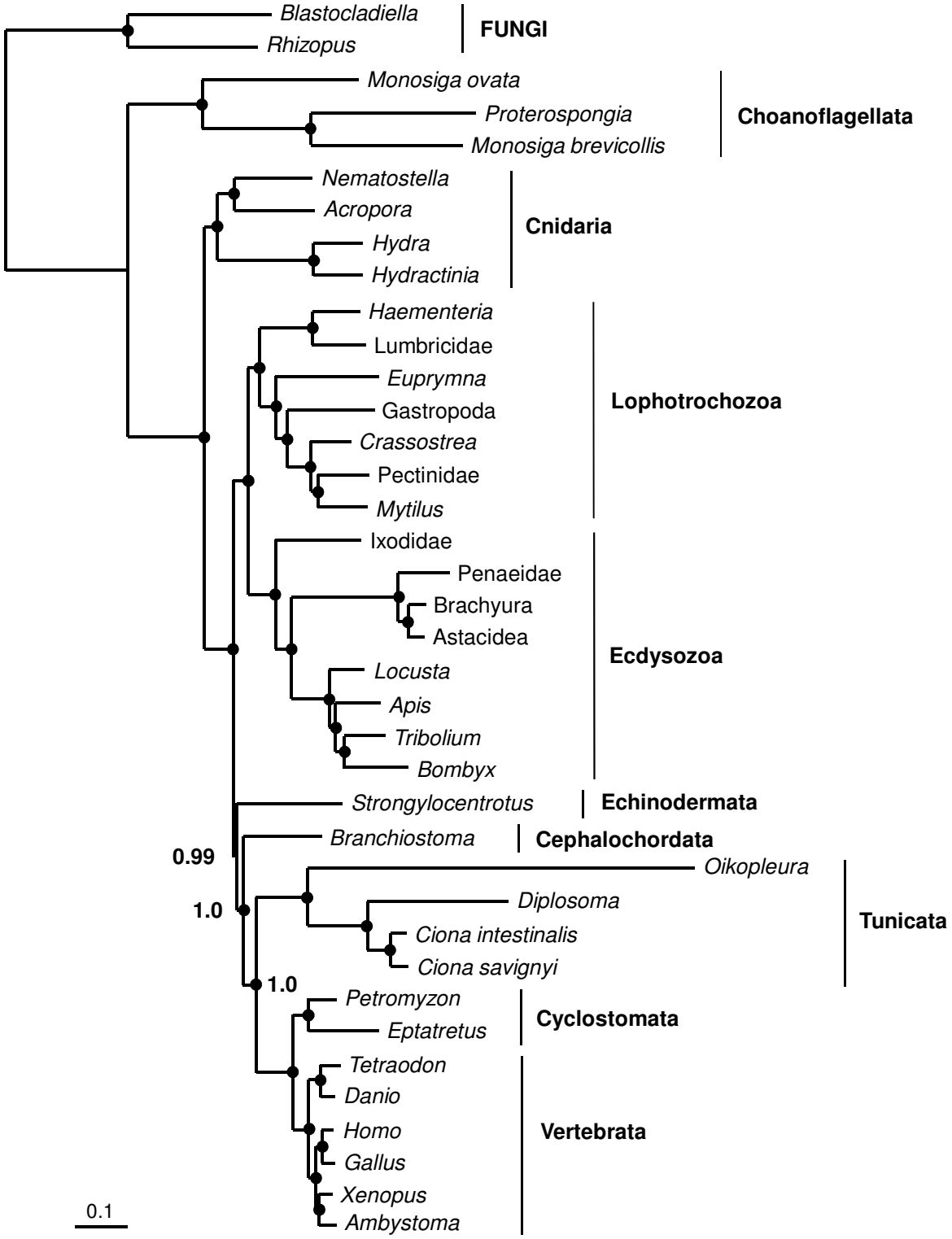

Blastocladiella
Rhizopus
FUNGI
Monosiga ovata
Proterospongia
Monosiga brevicollis
Choanoflagellata
Nematostella
Acropora
Hydra
Hydractinia
Cnidaria
Haementeria
Lumbricidae
Euprymna
Gastropoda
Crassostrea
Pectinidae
Mytilus
Lophotrochozoa
Ixodidae
Penaeidae
Brachyura
Astacidea
Locusta
Apis
Tribolium
Bombyx
Ecdysozoa
Strongylocentrotus
Echinodermata
Branchiostoma
Cephalochordata
0.99
Oikopleura
1.0
Diplosoma
Ciona intestinalis
Ciona savignyi
Tunicata
1.0
Petromyzon
Eptatretus
Cyclostomata
Tetraodon
Danio
Homo
Gallus
Xenopus
Ambystoma
Vertebrata
0.1

Figure 2

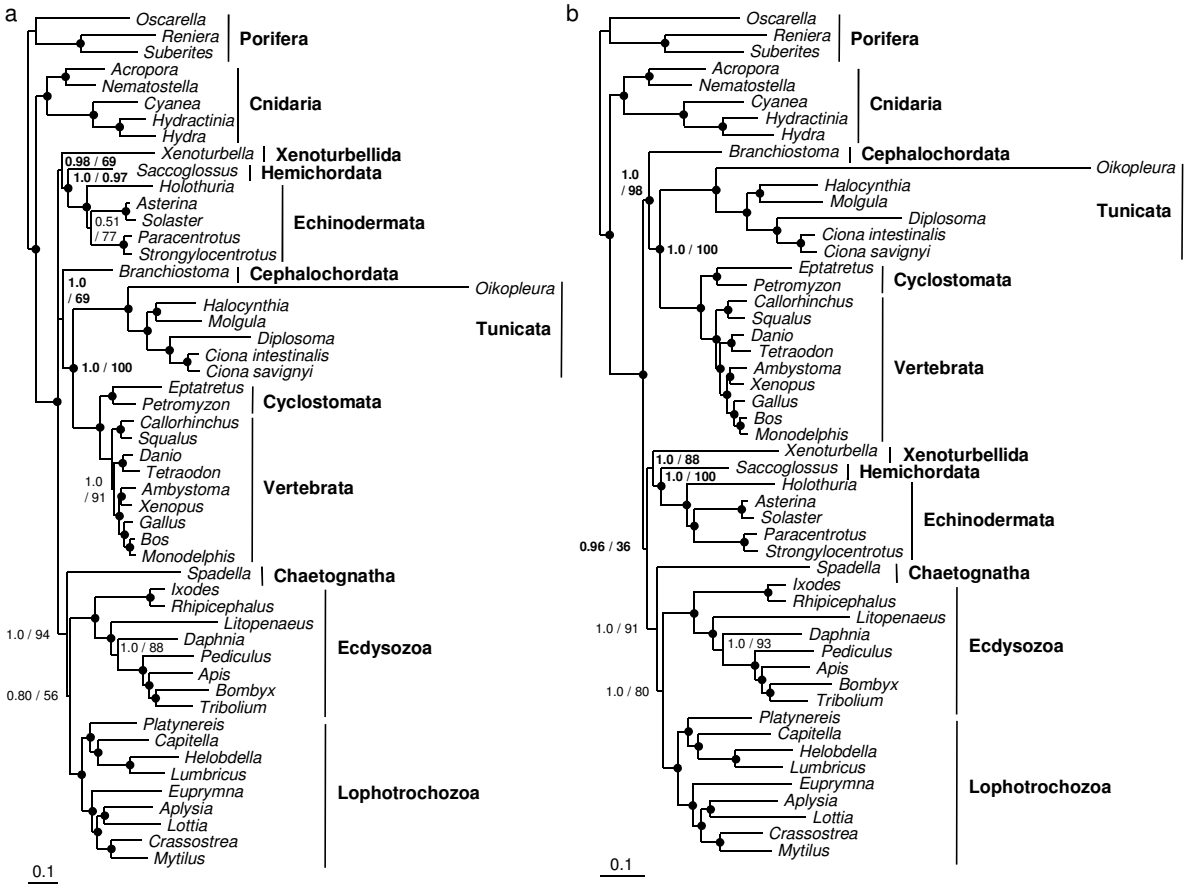

a
Oscarella
Reniera
Suberites
Porifera
Acropora
Nematostella
Cyanea
Hydractinia
Hydra
Cnidaria
0.98 / 69
Xenoturbella
Xenoturbellida
1.0 / 0.97
Saccoglossus
Hemichordata
Holothuria
Asterina
0.51
Solaster
/ 77
Paracentrotus
Strongylocentrotus
Echinodermata
Branchiostoma
Cephalochordata
1.0
/ 69
Oikopleura
Halocynthia
Molgula
Tunicata
Diplosoma
Ciona intestinalis
Ciona savignyi
1.0 / 100
Eptatretus
Petromyzon
Cyclostomata
Callorhinchus
Squalus
Danio
Tetraodon
1.0
/ 91
Ambystoma
Xenopus
Gallus
Bos
Monodelphis
Vertebrata
Spadella
Chaetognatha
Ixodes
Rhipicephalus
Litopenaeus
1.0 / 94
Daphnia
1.0 / 88
Pediculus
Apis
0.80 / 56
Bombyx
Tribolium
Ecdysozoa
Platynereis
Capitella
Helobdella
Lumbricus
Euprymna
Aplysia
Lottia
Crassostrea
Mytilus
Lophotrochozoa
0.1
b
Oscarella
Reniera
Suberites
Porifera
Acropora
Nematostella
Cyanea
Hydractinia
Hydra
Cnidaria
Branchiostoma
Cephalochordata
Oikopleura
1.0
/ 98
Halocynthia
Molgula
Tunicata
Diplosoma
Ciona intestinalis
Ciona savignyi
1.0 / 100
Eptatretus
Petromyzon
Cyclostomata
Callorhinchus
Squalus
Danio
Tetraodon
Ambystoma
Xenopus
Gallus
Bos
Monodelphis
Vertebrata
Xenoturbella
Xenoturbellida
1.0 / 88
Saccoglossus
Hemichordata
1.0 / 100
Holothuria
Asterina
Solaster
Paracentrotus
Strongylocentrotus
Echinodermata
0.96 / 36
Spadella
Chaetognatha
Ixodes
Rhipicephalus
Litopenaeus
1.0 / 91
Daphnia
1.0 / 93
Pediculus
Apis
Bombyx
1.0 / 80
Tribolium
Ecdysozoa
Platynereis
Capitella
Helobdella
Lumbricus
Euprymna
Aplysia
Lottia
Crassostrea
Mytilus
Lophotrochozoa
0.1

Figure 3

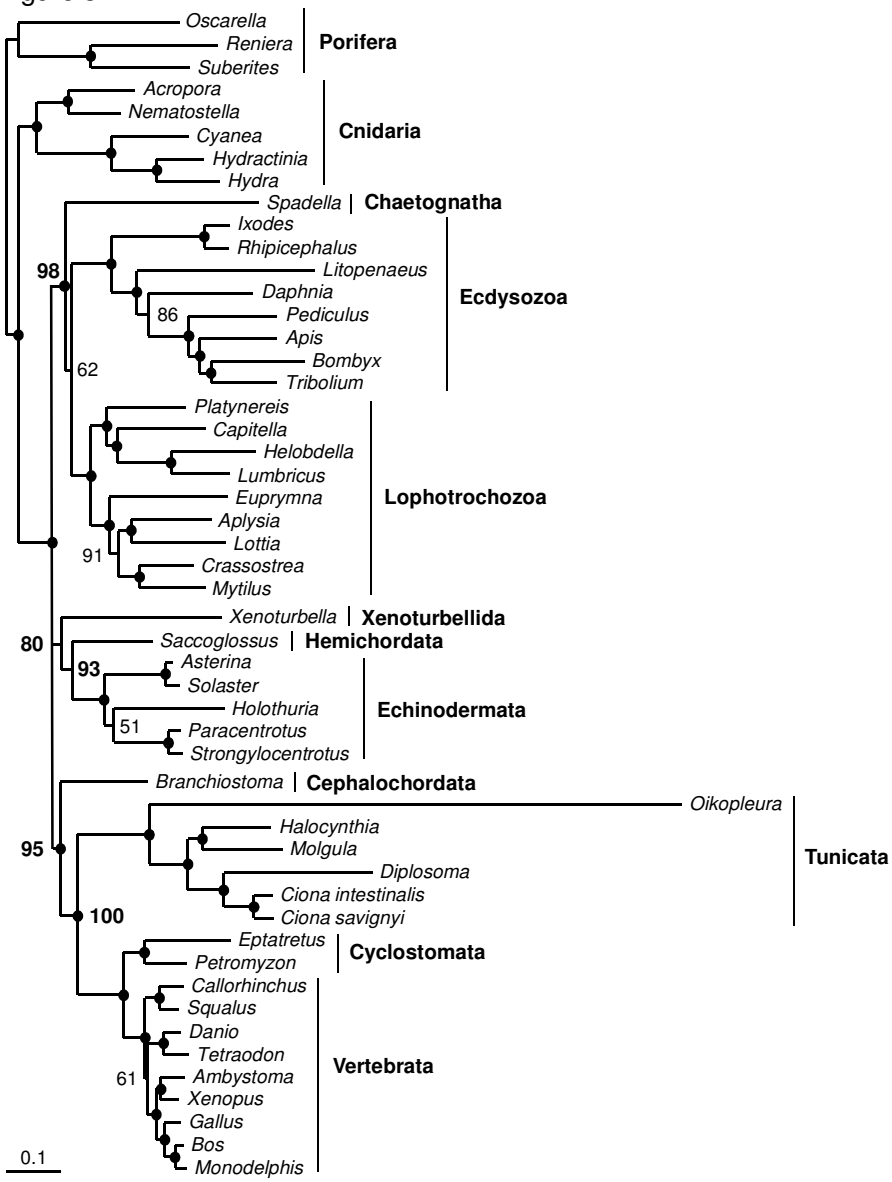
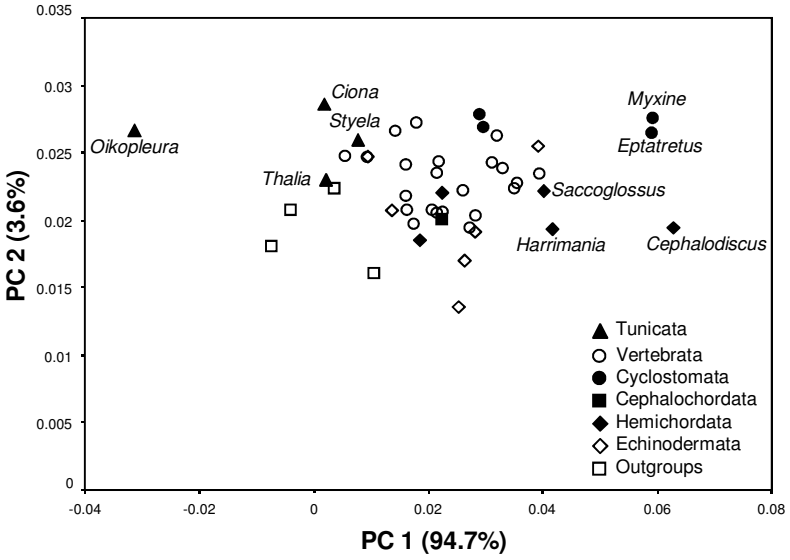
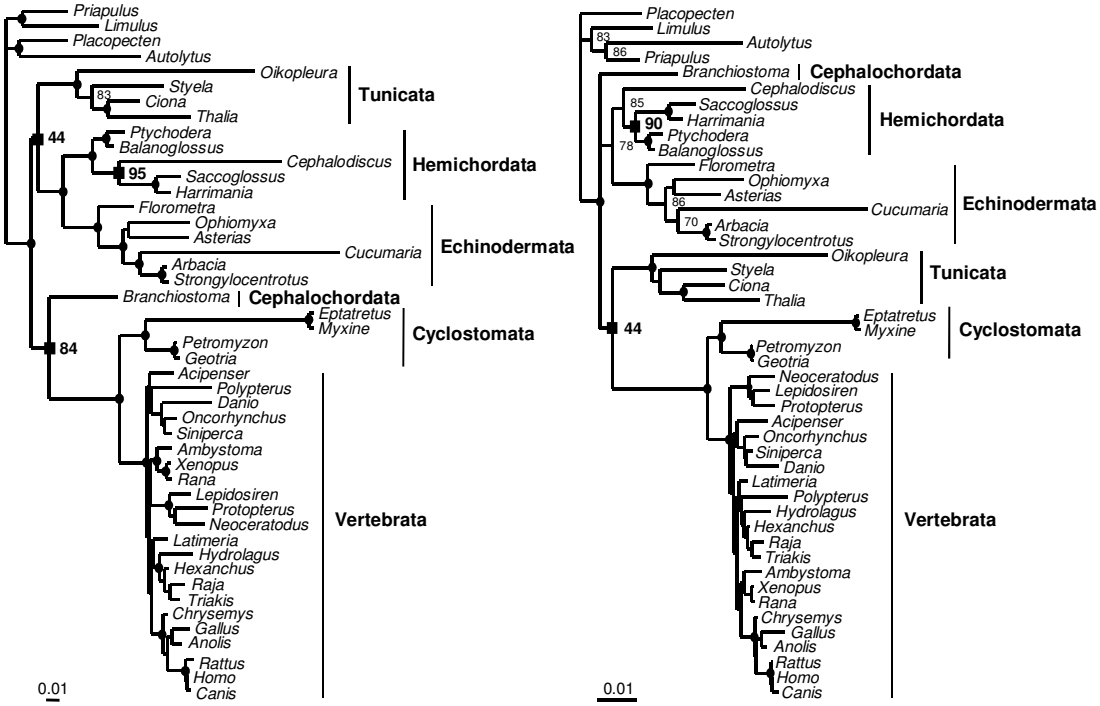

Oscarella
Reniera
Suberites
Porifera
Acropora
Nematostella
Cyanea
Hydractinia
Hydra
Cnidaria
Spadella
Chaetognatha
Ixodes
Rhipicephalus
Litopenaeus
Daphnia
Pediculus
Apis
Bombyx
Tribolium
Ecdysozoa
98
86
62
Platynereis
Capitella
Helobdella
Lumbricus
Euprymna
Aplysia
Lottia
Crassostrea
Mytilus
91
Lophotrochozoa
Xenoturbella
Xenoturbellida
80
Saccoglossus
Hemichordata
93
Asterina
Solaster
Holothuria
Paracentrotus
Strongylocentrotus
51
Echinodermata
Branchiostoma
Cephalochordata
Oikopleura
Halocynthia
Molgula
Diplosoma
Ciona intestinalis
Ciona savignyi
Tunicata
95
100
Eptatretus
Petromyzon
Cyclostomata
Callorhinchus
Squalus
Danio
Tetraodon
Ambystoma
Xenopus
Gallus
Bos
Monodelphis
61
Vertebrata
0.1

Figure 4

a


PC 2 (3.6%)
PC 1 (94.7%)
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
-0.04
-0.02
0
0.02
0.04
0.06
0.08
Ciona
Styela
Myxine
Oikopleura
Eptatretus
Thalia
Saccoglossus
Harrimania
Cephalodiscus
Tunicata
Vertebrata
Cyclostomata
Cephalochordata
Hemichordata
Echinodermata
Outgroups

b


Priapulus
Limulus
Placopecten
Autolytus
Oikopleura
Styela
Ciona
Thalia
Tunicata
Ptychodera
Balanoglossus
Cephalodiscus
Saccoglossus
Harrimania
Hemichordata
Florometra
Ophiomyxa
Asterias
Cucumaria
Arbacia
Strongylocentrotus
Echinodermata
Branchiostoma | Cephalochordata
Eptatretus
Myxine
Petromyzon
Geotria
Cyclostomata
Acipenser
Polypterus
Danio
Oncorhynchus
Siniperca
Ambystoma
Xenopus
Rana
Lepidosiren
Protopterus
Neoceratodus
Latimeria
Hydrolagus
Hexanchus
Raja
Triakis
Chrysemys
Gallus
Anolis
Rattus
Homo
Canis
Vertebrata
0.01
Placopecten
Limulus
Autolytus
Priapulus
Branchiostoma | Cephalochordata
Cephalodiscus
Saccoglossus
Harrimania
Ptychodera
Balanoglossus
Hemichordata
Florometra
Ophiomyxa
Asterias
Cucumaria
Arbacia
Strongylocentrotus
Echinodermata
Oikopleura
Styela
Ciona
Thalia
Tunicata
Eptatretus
Myxine
Petromyzon
Geotria
Cyclostomata
Neoceratodus
Lepidosiren
Protopterus
Acipenser
Oncorhynchus
Siniperca
Danio
Latimeria
Polypterus
Hydrolagus
Hexanchus
Raja
Triakis
Ambystoma
Xenopus
Rana
Chrysemys
Gallus
Anolis
Rattus
Homo
Canis
Vertebrata
0.01

Additional molecular evidence for the new chordate phylogeny

Frédéric Delsuc, Georgia Tsagkogeorga, Nicolas Lartillot & Hervé Philippe

Supplementary Material

Table S1: List of the 179 genes with names and numbers of amino-acid positions conserved for each gene alignment.

Gene Abbreviation	Gene name	Amino-acid positions
ar21	Actin-related protein 2/3 complex subunit 3	155
arc20	Actin-related protein 2/3 complex subunit 4	166
arp23	Actin-related protein 2/3 complex subunit 1b	200
atpsynthalpha-a-mt	F0F1 ATP synthase subunit alpha	464
cct-A	T complex protein 1 alpha subunit	515
cct-B	T complex protein 1 beta subunit	513
cct-D	T complex protein 1 delta subunit	489
cct-E	T complex protein 1 epsilon subunit	526
cct-G	T complex protein 1 gamma subunit	512
cct-N	T complex protein 1 eta subunit	528
cct-T	T complex protein 1 theta subunit	484
cct-Z	T complex protein 1 ? subunit	499
cpn60-mt	Heat shock protein HSP 60kDa mitochondrial	489
crfg	Nucleolar GTP binding protein 1	404
ef1-RF3	Release factor RF3	417
ef1-RFS	EF1alpha-like	399
ef2-EF2	Elongation factor EF2	806
ef2-EF4	EF2-like	567
ef2-U5	Elongation factor Tu family U5 snRNP specific protein	914
fibri	Fibrillarin	228
fpps	Farnesyl pyrophosphate synthase	250
glcn	N-acetyl glucosamine phophotransferase	331
grc5	60S ribosomal protein L10 QM protein	214
hsp70-E	Heat shock 70kDa protein form ER	594
hsp70-mt	Heat shock 70kDa protein, mitochondrial form	600
hsp90-C	Heat shock 90kDa protein cytosolic form	641
hsp90-E	Heat shock 90kDa protein form ER	673
hsp90-Z	TNF receptor-associated protein 1	474
if1a	Eukaryotic translation initiation factor 1a	118
if2b	Eukaryotic translation initiation factor 2b	194
if2g	Eukaryotic translation initiation factor 2g	451
if2p	Eukaryotic translation initiation factor 2p	590
if4a-a	Translation initiation factor 4A	390
if4a-b	Translation initiation factor 4A	369
if6	Eukaryotic translation initiation factor 6	239
ino1	Myo-inositol-1-phosphate synthase	442
l12e-A	40S ribosomal Protein S12	123
l12e-B	High mobility group like nuclear protein 2 NHP2	127
l12e-C	High mobility group like nuclear protein 2 NHP2-like protein 1	124
l12e-D	60S ribosomal Protein L7a	246
limif	Small CTD phosphatase	186

mcm-A	minichromosome family maintenance protein 5	617
mcm-B	minichromosome family maintenance protein 2	708
mcm-C	minichromosome family maintenance protein 3	525
mcm-D	minichromosome family maintenance protein 7	641
mcm-E	minichromosome family maintenance protein 4	627
mcm-F	minichromosome family maintenance protein 6	686
mcm-G	minichromosome family maintenance protein 9	398
mcm-H	minichromosome family maintenance protein 8	461
mra1	Ribosome biogenesis protein NEP1 C2F protein	187
nsf1-C	Vacuolar protein sorting factor 4b	400
nsf1-G	26S proteasome AAA-ATPase regulatory subunit 8	385
nsf1-I	putative 26S proteasome ATPase regulatory subunit 7	422
nsf1-J	26S proteasome AAA-ATPase regulatory subunit 6	385
nsf1-K	26S proteasome AAA-ATPase regulatory subunit 6a	403
nsf1-L	26S proteasome AAA-ATPase regulatory subunit 6b	380
nsf1-M	26S proteasome AAA-ATPase regulatory subunit 4	441
nsf1-N	Katanin p60 subunit A	265
nsf2-A	Transitional endoplasmic reticulum ATPase TER ATPase	762
nsf2-B	Nuclear VCP-like	451
nsf2-F	Vesicular fusion protein nsf2	497
orf2	putative 28 kDa protein	177
ornamtrans-a	Ornithine aminotransferase	371
pace2-A	XPA binding protein 1	209
pace2-B	Conserved hypothetical ATP binding protein	239
pace2-C	Conserved hypothetical ATP binding protein	229
pace4	protein chromosome 2 ORF 4	230
pace5-A	Shwachman-Badian-Diamond syndrome protein	237
pace6	programmed cell death protein 5	106
psma-A	20S proteasome beta subunit macropain zeta chain	225
psma-B	20S proteasome alpha 1a chain	242
psma-C	20S proteasome alpha 1b chain	252
psma-D	20S proteasome alpha 2 chain	228
psma-E	20S proteasome alpha 1c chain	216
psma-F	20S proteasome alpha 3 chain	241
psma-G	20S proteasome alpha 6 chain	241
psmb-H	20S proteasome alpha 1d chain	185
psmb-I	20S proteasome alpha 1e chain	205
psmb-J	20S proteasome alpha 1f chain	212
psmb-K	20S proteasome beta 7 chain	234
psmb-L	20S proteasome beta 6 chain	196
psmb-M	20S proteasome beta 5 chain	195
psmb-N	20S proteasome beta 4 chain	194
pyrdehydroe1b-B	Branched chain ketoacid dehydrogenase E1 beta	321
pyrdehydroe1b-mt	Pyruvate dehydrogenase E1 beta subunit	322
rad23	UV excision repair protein RAD23	216
rad51-A	DNA repair protein RAD51	315
rad51-B	DNA repair protein DMC1	320
rf1	Eukaryotic peptide chain release factor subunit 1	402
rla2-A	60S acidic ribosomal protein P2	92
rla2-B	60S acidic ribosomal protein P1	76
rpl1	60S ribosomal Protein 1	213
rpl11b	60S ribosomal Protein 11b	169
rpl12b	60S ribosomal Protein 12b	163
rpl13	60S ribosomal Protein 13	179
rpl14a	60S ribosomal Protein 14a	123

rpl15a	60S ribosomal Protein 15a	204
rpl16b	60S ribosomal Protein 16b	173
rpl17	60S ribosomal Protein 17	175
rpl18	60S ribosomal Protein 18	183
rpl19a	60S ribosomal Protein 19a	190
rpl2	60S ribosomal Protein 2	248
rpl20	60S ribosomal Protein 20	160
rpl21	60S ribosomal Protein 21	154
rpl22	60S ribosomal Protein 22	96
rpl23a	60S ribosomal Protein 23a	139
rpl24-A	60S ribosomal Protein 24a	123
rpl24-B	60S ribosomal Protein 24b	137
rpl25	60S ribosomal Protein 25	126
rpl26	60S ribosomal Protein 26	135
rpl27	60S ribosomal Protein 27	137
rpl3	60S ribosomal Protein 3	393
rpl30	60S ribosomal Protein 30	105
rpl31	60S ribosomal Protein 31	108
rpl32	60S ribosomal Protein 32	129
rpl33a	60S ribosomal Protein 33a	106
rpl34	60S ribosomal Protein 34	108
rpl35	60S ribosomal Protein 35	123
rpl36	60S ribosomal Protein 36	85
rpl37a	60S ribosomal Protein 37a	81
rpl38	60S ribosomal Protein 38	65
rpl39	60S ribosomal Protein 39	51
rpl42	60S ribosomal Protein 4	102
rpl43b	60S ribosomal Protein 43b	88
rpl4B	60S ribosomal Protein 4b	326
rpl5	60S ribosomal Protein 5	275
rpl6	60S ribosomal Protein 6	186
rpl7-A	60S ribosomal Protein 7a	207
rpl9	60S ribosomal Protein 9	183
rpo-A	RNA polymerase alpha subunit	684
rpo-B	RNA polymerase beta subunit	1419
rpo-C	RNA polymerase gamma subunit	1091
rpp0	60S acidic ribosomal protein P0 L10E	289
rps1	40S ribosomal Protein 1	244
rps10	40S ribosomal Protein 10	119
rps11	40S ribosomal Protein 11	140
rps13a	40S ribosomal Protein 13a	151
rps14	40S ribosomal Protein 14	149
rps15	40S ribosomal Protein 15	139
rps16	40S ribosomal Protein 16	138
rps17	40S ribosomal Protein 17	119
rps18	40S ribosomal Protein 18	152
rps19	40S ribosomal Protein 19	132
rps20	40S ribosomal Protein 20	106
rps22a	40S ribosomal Protein 22a	130
rps23	40S ribosomal Protein 23	143
rps24	40S ribosomal Protein 24	122
rps25	40S ribosomal Protein 25	92
rps26	40S ribosomal Protein 26	101
rps27	40S ribosomal Protein 27	84
rps27a	40S ribosomal Protein 27a	155

rps28a	40S ribosomal Protein 28a	61
rps29	40S ribosomal Protein 29	56
rrp46-A	Exosome component 5	165
rrp46-B	Exosome complex exonuclease Rrp41	210
sadhchydrolase-A	Adenosylhomocysteinase 89E	455
sadhchydrolase-E1	S-adenosylhomocysteine hydrolase	424
sap40	40S ribosomal protein SA 40kDa laminin receptor 1	212
Sra	Signal recognition particle receptor alpha subunit SR alpha	422
srp54	Signal recognition particle 54 kDa protein	492
Srs	Seryl tRNA synthetase	454
stbproptase2a-b	Protein phosphatase-2A catalytic subunit beta	302
stcproptase2a-c	Protein phosphatase 6	294
Suca	Succinyl-CoA ligase alpha chain mitochondrial precursor?	291
Tfiid	TATA box binding protein related factor 2	175
tif2a	Translation initiation factor 2 alpha	304
topo1	DNA topoisomerase I, mitochondrial precursor	520
u2snrnp	Splicing factor U2AF 35 kDa subunit	179
vacaatpasepl21-a	V-type ATP synthase subunit K	147
Vata	Vacuolar ATP synthase catalytic subunit A	610
Vatb	Vacuolar ATP synthase catalytic subunit B	479
Vatc	Vacuolar ATP synthase catalytic subunit C	336
Vate	Vacuolar ATP synthase catalytic subunit E	200
Vatpased	ATPase H+ transporting V1 subunit D	210
vdac2	Voltage-dependent anion channel 2	276
w09c	TGF beta inducible nuclear protein	260
Wrs	tryptophanyl-tRNA synthetase	379
Xpb	Helicase XPB subunit 2	631
yif1p	homolog of Yeast Golgi membrane protein	188

Table S2: List of Operational Taxonomic Units (OTUs) and Chimerical Sequences
In the manuscript, OTUs were indicated by the genus name of the most frequent species indicated (underlined for chimerical OTUs).

Echinodermata
Strongylocentrotus: *Strongylocentrotus purpuratus*
Paracentrotus: *Paracentrotus lividus*
Asterina: *Asterina pectinifera*
Holothuria: *Holothuria glaberrima*, *Apostichopus japonicus*
Solaster: *Solaster stimpsonii*

Hemichordata
Saccoglossus: *Saccoglossus kowalevskii*
Xenoturbella: *Xenoturbella bocki*

Cephalochordata
Branchiostoma: *Branchiostoma floridae*, *Branchiostoma belcheri*, *Branchiostoma lanceolatum*

Tunicata
Ciona savignyi: *Ciona savignyi*
Ciona intestinalis: *Ciona intestinalis*
Oikopleura: *Oikopleura dioica*
Molgula: *Molgula tectiformis*
Halocynthia: *Halocynthia roretzi*
Diplosoma: *Diplosoma listerianum*

Myxini
Eptatretus: *Eptatretus burgeri*, *Myxine glutinosa*

Petromyzontiformes
Petromyzon: *Petromyzon marinus*

Chondrichthyes
Squalus: *Squalus acanthias*
Callorhinchus: *Callorhinchus milii*

Actinopterygii
Danio: *Danio rerio*
Tetraodon: *Tetraodon nigroviridis*

Amphibia
Xenopus: *Xenopus tropicalis*, *Xenopus laevis*
Ambystoma: *Ambystoma mexicanum*, *Ambystoma tigrinum*

Aves
Gallus: *Gallus gallus*

Metatheria
Monodelphis: *Monodelphis domestica*

Placentalia
Bos: *Bos taurus*, *Canis familiaris*, *Homo sapiens*, *Macaca mulatta*, *Pan troglodytes*, *Pongo pygmaeus*, *Macaca fascicularis*, *Mus musculus*, *Rattus norvegicus*

Annelida
Capitella: *Capitella species-2004*
Helobdella: *Helobdella robusta*, *Haementeria depressa*
Lumbricus: *Lumbricus rubellus*, *Eisenia fetida*, *Eisenia andrei*
Platynereis: *Platynereis dumerilii*

Mollusca
Aplysia: *Aplysia californica*
Lottia: *Lottia gigantea*

Euprymna: *Euprymna scolopes*, *Idiosepius paradoxus*
Crassostrea: *Crassostrea virginica*, *Crassostrea gigas*
Mytilus: *Mytilus galloprovincialis*, *Mytilus californianus*, *Mytilus edulis*

Chaetognatha
Spadella: *Spadella cephaloptera*, *Flaccisagitta enflata*

#Arthropoda
Ixodes: *Ixodes scapularis*, *Ixodes pacificus*
Rhipicephalus: *Rhipicephalus microplus*, *Rhipicephalus appendiculatus*
Litopenaeus: *Litopenaeus vannamei*, *Litopenaeus setiferus*, *Penaeus monodon*, *Fenneropenaeus chinensis*, *Marsupenaeus japonicus*
Daphnia: *Daphnia pulex*, *Daphnia magna*
Tribolium: *Tribolium castaneum*
Apis: *Apis mellifera*
Bombyx: *Bombyx mori*
Pediculus: *Pediculus humanus*

Cnidaria
Nematostella: *Nematostella vectensis*
Hydra: *Hydra magnipapillata*, *Hydra vulgaris*
Acropora: *Acropora millepora*, *Acropora palmata*, *Montastraea faveolata*
Hydractinia: *Hydractinia echinata*, *Podocoryne carnea*
Cyanea: *Cyanea capillata*

Porifera
Reniera: *Reniera sp.*
Oscarella: *Oscarella carmela*, *Oscarella lobularis*, *Oscarella sp.*
Suberites: *Suberites domuncula*, *Suberites fuscus*

Table S3: Percentages of missing data in the 179-genes supermatrix

		ar21	arc20	arp23	atpsynthalpha-a-mt	cct-A	cct-B	cct-D	cct-E	cct-G	cct-N	cct-T	cct-Z	cpn60-mt	crfg	ef1-RF3	ef1-RFS	ef2-EF2	ef2-EF4	ef2-U5	fibri	fpps	glcn	grc5	hsp70-E	hsp70-mt	hsp90-C	hsp90-E	hsp90-Z	if1a	if2b	if2g	if2p
Acropora	*millepora*	0	0	4	6	100	73	25	52	53	64	70	13	46	100	100	100	45	100	100	55	62	86	0	47	31	0	25	100	0	53	33	100
Ambystoma	*mexicanum*	0	0	0	0	0	0	1	0	43	0	26	6	4	12	45	18	7	81	42	0	12	100	0	43	0	100	8	17	100	0	18	62
Apis	*mellifera*	1	0	0	0	0	0	0	0	0	0	0	42	0	0	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	13	0	0
Aplysia	*californica*	1	0	0	0	8	0	0	5	33	1	3	22	5	0	38	5	0	37	26	5	16	76	2	0	4	0	1	60	1	1	0	62
Asterina	*pectinifera*	100	0	12	100	55	100	100	66	100	100	57	64	100	100	100	100	39	100	100	100	100	100	0	80	100	67	100	60	0	100	100	100
Bombyx	*mori*	0	0	0	0	0	0	0	0	16	0	0	0	0	0	1	27	0	22	19	0	0	19	0	0	0	0	2	5	0	1	0	68
Bos	*taurus*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Branchiostoma	*floridae*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	6	0	0	0	0	1	0	0	2	0	6
Callorhinchus	*milii*	72	100	100	18	22	48	49	25	100	32	44	31	38	21	73	49	58	43	44	59	100	100	100	65	26	81	28	43	100	100	37	42
Capitella	*sp-2004*	12	0	22	0	3	0	4	12	5	2	5	0	0	14	4	0	0	2	2	2	1	21	6	0	0	0	9	32	34	10	2	6
Ciona	*intestinalis*	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	14	0	2	100	100	4	18	36	3	0	2	0
Ciona	*savignyi*	0	0	0	0	0	0	0	0	0	0	1	0	0	8	1	10	0	10	1	0	0	0	1	0	0	0	11	32	0	10	2	0
Crassostrea	*virginica*	1	0	0	3	33	19	74	23	82	13	27	3	66	55	100	100	4	100	100	100	100	100	1	0	100	10	0	80	1	23	68	100
Cyanea	*capillata*	25	100	100	100	100	100	100	100	100	100	100	78	100	100	100	100	89	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Danio	*rerio*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Daphnia	*pulex*	0	0	0	0	0	0	0	0	5	2	1	5	0	0	1	0	0	1	0	1	0	1	1	0	0	0	0	5	0	1	1	0
Diplosoma	*listerianum*	100	100	17	67	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	2	100	100	100	100	100	100	100	100	100
Eptatretus	*burgeri*	0	0	0	1	0	47	9	18	13	1	2	0	100	78	100	35	38	100	59	4	4	100	1	77	40	7	100	49	0	4	10	11
Euprymna	*scolopes*	15	1	0	8	9	27	61	73	64	44	65	34	67	5	37	49	8	32	56	7	82	100	1	68	10	5	82	30	1	11	60	52
Gallus	*gallus*	0	23	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	100	0	0	2	0	0	0	0	0	4	0	0	0
Halocynthia	*roretzi*	0	0	0	0	9	100	19	4	8	24	35	6	55	40	100	52	1	100	36	28	18	24	100	4	13	0	24	10	0	3	2	19
Helobdella	*robusta*	14	0	0	6	16	42	14	12	16	2	49	21	46	13	26	30	0	44	9	2	54	29	13	0	29	0	10	73	33	30	52	19
Holothuria	*glaberrima*	100	100	100	41	100	100	100	100	100	100	100	100	83	100	100	100	81	100	100	100	100	100	0	100	87	82	100	100	100	100	100	100
Hydra	*magnipapillata*	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	38	0	4	10	0	1	37	0	0	0	0	0	14	1	4	2	0
Hydractinia	*echinata*	100	100	0	10	17	39	23	100	35	49	30	6	47	43	100	100	0	100	67	56	100	100	2	43	100	1	36	100	100	100	16	100
Ixodes	*scapularis*	26	24	18	0	27	14	0	0	0	1	1	1	7	0	8	0	2	33	24	18	0	0	0	0	16	0	12	4	0	1	0	19
Litopenaeus	*vannamei*	1	3	20	29	66	1	1	45	29	1	0	21	66	100	100	100	3	100	100	0	100	100	0	1	61	8	100	100	0	4	12	100
Lottia	*gigantea*	100	23	18	0	29	32	32	12	52	5	63	35	82	45	1	50	10	0	66	5	57	32	16	3	42	0	10	69	10	30	21	45
Lumbricus	*rubellus*	0	0	13	0	27	100	34	72	60	64	37	70	39	80	100	100	0	100	85	100	100	100	1	33	100	17	100	100	0	46	100	100
Molgula	*tectiformis*	1	0	0	0	0	0	0	0	0	13	0	5	0	2	49	58	0	42	22	0	66	7	2	0	3	0	0	22	0	9	14	7
Monodelphis	*domestica*	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	12	0	0	0	0	0	0	0	3	0	0	100
Mytilus	*galloprovinciali*	1	100	0	24	100	11	51	53	15	67	100	0	64	100	100	8	25	100	100	21	100	100	2	44	80	5	100	100	100	100	57	100
Nematostella	*vectensis*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	7
Oikopleura	*dioica*	11	31	36	0	1	0	0	2	2	14	13	1	0	0	2	28	11	30	3	1	24	36	10	0	0	7	2	29	27	100	10	11
Oscarella	*carmela*	4	100	16	54	48	52	28	24	100	21	100	1	23	100	30	54	0	100	100	0	100	100	0	0	6	0	6	100	0	100	100	100
Paracentrotus	*lividus*	0	0	0	11	39	4	49	0	0	0	0	45	0	13	1	22	0	83	18	0	71	39	0	0	0	4	0	44	0	0	43	0
Pediculus	*humanus*	2	1	0	0	3	0	6	5	5	0	2	0	0	1	2	9	14	1	0	0	5	0	0	0	0	0	1	0	0	1	0	4
Petromyzon	*marinus*	0	0	0	0	78	10	0	0	37	0	100	100	0	21	36	29	0	16	55	3	0	28	0	3	17	15	0	100	100	17	8	55
Platynereis	*dumerilii*	100	100	0	28	100	42	10	64	45	100	100	44	13	30	100	100	35	100	100	23	100	100	100	0	17	17	100	100	0	0	73	100
Reniera	*sp.*	1	0	0	0	0	0	0	0	0	1	6	0	0	0	4	1	0	17	2	0	11	0	2	0	0	0	18	100	0	4	7	0
Rhipicephalus	*microplus*	0	0	100	0	0	0	0	0	0	1	1	21	13	0	69	48	0	100	41	0	6	24	0	0	0	16	0	6	0	1	1	56
Saccoglossus	*kowalevskii*	0	0	0	0	29	0	0	9	0	0	0	27	0	12	13	43	0	19	16	0	59	32	2	3	0	2	15	23	0	0	4	42
Solaster	*stimpsonii*	100	0	0	45	61	50	41	72	100	76	69	100	100	100	100	100	68	100	100	0	100	100	0	100	100	60	100	100	100	100	100	100
Spadella	*cephaloptera*	3	0	100	75	100	100	100	100	83	100	100	100	100	100	100	100	100	100	100	100	100	100	2	83	100	100	100	100	100	100	100	100
Squalus	*acanthias*	0	0	25	23	58	38	60	61	26	30	62	44	37	44	100	50	47	87	57	34	100	100	16	64	12	100	60	52	100	67	58	55
Strongylocentrotus	*purpuratus*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	44	0	0	14	0	0	0	0	0	0	0	0	0	0	1	0	0
Suberites	*domuncula*	0	0	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0	0	100	17	100	100	100	100	73	100
Tetraodon	*nigroviridis*	0	0	0	0	5	0	0	0	100	0	0	0	0	0	0	0	13	32	100	0	0	6	0	0	31	17	0	6	1	0	0	0
Tribolium	*castaneum*	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	10	0	0	0	100	0	1	0	0
Xenopus	*tropicalis*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	47	0	0	0
Xenoturbella	*bocki*	0	0	100	34	100	9	75	100	47	100	42	0	40	17	100	100	0	100	100	100	100	100	2	7	69	0	82	50	100	100	37	100
% missing positions		**17**	**20**	**18**	**15**	**30**	**27**	**25**	**28**	**35**	**26**	**32**	**25**	**30**	**32**	**44**	**44**	**18**	**50**	**45**	**26**	**45**	**49**	**10**	**23**	**31**	**19**	**33**	**48**	**29**	**32**	**30**	**46**
% missing OTUs		**14**	**18**	**14**	**6**	**18**	**16**	**12**	**14**	**20**	**16**	**18**	**12**	**14**	**22**	**35**	**29**	**6**	**37**	**29**	**20**	**33**	**39**	**8**	**10**	**20**	**10**	**24**	**31**	**25**	**25**	**16**	**33**
% chimeras	**chimeras**	**0**	**0**	**2**	**4**	**2**	**6**	**4**	**4**	**6**	**6**	**6**	**6**	**4**	**2**	**4**	**2**	**4**	**0**	**4**	**0**	**0**	**0**	**2**	**8**	**2**	**2**	**2**	**6**	**0**	**0**	**4**	**2**
# amino-acid sites	**length**	**155**	**166**	**200**	**464**	**515**	**513**	**489**	**526**	**512**	**528**	**484**	**499**	**489**	**404**	**417**	**399**	**806**	**567**	**914**	**228**	**250**	**331**	**214**	**594**	**600**	**641**	**673**	**474**	**118**	**194**	**451**	**590**

		if6	ino1	l12e-	l12e-	l12e-	l12e-	limif	mcm	mcm	mcm	mcm	mcm	mcm	mcm	mcm	mra1	nsf1-	nsf1-	nsf1-	nsf1-	nsf1-	nsf1-	nsf1-	nsf1-	nsf2-	nsf2-	nsf2-	orf2	orna	pace	pace	pace	pace	pace
Acropora	*millepora*	100	87	0	100	0	5	100	79	70	70	44	76	14	100	100	0	74	62	43	0	100	0	7	100	55	100	100	100	59	100	100	100	11	100
Ambystoma	*mexicanum*	40	59	1	0	0	0	100	100	51	52	5	10	28	100	100	100	25	9	35	38	61	25	35	100	19	100	71	21	28	70	0	100	100	25
Apis	*mellifera*	0	0	0	0	0	0	100	0	0	0	0	33	0	2	3	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	9	0
Aplysia	*californica*	24	0	4	100	0	0	33	18	19	42	17	8	32	43	79	43	36	5	6	17	0	0	6	5	17	100	29	1	4	0	44	20	79	54
Asterina	*pectinifera*	100	100	2	0	0	1	100	100	100	66	5	100	100	100	100	9	100	16	31	100	5	51	100	100	100	70	100	0	100	100	100	100	40	28
Bombyx	*mori*	0	41	0	0	0	0	0	13	47	14	1	34	37	100	70	0	0	0	9	0	0	0	14	33	0	0	81	0	5	19	0	15	3	8
Bos	*taurus*	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0
Branchiostoma	*floridae*	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Callorhinchus	*milii*	37	19	100	41	49	21	49	67	29	41	26	13	100	100	24	100	64	69	45	63	74	100	100	70	54	57	39	53	23	100	5	65	48	51
Capitella	*sp-2004*	0	0	0	22	0	24	100	0	0	3	0	3	3	1	0	7	6	0	3	11	0	2	7	12	0	0	1	0	0	9	9	5	11	37
Ciona	*intestinalis*	0	100	0	1	0	1	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	6	0	0	0	0	0	0	0
Ciona	*savignyi*	0	100	2	2	0	1	32	0	4	2	0	4	0	0	0	0	29	0	14	0	3	0	7	34	6	36	14	0	6	0	0	0	0	0
Crassostrea	*virginica*	0	100	0	100	0	0	100	73	74	67	71	100	100	67	100	53	25	100	51	100	53	100	6	100	74	100	60	10	100	100	100	100	47	100
Cyanea	*capillata*	100	100	100	100	100	77	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Danio	*rerio*	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	1	0	0	5	0	0	0	0	0	0	0	0	0	0
Daphnia	*pulex*	0	0	2	3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	1	1	0	0	1	0	0	1
Diplosoma	*listerianum*	17	100	2	100	100	13	100	100	100	100	100	100	100	100	100	100	100	100	61	100	100	100	56	100	100	100	100	100	100	100	100	100	100	100
Eptatretus	*burgeri*	41	70	0	100	0	0	100	100	20	28	100	100	54	100	100	0	35	20	100	6	19	100	2	43	59	100	100	0	30	100	100	100	0	100
Euprymna	*scolopes*	16	100	100	100	100	10	100	65	70	100	66	50	100	100	100	0	45	47	11	9	48	100	49	100	46	17	100	100	77	9	100	32	100	1
Gallus	*gallus*	0	100	0	0	0	0	0	0	0	0	100	0	0	0	0	9	100	0	0	0	0	100	0	0	0	0	0	0	0	0	56	0	0	0
Halocynthia	*roretzi*	100	100	0	100	100	1	35	100	32	18	0	51	61	100	47	74	48	1	100	100	100	23	7	100	12	31	58	0	12	100	0	25	0	100
Helobdella	*robusta*	27	100	0	62	31	23	42	18	6	47	34	35	14	75	100	35	100	30	9	17	9	18	4	4	8	60	43	1	22	44	10	51	57	36
Holothuria	*glaberrima*	100	100	0	100	60	2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Hydra	*magnipapillata*	0	0	6	0	0	1	35	17	0	20	13	47	0	66	10	0	5	0	0	0	0	0	0	12	0	33	1	0	0	3	7	0	0	2
Hydractinia	*echinata*	100	100	0	100	0	1	100	100	100	100	74	100	100	100	76	27	100	100	75	29	100	100	56	100	47	100	100	24	21	100	100	27	100	100
Ixodes	*scapularis*	31	3	0	0	0	0	100	21	16	29	9	0	21	59	10	26	0	0	20	0	0	0	1	21	11	0	6	0	0	11	38	0	55	9
Litopenaeus	*vannamei*	55	100	1	2	0	2	100	89	100	82	80	100	100	100	100	0	74	45	100	100	100	100	7	100	87	100	85	100	100	60	100	18	100	72
Lottia	*gigantea*	37	0	16	100	42	30	100	21	10	32	24	3	47	63	78	36	30	45	27	46	32	21	28	75	27	37	68	51	7	26	21	41	78	55
Lumbricus	*rubellus*	100	52	0	100	1	0	100	72	100	100	100	100	100	100	100	100	100	19	100	14	45	100	9	100	0	100	100	100	52	100	100	100	100	100
Molgula	*tectiformis*	4	100	0	3	0	0	6	20	21	30	0	51	13	100	100	0	8	0	0	0	1	0	0	24	0	42	37	0	0	100	38	0	0	1
Monodelphis	*domestica*	0	100	100	0	1	0	100	0	0	0	0	0	0	0	0	0	0	0	100	100	0	100	0	0	0	100	0	0	0	0	0	0	100	0
Mytilus	*galloprovinciali*	0	12	0	0	0	0	100	78	100	100	100	68	100	100	100	100	0	1	56	0	100	30	1	100	79	63	84	100	58	47	6	100	100	2
Nematostella	*vectensis*	0	0	0	0	0	2	7	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	1	8	0	0	0	0	0	0
Oikopleura	*dioica*	0	100	43	12	0	1	17	3	1	0	2	3	12	100	100	7	0	1	2	0	11	1	6	100	1	8	11	1	0	0	100	3	0	2
Oscarella	*carmela*	100	44	100	100	100	1	2	45	100	100	100	100	100	100	72	100	43	100	72	8	9	100	24	100	17	100	54	0	0	100	33	100	100	100
Paracentrotus	*lividus*	8	10	100	0	0	0	5	8	38	42	0	12	12	85	100	3	18	0	9	43	100	0	41	0	12	100	46	100	13	0	0	100	0	31
Pediculus	*humanus*	14	0	0	0	0	1	0	2	0	3	18	0	6	11	3	0	0	12	0	6	6	2	0	12	0	0	6	0	0	0	17	5	5	0
Petromyzon	*marinus*	0	21	0	0	0	0	7	29	15	61	18	100	26	63	19	100	2	0	0	8	0	1	0	50	0	46	34	53	45	76	36	100	0	58
Platynereis	*dumerilii*	0	100	100	100	0	2	100	57	64	30	36	59	100	100	100	100	100	0	100	100	32	100	31	100	29	100	100	0	100	100	100	100	100	100
Reniera	*sp.*	0	7	0	0	0	0	13	0	1	1	15	11	3	62	0	10	1	0	8	1	0	0	11	12	2	31	16	42	0	0	0	64	0	0
Rhipicephalus	*microplus*	0	81	0	5	0	0	0	0	100	17	40	44	32	100	100	100	15	0	0	0	0	0	0	100	0	100	89	0	0	0	100	100	100	1
Saccoglossus	*kowalevskii*	0	0	0	0	0	0	100	8	0	30	0	18	4	32	55	0	0	0	0	0	0	0	0	51	10	33	57	0	0	55	7	0	77	0
Solaster	*stimpsonii*	18	100	2	100	100	11	100	100	100	100	71	100	100	100	100	100	100	34	100	100	42	35	31	100	77	100	100	100	100	100	100	100	100	100
Spadella	*cephaloptera*	100	100	0	0	0	27	100	100	100	100	100	100	100	100	100	13	83	100	100	71	52	72	100	100	100	100	100	13	100	100	100	100	100	100
Squalus	*acanthias*	100	56	1	0	100	0	13	37	33	100	100	100	100	100	100	7	50	51	33	49	0	50	51	60	60	100	77	100	100	38	100	9	37	59
Strongylocentrotus	*purpuratus*	0	10	2	0	0	0	100	0	0	14	0	0	0	0	7	0	100	0	0	0	0	0	0	0	0	5	0	0	13	0	0	0	0	0
Suberites	*domuncula*	100	100	0	100	100	0	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Tetraodon	*nigroviridis*	100	0	0	100	0	0	0	100	0	3	5	100	12	35	100	0	0	0	0	0	2	100	0	100	5	0	1	3	0	0	0	100	4	0
Tribolium	*castaneum*	0	1	0	0	0	0	0	0	2	0	0	0	0	0	1	1	0	0	0	0	0	0	1	3	0	0	1	0	0	0	0	0	0	0
Xenopus	*tropicalis*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Xenoturbella	*bocki*	61	100	0	100	100	0	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	29	89	100	100	100	100	0	39	100	100	100	100	100
% missing positions		**32**	**54**	**15**	**38**	**21**	**5**	**57**	**42**	**40**	**42**	**37**	**46**	**44**	**62**	**58**	**35**	**40**	**29**	**36**	**32**	**32**	**38**	**23**	**56**	**30**	**52**	**49**	**29**	**32**	**44**	**44**	**47**	**44**	**40**
% missing OTUs		**24**	**43**	**14**	**35**	**18**	**0**	**51**	**25**	**27**	**25**	**24**	**33**	**35**	**49**	**47**	**27**	**25**	**20**	**24**	**24**	**22**	**31**	**12**	**45**	**14**	**41**	**27**	**24**	**22**	**35**	**37**	**39**	**33**	**29**
% chimeras	**chimeras**	**2**	**0**	**0**	**0**	**0**	**2**	**0**	**2**	**0**	**2**	**4**	**2**	**4**	**0**	**0**	**0**	**2**	**4**	**0**	**2**	**0**	**0**	**2**	**0**	**8**	**0**	**0**	**2**	**2**	**0**	**0**	**0**	**0**	**0**
# amino-acid sites	**length**	**239**	**442**	**123**	**127**	**124**	**246**	**186**	**617**	**708**	**525**	**641**	**627**	**686**	**398**	**461**	**187**	**400**	**385**	**422**	**385**	**403**	**380**	**441**	**265**	**762**	**451**	**497**	**177**	**371**	**209**	**239**	**229**	**230**	**237**

		pace	psm	psm	psm	psm	psm	psm	psm	psm	psm	psm	psm	psm	psm	psm	pyrd	pyrd	rad2	rad5	rad5	rf1	rla2-	rla2-	rpl1	rpl11	rpl12	rpl13	rpl14	rpl15	rpl16	rpl17	rpl18	rpl19	rpl2
Acropora	***millepora***	3	0	10	1	1	0	100	100	100	0	8	0	100	35	100	100	71	25	29	72	100	10	0	100	0	1	0	0	52	0	1	1	0	0
Ambystoma	***mexicanum***	0	0	3	0	0	0	0	0	41	0	100	0	0	0	0	2	0	9	40	100	55	3	0	0	0	0	0	0	0	0	0	0	0	0
Apis	***mellifera***	0	0	0	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0	0	100	40	3	0	0	0	0	1	0	0	0	0	0	0	0
Aplysia	***californica***	100	17	0	5	0	73	0	0	0	0	0	10	0	0	0	0	7	0	1	88	19	1	4	0	0	0	0	0	0	0	0	1	2	0
Asterina	***pectinifera***	0	0	1	0	0	36	100	26	0	14	22	0	0	21	6	100	55	62	100	100	36	2	0	0	0	0	0	0	0	0	0	0	0	16
Bombyx	***mori***	0	0	0	20	0	4	0	0	0	0	0	0	0	24	1	0	0	4	0	0	0	2	0	0	0	0	0	0	0	0	0	3	0	0
Bos	***taurus***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Branchiostoma	***floridae***	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	39	0	0	3	0	0	0	0	0	0	0	1	0	0	0
Callorhinchus	***milii***	100	81	35	100	46	46	69	58	100	100	37	59	100	100	100	31	18	10	24	100	63	100	100	10	9	56	100	100	28	57	37	100	38	62
Capitella	***sp-2004***	24	14	23	21	4	10	4	31	40	23	21	5	1	10	1	13	11	30	2	12	7	11	24	0	24	7	0	0	0	25	1	15	0	5
Ciona	***intestinalis***	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	26	0	33	0	29	0	1	0	0	0	0	1	0	0	0	1	1	0	0
Ciona	***savignyi***	45	0	0	0	0	0	0	0	1	0	0	0	0	0	0	35	0	0	0	0	0	5	0	0	0	0	1	0	0	0	1	1	0	0
Crassostrea	***virginica***	100	25	3	72	15	100	100	0	29	100	100	0	71	64	55	100	100	48	100	100	100	0	0	0	0	0	0	0	0	0	0	1	0	0
Cyanea	***capillata***	100	100	29	100	100	100	48	100	100	100	2	100	100	100	100	100	100	100	100	100	100	4	3	19	0	0	0	0	64	0	1	1	100	100
Danio	***rerio***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Daphnia	***pulex***	13	0	0	0	0	0	0	3	0	0	1	0	2	0	0	0	0	0	2	100	0	2	0	0	0	0	1	0	0	0	0	0	0	0
Diplosoma	***listerianum***	100	100	100	100	100	100	100	100	1	100	100	100	18	26	100	100	100	48	100	100	100	17	0	0	100	0	2	0	0	0	1	0	33	0
Eptatretus	***burgeri***	1	0	0	87	0	0	53	0	0	22	12	12	0	0	0	100	31	52	100	100	100	0	0	0	0	0	0	0	0	0	0	0	0	0
Euprymna	***scolopes***	1	8	26	3	23	8	0	14	100	0	21	14	0	0	9	53	6	52	29	100	40	34	0	0	70	21	0	100	0	0	0	0	1	0
Gallus	***gallus***	0	0	0	0	0	4	0	0	0	0	0	0	100	0	0	0	0	9	0	0	100	0	0	0	0	0	0	0	0	100	100	11	0	0
Halocynthia	***roretzi***	1	61	1	0	28	100	100	100	100	100	0	100	0	100	69	100	56	4	100	100	0	100	0	100	100	0	15	100	100	100	100	34	49	10
Helobdella	***robusta***	13	41	10	6	4	17	69	55	37	9	21	42	13	30	39	15	21	78	2	0	12	3	100	0	1	4	1	28	0	0	15	0	39	17
Holothuria	***glaberrima***	100	100	100	100	100	100	100	100	100	25	100	100	100	100	100	100	100	100	100	100	100	3	0	0	100	0	0	0	0	0	12	0	0	1
Hydra	***magnipapillata***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	4	3	0	0	5	0	0	0	2	0	1	0	0	0	1	1	0	0
Hydractinia	***echinata***	27	0	10	0	85	43	23	10	0	27	8	100	1	0	0	100	56	44	100	100	54	3	0	0	2	0	1	0	0	0	1	1	0	5
Ixodes	***scapularis***	0	0	1	25	4	0	0	16	0	6	0	0	0	0	0	0	10	0	1	0	3	3	22	0	0	0	0	0	0	0	0	0	0	0
Litopenaeus	***vannamei***	1	9	71	100	0	1	36	0	0	0	30	28	47	58	0	100	100	31	100	100	100	5	0	0	0	0	1	0	0	0	1	0	2	0
Lottia	***gigantea***	100	0	10	48	37	80	83	53	49	25	34	28	33	66	0	22	29	80	13	76	25	39	100	0	24	62	34	100	0	72	20	73	38	18
Lumbricus	***rubellus***	100	27	1	100	100	100	41	0	0	0	0	1	0	40	45	100	43	0	36	83	60	3	0	0	0	1	1	0	0	0	0	0	0	0
Molgula	***tectiformis***	2	100	0	2	0	0	0	0	0	0	0	0	0	0	0	45	0	3	0	100	0	4	0	0	0	0	0	100	1	0	1	1	0	0
Monodelphis	***domestica***	0	0	0	0	0	0	0	0	9	1	0	0	0	0	100	100	0	0	0	0	100	100	0	0	0	0	0	0	0	0	0	0	100	0
Mytilus	***galloprovincial.***	2	0	100	100	27	12	0	100	0	0	0	49	4	18	0	100	100	50	0	100	13	4	0	0	0	0	0	0	0	0	0	10	0	0
Nematostella	***vectensis***	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	45	0	3	0	0	2	1	0	0	0	0	1	1	0	0
Oikopleura	***dioica***	1	0	0	12	1	0	1	0	4	27	20	7	1	0	6	2	2	5	1	1	13	7	1	0	0	0	1	4	11	6	1	4	0	3
Oscarella	***carmela***	100	57	100	100	53	100	100	100	100	100	100	0	0	100	0	100	0	48	100	100	54	100	9	0	4	0	100	100	100	0	0	1	100	0
Paracentrotus	***lividus***	100	0	6	4	0	19	0	24	0	0	100	0	100	0	6	52	36	100	0	100	51	5	100	100	100	100	100	100	0	0	0	0	0	6
Pediculus	***humanus***	100	0	0	8	3	0	3	10	0	0	4	0	1	0	0	0	0	9	0	5	20	3	0	0	5	7	0	0	0	0	0	0	0	0
Petromyzon	***marinus***	0	0	19	0	0	0	0	0	100	23	39	0	0	0	0	57	34	71	0	55	5	0	0	0	0	0	0	0	0	0	0	0	0	0
Platynereis	***dumerilii***	100	5	100	100	100	0	0	100	100	100	100	0	36	100	100	30	100	100	100	100	57	100	4	0	57	100	0	0	0	100	100	100	0	0
Reniera	***sp.***	33	0	0	2	0	0	0	0	1	1	0	0	0	0	37	55	0	50	0	4	0	0	0	0	0	0	1	0	0	0	0	1	0	0
Rhipicephalus	***microplus***	0	0	1	0	0	0	100	0	0	0	0	0	0	0	0	0	22	0	0	40	49	3	3	0	0	0	0	0	0	0	0	1	8	0
Saccoglossus	***kowalevskii***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	0	17	2	64	0	8	0	0	0	0	0	0	0	0	0	0	0	0
Solaster	***stimpsonii***	100	100	100	23	100	100	100	100	100	10	100	100	0	100	100	100	100	100	100	100	100	100	100	0	0	100	0	0	100	0	100	100	51	0
Spadella	***cephaloptera***	100	100	100	7	100	0	54	18	0	0	0	100	0	100	60	100	100	100	48	100	40	46	0	0	2	0	1	2	0	3	0	1	0	15
Squalus	***acanthias***	100	14	27	48	6	21	0	100	0	100	100	0	1	3	11	0	0	69	100	100	66	2	0	31	100	72	100	0	100	0	0	1	100	6
Strongylocentrotus	***purpuratus***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	100	0	5	5	4	0	0	0	0	0	0	0	0	0	0	0	0
Suberites	***domuncula***	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	2	0	0	1	0	0	2	7	0	0	0	0	4
Tetraodon	***nigroviridis***	1	0	0	0	100	10	19	0	46	0	0	100	35	0	0	0	0	7	100	0	0	0	0	100	0	0	0	25	0	100	0	0	2	0
Tribolium	***castaneum***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	3	0	0	0	0	0	0	0	0	3	0	0	0
Xenopus	***tropicalis***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	76	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Xenoturbella	***bocki***	100	100	100	0	100	16	100	5	100	100	0	100	100	25	100	100	100	100	100	100	100	3	4	0	0	0	0	0	0	0	1	1	74	0
% missing positions		**37**	**23**	**23**	**27**	**26**	**26**	**32**	**28**	**29**	**24**	**25**	**25**	**21**	**26**	**26**	**45**	**32**	**36**	**36**	**63**	**39**	**17**	**11**	**9**	**14**	**10**	**9**	**15**	**11**	**11**	**10**	**9**	**14**	**5**
% missing OTUs		**33**	**16**	**18**	**20**	**20**	**18**	**22**	**22**	**24**	**20**	**20**	**20**	**16**	**18**	**20**	**35**	**22**	**18**	**31**	**49**	**24**	**12**	**10**	**8**	**10**	**6**	**8**	**14**	**8**	**8**	**8**	**6**	**8**	**2**
% chimeras	**chimeras**	**2**	**2**	**0**	**0**	**0**	**4**	**0**	**2**	**0**	**2**	**2**	**2**	**2**	**2**	**2**	**0**	**2**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**2**	**0**	**0**	**0**	**0**	**0**	**2**
# amino-acid sites	**length**	**106**	**225**	**242**	**252**	**228**	**216**	**241**	**241**	**185**	**205**	**212**	**234**	**196**	**195**	**194**	**321**	**322**	**216**	**315**	**320**	**402**	**92**	**76**	**213**	**169**	**163**	**179**	**123**	**204**	**173**	**175**	**183**	**190**	**248**

		rpl20	rpl21	rpl22	rpl23	rpl24-	rpl24-	rpl25	rpl26	rpl27	rpl3	rpl30	rpl31	rpl32	rpl33	rpl34	rpl35	rpl36	rpl37	rpl38	rpl39	rpl42	rpl43	rpl4	rpl5	rpl6	rpl7-	rpl9	rpo-	rpo-	rpo-	rpp0	rps1	rps1	rps1
Acropora	***millepora***	6	1	0	1	0	2	0	1	10	0	100	0	0	0	0	1	1	0	49	0	100	0	21	0	5	0	1	87	100	100	0	0	2	100
Ambystoma	***mexicanum***	0	0	0	0	3	0	0	0	0	0	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	100	100	0	0	0	0
Apis	***mellifera***	0	1	0	3	0	20	0	0	0	0	0	0	0	0	1	0	0	0	100	100	0	1	0	0	2	0	0	0	4	43	0	0	0	0
Aplysia	***californica***	0	0	0	1	16	0	0	0	0	0	0	0	0	0	2	7	0	0	32	4	0	1	0	0	1	0	1	23	4	13	0	0	3	0
Asterina	***pectinifera***	0	1	0	0	0	0	0	0	0	100	0	0	0	0	0	1	1	100	0	0	0	0	100	1	1	0	0	100	100	100	0	0	0	0
Bombyx	***mori***	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2	0	0	0	31	58	0	0	0	0
Bos	***taurus***	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Branchiostoma	***floridae***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Callorhinchus	***milii***	78	71	21	48	100	58	22	0	31	29	100	61	53	28	100	100	26	100	100	100	100	57	45	23	18	37	100	15	86	12	6	46	18	71
Capitella	***sp-2004***	4	0	55	1	33	20	21	0	1	0	3	25	2	6	41	38	22	42	12	100	58	1	7	1	11	0	7	0	2	1	17	17	18	1
Ciona	***intestinalis***	0	0	0	1	0	0	0	0	0	0	0	0	2	0	0	0	0	9	0	0	0	1	0	0	0	0	0	0	0	50	0	0	0	0
Ciona	***savignyi***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	1	0	0	0	0	0	0	25	0	0	0	0	0
Crassostrea	***virginica***	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	42	42	4	1	7	0	9	1	0	1	100	9	100	21	0	0	0
Cyanea	***capillata***	2	1	100	1	14	0	1	0	0	75	0	0	0	0	0	1	0	0	2	0	0	0	100	75	21	0	4	100	100	100	44	100	100	0
Danio	***rerio***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	11	0	100	0	0	0	0
Daphnia	***pulex***	0	0	5	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	2	0
Diplosoma	***listerianum***	0	100	100	1	11	100	1	100	0	100	1	0	0	0	0	0	0	12	0	0	0	100	100	100	100	10	100	100	100	100	0	4	1	4
Eptatretus	***burgeri***	0	0	100	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	8	0	0	0	0	0	100	59	50	0	0	0	0
Euprymna	***scolopes***	0	3	1	3	0	0	0	1	1	0	57	0	0	0	10	0	100	48	100	100	100	100	11	0	2	0	0	100	100	100	9	0	5	1
Gallus	***gallus***	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	5	100	0	0	8	0	0
Halocynthia	***roretzi***	100	0	9	100	100	100	100	100	3	1	0	0	5	100	100	100	0	100	100	100	2	100	26	0	100	100	3	100	75	65	0	100	0	100
Helobdella	***robusta***	36	0	21	1	3	52	6	0	1	1	0	6	0	0	100	2	20	1	43	0	0	2	0	4	26	5	0	53	2	66	25	0	1	23
Holothuria	***glaberrima***	0	1	0	0	0	100	0	0	0	0	0	0	0	2	0	3	40	2	0	0	0	0	1	0	1	0	0	100	100	100	0	0	9	0
Hydra	***magnipapillata***	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	1	20	11	27	0	0	2	0
Hydractinia	***echinata***	0	1	0	1	0	26	0	0	0	2	3	0	0	0	3	1	0	0	0	0	2	0	1	0	3	0	1	100	100	100	0	1	2	0
Ixodes	***scapularis***	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28	11	2	0	0	0	0
Litopenaeus	***vannamei***	0	0	0	0	2	100	6	0	0	0	0	0	1	0	3	0	0	0	0	0	0	0	0	0	3	0	0	100	84	100	16	0	1	0
Lottia	***gigantea***	74	100	55	41	32	54	22	58	31	12	12	63	21	6	100	38	46	46	43	100	67	51	16	23	45	31	24	43	4	100	71	32	18	22
Lumbricus	***rubellus***	0	0	0	0	0	13	0	0	0	0	0	0	0	5	0	0	0	41	0	0	0	0	0	0	1	1	2	100	100	100	0	0	0	1
Molgula	***tectiformis***	0	0	0	0	1	0	0	0	0	0	0	20	0	0	0	0	0	9	0	0	0	1	0	0	0	0	2	65	73	100	0	0	0	0
Monodelphis	***domestica***	100	0	0	100	1	0	0	0	0	0	3	0	0	0	0	1	0	0	0	100	100	1	0	0	0	0	0	100	0	0	0	0	0	0
Mytilus	***galloprovincial.***	0	0	0	0	0	78	0	0	0	0	0	0	0	4	10	1	41	57	3	10	0	0	0	1	1	8	1	100	100	100	0	2	0	0
Nematostella	***vectensis***	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	3	0	0	1	0	0	0	0	2	0
Oikopleura	***dioica***	60	0	0	0	0	0	2	0	1	0	0	1	7	37	0	0	60	2	2	100	29	0	2	1	9	2	0	0	0	2	1	0	42	0
Oscarella	***carmela***	1	0	0	100	100	100	0	100	100	1	100	100	100	3	0	1	100	100	0	100	100	100	33	0	2	0	100	100	100	100	0	0	39	0
Paracentrotus	***lividus***	75	100	0	0	100	0	100	100	1	34	100	100	100	100	3	100	100	100	100	0	100	0	0	100	100	100	0	80	10	65	11	100	0	100
Pediculus	***humanus***	2	1	0	0	0	0	0	0	0	0	3	5	0	1	20	1	20	0	100	100	0	0	0	0	8	16	0	0	2	0	5	2	0	0
Petromyzon	***marinus***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	59	63	0	0	0	0
Platynereis	***dumerilii***	0	100	0	4	100	100	100	0	100	0	100	100	0	100	100	2	100	100	100	100	100	100	0	49	100	100	100	100	58	100	0	100	100	0
Reniera	***sp.***	0	3	12	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	2	0	0	0	0	0	2	0	1	17	2	19	1	0	0	0
Rhipicephalus	***microplus***	0	0	0	0	0	0	5	1	0	0	20	0	0	0	0	100	100	1	0	100	100	0	0	0	0	0	0	57	100	75	0	1	0	1
Saccoglossus	***kowalevskii***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	0	0	16	5	41	0	0	0	0
Solaster	***stimpsonii***	100	100	0	100	100	100	0	100	0	4	100	100	100	100	100	100	100	100	100	100	100	19	0	1	100	0	32	100	100	100	8	5	0	100
Spadella	***cephaloptera***	0	1	0	0	0	100	54	0	1	6	0	0	0	0	0	1	0	0	0	0	0	0	100	22	1	1	1	100	100	100	19	14	3	0
Squalus	***acanthias***	100	0	100	19	0	0	22	2	0	0	0	100	0	100	100	100	100	0	0	100	100	0	0	0	26	100	100	100	100	82	14	100	0	24
Strongylocentrotus	***purpuratus***	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	24	0	0	0	0
Suberites	***domuncula***	0	0	0	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	100	100	0	0	0	0	4	0	1	100	38	100	0	0	0	0
Tetraodon	***nigroviridis***	0	0	100	100	0	7	0	17	0	0	100	6	0	0	0	7	6	0	100	100	100	27	16	0	64	0	0	8	12	6	0	0	0	0
Tribolium	***castaneum***	0	1	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	2	0	0	0	0	0	0	0	0	0
Xenopus	***tropicalis***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0
Xenoturbella	***bocki***	0	3	100	3	0	0	21	0	0	0	0	0	0	0	100	3	0	0	2	0	0	0	1	0	0	0	0	100	100	100	0	1	0	0
% missing positions		**15**	**12**	**17**	**12**	**14**	**22**	**10**	**11**	**6**	**7**	**16**	**13**	**8**	**12**	**19**	**14**	**19**	**20**	**24**	**34**	**25**	**13**	**11**	**8**	**15**	**10**	**11**	**54**	**47**	**56**	**5**	**12**	**7**	**11**
% missing OTUs		**8**	**10**	**14**	**10**	**12**	**16**	**6**	**10**	**4**	**4**	**14**	**10**	**6**	**10**	**18**	**12**	**14**	**14**	**20**	**33**	**22**	**10**	**8**	**4**	**10**	**8**	**10**	**43**	**33**	**41**	**0**	**10**	**4**	**8**
% chimeras	**chimeras**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**2**	**0**	**2**	**2**	**0**	**0**	**0**
# amino-acid sites	**length**	**160**	**154**	**96**	**139**	**123**	**137**	**126**	**135**	**137**	**393**	**105**	**108**	**129**	**106**	**108**	**123**	**85**	**81**	**65**	**51**	**102**	**88**	**326**	**275**	**186**	**207**	**183**	**684**	**1419**	**1091**	**289**	**244**	**119**	**140**

		rps1	rps1	rps1	rps1	rps1	rps1	rps1	rps2	rps2	rps2	rps2	rps2	rps2	rps2	rps2	rps2	rps2	rrp4	rrp4	sadh	sadh	sap4	sra	srp5	srs	stbp	stcpr	suca	tfiid	tif2a	topo	u2sn	vaca	vata
Acropora	***millepora***	0	0	0	0	0	0	0	1	0	0	0	0	5	2	0	0	4	0	72	100	20	0	66	100	87	0	100	24	1	2	33	0	100	71
Ambystoma	***mexicanum***	0	100	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	13	100	0	0	83	0	17	0	0	16	100	6	56	0	100	74
Apis	***mellifera***	0	3	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0
Aplysia	***californica***	0	1	0	0	2	3	0	0	0	0	3	4	1	0	0	0	0	1	14	18	1	0	17	42	2	21	4	0	27	0	17	21	24	19
Asterina	***pectinifera***	0	0	0	0	0	0	0	0	1	0	0	0	100	3	0	0	100	8	7	57	0	100	100	100	0	23	100	100	24	100	0	0	100
Bombyx	***mori***	0	0	0	0	0	0	0	1	0	0	0	0	0	2	0	0	62	20	8	1	0	0	0	0	0	13	0	0	0	1	18	63	0	0
Bos	***taurus***	0	8	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Branchiostoma	***floridae***	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	2	2	0	0	0
Callorhinchus	***milii***	40	100	100	43	27	100	100	100	65	100	43	33	100	100	39	34	100	100	74	78	17	61	38	63	27	45	100	46	100	60	100	56	100	25
Capitella	***sp-2004***	13	15	1	5	12	16	0	0	0	0	43	1	3	8	0	2	39	0	0	16	2	0	12	1	6	100	6	0	1	24	0	8	10	4
Ciona	***intestinalis***	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	8	0	2	0	100	1	0	0	0	0	0	0	0	0	0	0	1	0	0
Ciona	***savignyi***	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	20	0	0	0	0	0	0	16	1	0	0	0
Crassostrea	***virginica***	1	1	0	0	0	0	3	0	0	1	2	4	0	0	0	0	14	100	100	100	1	4	83	78	100	51	67	53	75	100	69	1	37	100
Cyanea	***capillata***	0	100	0	100	0	0	8	0	0	0	0	0	100	0	0	2	4	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Danio	***rerio***	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Daphnia	***pulex***	0	0	0	0	0	0	1	0	0	0	0	0	1	2	1	0	0	0	0	0	0	0	0	0	0	0	3	0	0	1	6	0	0	1
Diplosoma	***listerianum***	0	100	1	1	100	0	7	28	0	0	0	1	0	2	2	0	0	100	100	59	50	0	100	100	100	60	100	31	100	100	100	100	0	100
Eptatretus	***burgeri***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	23	100	0	100	100	100	6	36	0	100	11	61	100	100	62
Euprymna	***scolopes***	0	0	0	0	8	0	0	0	12	1	0	100	100	100	0	0	100	100	100	19	63	10	100	19	54	0	49	37	100	46	10	23	0	100
Gallus	***gallus***	16	0	100	0	0	100	100	0	0	0	100	0	0	0	0	0	0	100	100	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0
Halocynthia	***roretzi***	100	100	14	100	100	100	100	100	100	100	100	100	100	100	2	100	100	100	100	0	9	2	100	0	12	52	12	100	1	54	0	0	0	46
Helobdella	***robusta***	0	1	0	66	0	15	13	2	1	0	26	4	1	0	0	0	9	1	14	38	24	0	46	49	31	17	54	13	26	73	82	33	25	45
Holothuria	***glaberrima***	5	0	0	0	7	1	1	0	5	0	0	0	0	1	0	0	5	100	100	100	56	0	100	100	100	100	100	100	100	100	100	100	100	100
Hydra	***magnipapillata***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	4	0	0	4	1	0	8	1	1	0	2	0	0	0	5	0	0	9
Hydractinia	***echinata***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	4	100	100	100	43	0	63	100	100	100	100	100	0	41	100	0	10	100
Ixodes	***scapularis***	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	2	0	18	0	15	0	0	25	0	1	0	86	0	25	1	4	0	0	8
Litopenaeus	***vannamei***	0	0	0	0	0	0	0	1	0	1	0	0	2	0	1	0	4	100	100	100	0	0	63	100	27	85	100	78	100	3	100	100	0	30
Lottia	***gigantea***	40	32	49	66	13	57	0	37	100	28	43	33	2	10	0	100	100	0	0	55	9	1	46	32	0	17	47	16	23	36	72	44	0	21
Lumbricus	***rubellus***	0	1	0	0	0	0	0	0	0	0	2	2	0	0	0	0	0	5	100	100	2	0	100	90	63	49	100	100	62	67	100	100	44	38
Molgula	***tectiformis***	0	0	0	0	0	0	1	1	1	0	0	1	0	0	2	0	0	100	0	10	1	0	49	1	0	0	0	0	0	100	0	0	100	56
Monodelphis	***domestica***	0	100	0	0	0	100	100	0	0	1	0	100	1	2	0	2	4	0	0	0	0	100	0	0	0	100	100	0	100	0	0	0	100	0
Mytilus	***galloprovincial***	5	1	0	0	0	0	0	0	0	1	4	4	0	1	0	0	100	29	100	100	100	0	75	55	100	29	2	0	63	6	100	1	0	100
Nematostella	***vectensis***	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	2	0	0	0	0
Oikopleura	***dioica***	0	5	1	0	10	0	2	1	4	0	62	11	34	2	26	0	100	100	45	6	20	0	2	3	1	8	10	6	1	53	2	9	13	2
Oscarella	***carmela***	0	1	100	0	100	100	100	100	100	100	2	100	100	2	0	100	100	100	100	100	18	0	100	100	24	100	100	0	100	100	100	1	100	56
Paracentrotus	***lividus***	100	0	100	100	1	100	0	100	0	0	100	0	100	0	100	100	0	0	0	15	62	5	41	5	56	0	9	0	0	27	0	0	0	27
Pediculus	***humanus***	5	0	0	41	12	1	0	0	0	1	0	3	1	56	1	0	39	33	4	0	2	0	2	5	1	67	0	0	0	4	0	8	0	4
Petromyzon	***marinus***	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	40	2	0	0	46	0	8	29	0	0	1	22	0	58	31
Platynereis	***dumerilii***	1	1	100	100	100	100	100	0	0	0	2	100	100	100	100	100	100	100	100	100	36	100	100	61	49	100	100	16	100	12	100	0	34	100
Reniera	***sp.***	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	3	2	22	12	7	7	0	21	0	11	0	1	14	12	2	22	1	0	0
Rhipicephalus	***microplus***	0	0	0	0	2	0	0	0	0	1	0	0	3	100	1	100	0	0	0	56	0	0	28	17	1	100	0	0	21	1	26	26	0	10
Saccoglossus	***kowalevskii***	0	0	0	0	0	0	0	1	0	0	0	3	0	0	0	0	0	0	0	33	0	0	56	20	4	100	2	0	19	29	9	8	0	9
Solaster	***stimpsonii***	100	100	100	100	100	0	100	0	100	100	100	100	100	0	100	100	100	100	100	100	100	0	100	100	100	100	100	100	100	100	100	100	100	62
Spadella	***cephaloptera***	0	0	1	0	0	0	0	0	0	0	0	0	0	2	0	8	0	100	100	100	100	0	100	100	100	100	100	100	100	100	100	7	100	100
Squalus	***acanthias***	100	100	0	0	100	21	0	0	100	100	0	0	0	100	0	0	0	100	100	51	56	0	44	53	43	0	44	0	82	36	78	2	0	100
Strongylocentrotus	***purpuratus***	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	8	0	0
Suberites	***domuncula***	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	100	100	100	100	0	0	100	100	100	100	100	100	100	100	100	100	100	100
Tetraodon	***nigroviridis***	0	100	100	0	0	26	13	100	0	100	16	1	1	0	10	0	100	100	0	7	0	100	5	0	0	0	9	0	18	0	8	0	0	0
Tribolium	***castaneum***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Xenopus	***tropicalis***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	69	0	100	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0
Xenoturbella	***bocki***	0	0	0	1	0	0	1	0	0	2	0	0	3	0	0	0	100	100	61	100	2	0	100	100	63	100	100	1	100	100	100	0	100	
% missing positions		**10**	**19**	**15**	**14**	**14**	**16**	**15**	**11**	**12**	**12**	**13**	**14**	**17**	**18**	**9**	**15**	**26**	**47**	**42**	**43**	**23**	**9**	**45**	**38**	**33**	**38**	**39**	**25**	**40**	**32**	**42**	**24**	**29**	**39**
% missing OTUs		**8**	**18**	**14**	**10**	**12**	**14**	**14**	**10**	**10**	**12**	**8**	**12**	**16**	**16**	**6**	**14**	**24**	**43**	**35**	**31**	**12**	**8**	**27**	**25**	**22**	**27**	**29**	**18**	**31**	**20**	**29**	**18**	**24**	**25**
% chimeras	**chimeras**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**2**	**0**	**0**	**4**	**6**	**2**	**2**	**2**	**2**	**0**	**0**	**4**	**0**	**2**	**2**
# amino-acid sites	**length**	**151**	**149**	**139**	**138**	**119**	**152**	**132**	**106**	**130**	**143**	**122**	**92**	**101**	**84**	**155**	**61**	**56**	**165**	**210**	**455**	**424**	**212**	**422**	**492**	**454**	**302**	**294**	**291**	**175**	**304**	**520**	**179**	**147**	**610**

		vatb	vatc	vate	vatp ased	vdac 2	w09c	wrs	xpb	yif1p	% missing positions	% missing genes	% chimeras
Acropora	***millepora***	82	28	5	10	0	0	35	100	30	**52**	**23**	**2**
Ambystoma	***mexicanum***	0	45	10	3	0	0	62	77	21	**33**	**13**	**7**
Apis	***mellifera***	0	0	0	0	0	0	0	0	2	**4**	**2**	**0**
Aplysia	***californica***	0	9	8	2	0	0	17	9	29	**14**	**2**	**0**
Asterina	***pectinifera***	100	100	100	100	0	100	100	100	100	**64**	**37**	**1**
Bombyx	***mori***	0	0	0	0	7	0	13	25	71	**10**	**1**	**0**
Bos	***taurus***	0	0	0	0	0	0	0	0	0	**1**	**1**	**0**
Branchiostoma	***floridae***	0	0	0	1	0	0	0	0	0	**0**	**0**	**0**
Callorhinchus	***milii***	33	52	100	63	47	38	20	64	51	**54**	**29**	**0**
Capitella	***sp-2004***	4	18	5	24	0	19	1	0	0	**7**	**2**	**0**
Ciona	***intestinalis***	0	9	4	0	0	0	2	0	36	**7**	**3**	**0**
Ciona	***savignyi***	17	0	0	0	0	0	0	0	0	**4**	**1**	**0**
Crassostrea	***virginica***	66	51	100	13	1	0	35	100	100	**52**	**26**	**5**
Cyanea	***capillata***	100	100	28	100	82	100	100	100	100	**86**	**66**	**0**
Danio	***rerio***	0	0	0	0	0	0	0	0	0	**4**	**2**	**0**
Daphnia	***pulex***	0	0	0	0	0	0	0	0	1	**2**	**1**	**0**
Diplosoma	***listerianum***	100	100	30	100	0	100	100	100	100	**82**	**62**	**0**
Eptatretus	***burgeri***	100	68	2	100	1	0	100	29	1	**42**	**21**	**0**
Euprymna	***scolopes***	19	100	1	100	0	0	34	73	100	**45**	**22**	**8**
Gallus	***gallus***	0	0	0	0	0	0	0	0	100	**12**	**10**	**0**
Halocynthia	***roretzi***	44	1	13	4	0	45	51	100	100	**45**	**40**	**0**
Helobdella	***robusta***	27	60	100	27	55	55	0	50	100	**27**	**4**	**4**
Holothuria	***glaberrima***	100	100	100	100	100	87	100	100	100	**81**	**60**	**0**
Hydra	***magnipapillata***	0	0	0	0	0	0	2	37	60	**5**	**0**	**2**
Hydractinia	***echinata***	100	100	100	100	1	100	100	62	38	**56**	**32**	**6**
Ixodes	***scapularis***	8	0	6	0	0	0	0	21	0	**8**	**1**	**0**
Litopenaeus	***vannamei***	66	29	8	61	42	0	100	100	100	**53**	**26**	**11**
Lottia	***gigantea***	39	81	0	29	38	39	5	0	0	**34**	**7**	**1**
Lumbricus	***rubellus***	38	64	0	0	1	18	100	100	100	**56**	**30**	**4**
Molgula	***tectiformis***	8	1	0	0	0	0	0	35	0	**17**	**6**	**1**
Monodelphis	***domestica***	0	0	0	0	0	0	0	0	0	**15**	**17**	**0**
Mytilus	***galloprovincial***	25	11	0	60	5	100	14	58	10	**51**	**25**	**2**
Nematostella	***vectensis***	0	1	0	1	0	0	0	0	0	**1**	**0**	**0**
Oikopleura	***dioica***	3	12	14	1	100	0	7	4	21	**11**	**6**	**0**
Oscarella	***carmela***	20	0	100	1	0	100	0	100	100	**59**	**51**	**1**
Paracentrotus	***lividus***	0	32	0	0	0	5	49	52	8	**30**	**25**	**0**
Pediculus	***humanus***	4	0	0	7	0	0	0	0	11	**4**	**2**	**0**
Petromyzon	***marinus***	17	29	0	4	0	0	12	33	31	**23**	**6**	**0**
Platynereis	***dumerilii***	100	34	100	100	0	100	100	100	100	**67**	**59**	**0**
Reniera	***sp.***	0	28	0	0	0	25	0	29	2	**7**	**1**	**0**
Rhipicephalus	***microplus***	0	0	20	0	0	0	69	100	0	**25**	**12**	**5**
Saccoglossus	***kowalevskii***	0	0	0	45	0	0	0	30	0	**11**	**2**	**0**
Solaster	***stimpsonii***	100	100	0	100	100	100	100	100	12	**81**	**71**	**0**
Spadella	***cephaloptera***	100	100	38	100	19	100	100	100	100	**75**	**49**	**1**
Squalus	***acanthias***	19	46	16	77	18	65	81	80	28	**54**	**28**	**0**
Strongylocentrotus	***purpuratus***	0	0	0	0	100	0	0	4	0	**5**	**4**	**0**
Suberites	***domuncula***	100	100	100	0	100	100	100	100	100	**77**	**61**	**0**
Tetraodon	***nigroviridis***	0	10	0	0	1	23	0	2	0	**17**	**16**	**0**
Tribolium	***castaneum***	0	0	0	0	0	0	0	0	0	**2**	**2**	**0**
Xenopus	***tropicalis***	0	0	0	0	0	0	0	0	0	**1**	**1**	**0**
Xenoturbella	***bocki***	100	80	100	0	0	100	100	100	100	**65**	**45**	**0**
% missing positions		**32**	**33**	**24**	**28**	**16**	**30**	**35**	**47**	**40**			
% missing OTUs		**22**	**18**	**20**	**20**	**10**	**22**	**25**	**31**	**31**			
% chimeras	**chimeras**	**4**	**0**	**0**	**0**	**2**	**4**	**0**	**0**	**0**			
# amino-acid sites	**length**	**479**	**336**	**200**	**210**	**276**	**260**	**379**	**631**	**188**			

Table S4: Percentages of missing data in the 106-genes supermatrix

		ar21	arc20	arp23	atp syn thal pha-mt	cct-A	cct-B	cct-D	cct-E	cct-G	cct-N	cct-T	cct-Z	cpn 60-mt	ef2-EF2	fibri	grc5	hsp 70-E	hsp 70-mt	hsp 90-C	if2g	if4a-a	l12e-A	l12e-C	l12e-D	nsf 1-G	nsf 1-M	nsf 2-A	ps ma-A	ps ma-B	ps ma-C	ps ma-D	ps ma-E	ps mb-I	ps mb-J
Acropora	***millepora***	0	0	4	6	100	73	25	52	53	64	70	13	46	45	55	0	47	31	0	33	100	0	0	5	62	7	55	0	10	1	1	0	0	8
Ambystoma	***mexicanum***	0	0	0	0	0	0	1	0	43	0	26	6	4	7	0	0	43	0	100	18	46	1	0	0	9	35	19	0	3	0	0	0	0	100
Apis	***mellifera***	1	0	0	0	0	0	0	0	0	0	0	42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Aplysia	***californica***	1	0	0	0	8	0	0	5	33	1	3	22	5	0	5	2	0	4	0	0	4	4	0	0	5	6	17	17	0	5	0	73	0	0
Asterina	***pectinifera***	100	0	12	100	55	100	100	66	100	100	57	64	100	39	100	0	80	100	67	100	47	2	0	1	16	100	100	0	1	0	0	36	14	22
Bombyx	***mori***	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0	20	0	4	0	0
Bos	***taurus***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0
Branchiostoma	***floridae***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Callorhinchus	***milii***	72	100	100	18	22	48	49	25	100	32	44	31	38	58	59	100	65	26	81	37	30	100	49	21	69	100	54	81	35	100	46	46	100	37
Capitella	***sp.-2004***	12	0	22	0	3	0	4	12	5	2	5	0	0	0	2	6	0	0	0	2	18	0	0	24	0	7	0	14	23	21	4	10	23	21
Ciona	***intestinalis***	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2	100	100	4	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Ciona	***savignyi***	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	2	0	2	0	1	0	7	6	0	0	0	0	0	0	0
Crassostrea	***virginica***	1	0	0	3	33	19	74	23	82	13	27	3	66	4	100	1	0	100	10	68	42	0	0	0	100	6	74	25	3	72	15	100	100	100
Cyanea	***capillata***	25	100	100	100	100	100	100	100	100	100	100	78	100	89	100	100	100	100	100	100	100	100	100	77	100	100	100	100	29	100	100	100	100	2
Danio	***rerio***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Daphnia	***pulex***	0	0	0	0	0	0	0	0	5	2	1	5	0	0	1	1	0	0	0	1	0	2	0	1	0	0	0	0	0	0	0	0	0	1
Diplosoma	***listerianum***	100	100	17	67	100	100	100	100	100	100	100	100	100	100	100	2	100	100	100	100	100	2	100	13	100	56	100	100	100	100	100	100	100	100
Eptatretus	***burgeri***	0	0	0	1	0	47	9	18	13	1	2	0	100	38	4	1	77	40	7	10	100	0	0	0	20	2	59	0	0	87	0	0	22	12
Euprymna	***scolopes***	15	1	0	8	9	27	61	73	64	44	65	34	67	8	7	1	68	10	5	60	66	100	100	10	47	49	46	8	26	3	23	8	0	21
Gallus	***gallus***	0	23	0	0	0	0	0	0	100	0	0	0	0	0	100	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0
Halocynthia	***roretzi***	0	0	0	0	9	100	19	4	8	24	35	6	55	1	28	100	4	13	0	2	0	0	100	1	1	7	12	61	1	0	28	100	100	0
Helobdella	***robusta***	14	0	0	6	16	42	14	12	16	2	49	21	46	0	2	13	0	29	0	52	0	0	31	23	30	4	8	41	10	6	4	17	9	21
Holothuria	***glaberrima***	100	100	100	41	100	100	100	100	100	100	100	100	83	81	100	0	100	87	82	100	100	0	60	2	100	100	100	100	100	100	100	100	25	100
Hydra	***magnipapillata***	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	6	0	1	0	0	0	0	0	0	0	0	0	0
Hydractinia	***echinata***	100	100	0	10	17	39	23	100	35	49	30	6	47	0	56	2	43	100	1	16	14	0	0	1	100	56	47	0	10	0	85	43	27	8
Ixodes	***scapularis***	26	24	18	0	27	14	0	0	0	1	1	1	7	2	18	0	0	16	0	0	0	0	0	0	0	1	11	0	1	25	4	0	6	0
Litopenaeus	***vannamei***	1	3	20	29	66	1	1	45	29	1	0	21	66	3	0	0	1	61	8	12	3	1	0	2	45	7	87	9	71	100	0	1	0	30
Lottia	***gigantea***	100	23	18	0	29	32	32	12	52	5	63	35	82	10	5	16	3	42	0	21	1	16	42	30	45	28	27	0	10	48	37	80	25	34
Lumbricus	***rubellus***	0	0	13	0	27	100	34	72	60	64	37	70	39	0	100	1	33	100	17	100	81	0	1	0	19	9	0	27	1	100	100	100	0	0
Molgula	***tectiformis***	1	0	0	0	0	0	0	0	0	13	0	5	0	0	0	2	0	3	0	14	0	0	0	0	0	0	0	100	0	2	0	0	0	0
Monodelphis	***domestica***	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	1	0	0	0	0	0	0	0	0	0	1	0
Mytilus	***galloprovincial***	1	100	0	24	100	11	51	53	15	67	100	0	64	25	21	2	44	80	5	57	4	0	0	0	1	1	79	0	100	100	27	12	0	0
Nematostella	***vectensis***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0
Oikopleura	***dioica***	11	31	36	0	1	0	0	2	2	14	13	1	0	11	1	10	0	0	7	10	0	43	0	1	1	6	1	0	0	12	1	0	27	20
Oscarella	***carmela***	4	100	16	54	48	52	28	24	100	21	100	1	23	0	0	0	0	6	0	100	34	100	100	1	100	24	17	57	100	100	53	100	100	100
Paracentrotus	***lividus***	0	0	0	11	39	4	49	0	0	0	0	45	0	0	0	0	0	0	4	43	25	100	0	0	0	41	12	0	6	4	0	19	0	100
Pediculus	***humanus***	2	1	0	0	3	0	6	5	5	0	2	0	0	14	0	0	0	0	0	0	0	0	0	1	12	0	0	0	0	8	3	0	0	4
Petromyzon	***marinus***	0	0	0	0	78	10	0	0	37	0	100	100	0	0	3	0	3	17	15	8	28	0	0	0	0	0	0	0	19	0	0	0	23	39
Platynereis	***dumerilii***	100	100	0	28	100	42	10	64	45	100	100	44	13	35	23	100	0	17	17	73	25	100	0	2	0	31	29	5	100	100	100	0	100	100
Reniera	***sp.***	1	0	0	0	0	0	0	0	0	1	6	0	0	0	0	2	0	0	0	7	0	0	0	0	0	11	2	0	0	2	0	0	1	0
Rhipicephalus	***microplus***	0	0	100	0	0	0	0	0	0	1	1	21	13	0	0	0	0	0	16	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Saccoglossus	***kowalevskii***	0	0	0	0	29	0	0	9	0	0	0	27	0	0	0	2	3	0	2	4	0	0	0	0	0	0	10	0	0	0	0	0	0	0
Solaster	***stimpsonii***	100	0	0	45	61	50	41	72	100	76	69	100	100	68	0	0	100	100	60	100	100	2	100	11	34	31	77	100	100	23	100	100	10	100
Spadella	***cephaloptera***	3	0	100	75	100	100	100	100	83	100	100	100	100	100	100	2	83	100	100	100	100	0	0	27	100	100	100	100	100	7	100	0	0	0
Squalus	***acanthias***	0	0	25	23	58	38	60	61	26	30	62	44	37	47	34	16	64	12	100	58	41	1	100	0	51	51	60	14	27	48	6	21	100	100
Strongylocentrotus	***purpuratus***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
Suberites	***domuncula***	0	0	100	100	100	100	100	100	100	100	100	100	100	100	100	0	0	100	17	73	100	0	100	0	100	100	100	100	100	100	100	100	100	100
Tetraodon	***nigroviridis***	0	0	0	0	5	0	0	0	100	0	0	0	0	13	0	0	0	31	17	0	0	0	0	0	0	0	5	0	0	0	100	10	0	0
Tribolium	***castaneum***	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Xenopus	***tropicalis***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Xenoturbella	***bocki***	0	0	100	34	100	9	75	100	47	100	42	0	40	0	100	2	7	69	0	37	100	0	100	0	100	89	100	100	100	0	100	16	100	0
% missing positions		**17**	**20**	**18**	**15**	**30**	**27**	**25**	**28**	**35**	**26**	**32**	**25**	**30**	**18**	**26**	**10**	**23**	**31**	**19**	**30**	**28**	**15**	**21**	**5**	**29**	**23**	**30**	**23**	**23**	**27**	**26**	**26**	**24**	**25**
% missing OTUs		**14**	**18**	**14**	**6**	**18**	**16**	**12**	**14**	**20**	**16**	**18**	**12**	**14**	**6**	**20**	**8**	**10**	**20**	**10**	**16**	**18**	**14**	**18**	**0**	**20**	**12**	**14**	**16**	**18**	**20**	**20**	**18**	**20**	**20**
% chimeras		**0**	**0**	**2**	**4**	**2**	**6**	**4**	**4**	**6**	**6**	**6**	**6**	**4**	**4**	**0**	**2**	**8**	**2**	**2**	**4**	**2**	**0**	**0**	**2**	**4**	**2**	**8**	**2**	**0**	**0**	**0**	**4**	**2**	**2**
# amino-acid sites		**155**	**166**	**200**	**464**	**515**	**513**	**489**	**526**	**512**	**528**	**484**	**499**	**489**	**806**	**228**	**214**	**594**	**600**	**641**	**451**	**390**	**123**	**124**	**246**	**385**	**441**	**762**	**225**	**242**	**252**	**228**	**216**	**205**	**212**

		psmb-K	psmb-L	psmb-M	psmb-N	rad23	rla2-A	rla2-B	rpl1	rpl1b	rpl2b	rpl3	rpl4a	rpl5a	rpl6b	rpl7	rpl8	rpl9a	rpl2	rpl20	rpl21	rpl22	rpl23a	rpl24-A	rpl24-B	rpl25	rpl26	rpl27	rpl3	rpl30	rpl31	rpl32	rpl33a	rpl34	rpl35
Acropora	***millepora***	0	100	35	100	25	10	0	100	0	1	0	0	52	0	1	1	0	0	6	1	0	1	0	2	0	1	10	0	100	0	0	0	0	1
Ambystoma	***mexicanum***	0	0	0	0	9	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	26	0	0	0	0	0
Apis	***mellifera***	0	0	0	0	0	3	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	3	0	20	0	0	0	0	0	0	0	0	0	0
Aplysia	***californica***	10	0	0	0	0	1	4	0	0	0	0	0	0	0	0	1	2	0	0	0	0	1	16	0	0	0	0	0	0	0	0	0	2	7
Asterina	***pectinifera***	0	0	21	6	62	2	0	0	0	0	0	0	0	0	0	0	0	16	0	1	0	0	0	0	0	0	0	100	0	0	0	0	0	1
Bombyx	***mori***	0	0	24	1	4	2	0	0	0	0	0	0	0	0	0	3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Bos	***taurus***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Branchiostoma	***floridae***	0	0	0	0	0	0	3	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Callorhinchus	***milii***	59	100	100	100	10	100	100	10	9	56	100	100	28	57	37	100	38	62	78	71	21	48	100	58	22	0	31	29	100	61	53	28	100	100
Capitella	***sp.-2004***	5	1	10	1	30	11	24	0	24	7	0	0	0	25	1	15	0	5	4	0	55	1	33	20	21	0	1	0	3	25	2	6	41	38
Ciona	***intestinalis***	0	0	0	0	33	1	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	0	0	0
Ciona	***savignyi***	0	0	0	0	0	5	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Crassostrea	***virginica***	0	71	64	55	48	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Cyanea	***capillata***	100	100	100	100	100	4	3	19	0	0	0	0	64	0	1	1	100	100	2	1	100	1	14	0	1	0	0	75	0	0	0	0	0	1
Danio	***rerio***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
Daphnia	***pulex***	0	2	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	5	6	0	0	0	0	0	0	0	0	0	0	0	0
Diplosoma	***listerianum***	100	18	26	100	48	17	0	0	100	0	2	0	0	0	1	0	33	0	0	100	100	1	11	100	1	100	0	100	1	0	0	0	0	0
Eptatretus	***burgeri***	12	0	0	0	52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	3	0	0	0	0	0	0	0	0	0	0	0
Euprymna	***scolopes***	14	0	0	9	52	34	0	0	70	21	0	100	0	0	0	0	1	0	0	3	1	3	0	0	0	1	1	0	57	0	0	0	10	0
Gallus	***gallus***	0	100	0	0	9	0	0	0	0	0	0	0	0	100	100	11	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Halocynthia	***roretzi***	100	0	100	69	4	100	0	100	100	0	15	100	100	100	100	34	49	10	100	0	9	100	100	100	100	100	3	1	0	0	5	100	100	100
Helobdella	***robusta***	42	13	30	39	78	3	100	0	1	4	1	28	0	0	15	0	39	17	36	0	21	1	3	52	6	0	1	1	0	6	0	0	100	2
Holothuria	***glaberrima***	100	100	100	100	100	3	0	0	100	0	0	0	0	0	12	0	0	1	0	1	0	0	0	100	0	0	0	0	0	0	0	2	0	3
Hydra	***magnipapillata***	0	0	0	0	3	0	0	0	2	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Hydractinia	***echinata***	100	1	0	0	44	3	0	0	2	0	1	0	0	0	1	1	0	5	0	1	0	1	0	26	0	0	0	2	3	0	0	0	3	1
Ixodes	***scapularis***	0	0	0	0	0	3	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Litopenaeus	***vannamei***	28	47	58	0	31	5	0	0	0	0	1	0	0	0	1	0	2	0	0	0	0	0	2	100	6	0	0	0	0	0	1	0	3	0
Lottia	***gigantea***	28	33	66	0	80	39	100	0	24	62	34	100	0	72	20	73	38	18	74	100	55	41	32	54	22	58	31	12	12	63	21	6	100	38
Lumbricus	***rubellus***	1	0	40	45	0	3	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	13	0	0	0	0	0	0	0	5	0	0
Molgula	***tectiformis***	0	0	0	0	3	4	0	0	0	0	0	100	1	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	20	0	0	0	0
Monodelphis	***domestica***	0	0	0	100	0	100	0	0	0	0	0	0	0	0	0	0	100	0	100	0	0	100	1	0	0	0	0	0	3	0	0	0	0	1
Mytilus	***galloprovincial***	49	4	18	0	50	4	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	78	0	0	0	0	0	0	0	4	10	1
Nematostella	***vectensis***	0	0	0	0	0	3	0	0	2	1	0	0	0	0	1	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1
Oikopleura	***dioica***	7	1	0	6	5	7	1	0	0	0	1	4	11	6	1	4	0	3	60	0	0	0	0	0	2	0	1	0	0	1	7	37	0	0
Oscarella	***carmela***	0	0	100	0	48	100	9	0	4	0	100	100	100	0	0	1	100	0	1	0	0	100	100	100	0	100	100	1	100	100	100	3	0	1
Paracentrotus	***lividus***	0	100	0	6	100	5	100	100	100	100	100	100	0	0	0	0	0	6	75	100	0	0	100	0	100	100	1	34	100	100	100	100	3	100
Pediculus	***humanus***	0	1	0	0	9	3	0	0	5	7	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	3	5	0	1	20	1
Petromyzon	***marinus***	0	0	0	0	71	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Platynereis	***dumerilii***	0	36	100	100	100	100	4	0	57	100	0	0	0	100	100	100	0	0	0	100	0	4	100	100	100	0	100	0	100	100	0	100	100	2
Reniera	***sp.***	0	0	0	37	50	0	0	0	0	0	1	0	0	0	0	1	0	0	0	3	12	0	1	0	0	0	0	2	0	0	0	0	0	0
Rhipicephalus	***microplus***	0	0	0	0	0	3	3	0	0	0	0	0	0	0	0	1	8	0	0	0	0	0	0	0	5	1	0	0	20	0	0	0	0	100
Saccoglossus	***kowalevskii***	0	0	0	0	17	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Solaster	***stimpsonii***	100	0	100	100	100	100	100	0	0	100	0	0	100	0	100	100	51	0	100	100	0	100	100	100	0	100	0	4	100	100	100	100	100	100
Spadella	***cephaloptera***	100	0	100	60	100	46	0	0	2	0	1	2	0	3	0	1	0	15	0	1	0	0	0	100	54	0	1	6	0	0	0	0	0	1
Squalus	***acanthias***	0	1	3	11	69	2	0	31	100	72	100	0	100	0	0	1	100	6	100	0	100	19	0	0	22	2	0	0	0	100	0	100	100	100
Strongylocentrotus	***purpuratus***	0	0	0	0	100	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0
Suberites	***domuncula***	100	100	100	100	100	2	0	0	1	0	0	2	7	0	0	0	0	4	0	0	0	1	0	0	0	0	0	2	0	0	0	0	0	0
Tetraodon	***nigroviridis***	100	35	0	0	7	0	0	100	0	0	0	25	0	100	0	0	2	0	0	0	100	100	0	7	0	17	0	0	100	6	0	0	0	7
Tribolium	***castaneum***	0	0	0	0	0	3	0	0	0	0	0	0	0	0	3	0	0	0	0	1	0	0	0	0	4	0	0	0	0	0	0	0	0	0
Xenopus	***tropicalis***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Xenoturbella	***bocki***	100	100	25	100	100	3	4	0	0	0	0	0	0	0	1	1	74	0	0	3	100	3	0	0	21	0	0	0	0	0	0	0	100	3
% missing positions		**25**	**21**	**26**	**26**	**36**	**17**	**11**	**9**	**14**	**10**	**9**	**15**	**11**	**11**	**10**	**9**	**14**	**5**	**15**	**12**	**17**	**12**	**14**	**22**	**10**	**11**	**6**	**7**	**16**	**13**	**8**	**12**	**19**	**14**
% missing OTUs		**20**	**16**	**18**	**20**	**18**	**12**	**10**	**8**	**10**	**6**	**8**	**14**	**8**	**8**	**8**	**6**	**8**	**2**	**8**	**10**	**14**	**10**	**12**	**16**	**6**	**10**	**4**	**4**	**14**	**10**	**6**	**10**	**18**	**12**
% chimeras		**2**	**2**	**2**	**2**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**2**	**0**	**0**	**0**	**0**	**0**	**2**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
# amino-acid sites		**234**	**196**	**195**	**194**	**216**	**92**	**76**	**213**	**169**	**163**	**179**	**123**	**204**	**173**	**175**	**183**	**190**	**248**	**160**	**154**	**96**	**139**	**123**	**137**	**126**	**135**	**137**	**393**	**105**	**108**	**129**	**106**	**108**	**123**

		rpl36	rpl37a	rpl38	rpl4 3b	rpl4 B	rpl5	rpl6	rpl7A	rpl9	rpp0	rps1	rps10	rps11	rps13a	rps14	rps15	rps16	rps17	rps18	rps19	rps20	rps22a	rps23	rps24	rps25	rps26	rps27	rps27a	rps28a	sadhchydrolase-E1	sap40	suca	tif2a	u2snrnp
Acropora	***millepora***	1	0	49	0	21	0	5	0	1	0	0	2	100	0	0	0	0	0	0	1	0	0	0	0	5	2	0	0	20	0	24	2	0
Ambystoma	***mexicanum***	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	16	6	0
Apis	***mellifera***	0	0	100	1	0	0	2	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	11	0
Aplysia	***californica***	0	0	32	1	0	0	1	0	1	0	0	3	0	0	1	0	0	2	3	0	0	0	0	3	4	1	0	0	0	1	0	0	21
Asterina	***pectinifera***	1	100	0	0	100	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	100	3	0	57	0	100	24	0
Bombyx	***mori***	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2	0	1	0	0	1	63
Bos	***taurus***	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Branchiostoma	***floridae***	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	0
Callorhinchus	***milii***	26	100	100	57	45	23	18	37	100	6	46	18	71	40	100	100	43	27	100	100	100	65	100	43	33	100	100	39	34	17	61	46	60	56
Capitella	***sp.-2004***	22	42	12	1	7	1	11	0	7	17	17	18	1	13	15	1	5	12	16	0	0	0	0	43	1	3	8	0	2	2	0	0	24	8
Ciona	***intestinalis***	0	9	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	8	1	0	0	0	1
Ciona	***savignyi***	0	9	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	1
Crassostrea	***virginica***	0	42	42	7	0	9	1	0	1	21	0	0	0	1	1	0	0	0	3	0	0	1	2	4	0	0	0	0	1	4	53	100	1
Cyanea	***capillata***	0	0	2	0	100	75	21	0	4	44	100	100	0	0	100	0	100	0	0	8	0	0	0	0	0	100	0	0	2	100	100	100	100	100
Danio	***rerio***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0
Daphnia	***pulex***	0	0	0	0	0	0	2	0	0	0	0	2	0	0	0	0	0	0	0	0	1	0	0	0	0	1	2	1	0	0	0	1	0
Diplosoma	***listerianum***	0	12	0	100	100	100	100	10	100	0	4	1	4	0	100	1	1	100	0	7	28	0	0	0	1	0	2	2	0	50	0	31	100	100
Eptatretus	***burgeri***	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	11	100
Euprymna	***scolopes***	100	48	100	100	11	0	2	0	0	9	0	5	1	0	0	0	0	8	0	0	0	12	1	0	100	100	100	0	0	63	10	37	46	23
Gallus	***gallus***	0	0	14	0	0	0	0	0	0	0	8	0	0	16	0	100	0	0	100	100	0	0	0	100	0	0	0	0	0	0	0	0	7	0
Halocynthia	***roretzi***	0	100	100	100	26	0	100	100	3	0	100	0	100	100	100	14	100	100	100	100	100	100	100	100	100	100	100	2	100	9	2	100	54	0
Helobdella	***robusta***	20	1	43	2	0	4	26	5	0	25	0	1	23	0	1	0	66	0	15	13	2	1	0	26	4	1	0	0	0	24	0	13	73	33
Holothuria	***glaberrima***	40	2	0	0	1	0	1	0	0	0	0	9	0	5	0	0	0	7	1	1	0	5	0	0	0	0	1	0	0	56	0	100	100	100
Hydra	***magnipapillata***	0	0	0	0	0	0	3	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	1	0	0	0	0
Hydractinia	***echinata***	0	0	0	0	1	0	3	0	1	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	43	0	100	41	0
Ixodes	***scapularis***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	2	0	0	0	1	0
Litopenaeus	***vannamei***	0	0	0	0	0	0	3	0	0	16	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	2	0	1	0	0	0	78	3	100
Lottia	***gigantea***	46	46	43	51	16	23	45	31	24	71	32	18	22	40	32	49	66	13	57	0	37	100	28	43	33	2	10	0	100	9	1	16	36	44
Lumbricus	***rubellus***	0	41	0	0	0	0	1	1	2	0	0	0	1	0	1	0	0	0	0	0	0	0	0	2	2	0	0	0	0	2	0	100	67	100
Molgula	***tectiformis***	0	9	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	2	0	1	0	0	100	0
Monodelphis	***domestica***	0	0	0	1	0	0	0	0	0	0	0	0	0	100	0	0	0	100	100	0	0	1	0	100	1	2	0	2	0	100	0	0	0	0
Mytilus	***galloprovincial***	41	57	3	0	0	1	1	8	1	0	2	0	0	5	1	0	0	0	0	0	0	0	1	4	4	0	1	0	0	100	0	0	6	1
Nematostella	***vectensis***	0	0	0	0	0	0	3	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	0	2	0
Oikopleura	***dioica***	60	2	2	0	2	1	9	2	0	1	0	42	0	0	5	1	0	10	0	2	1	4	0	62	11	34	2	26	0	20	0	6	53	9
Oscarella	***carmela***	100	100	0	100	33	0	2	0	100	0	0	39	0	0	1	100	0	100	100	100	100	100	100	2	100	100	2	0	100	18	0	0	100	1
Paracentrotus	***lividus***	100	100	100	0	0	100	100	100	0	11	100	0	100	100	0	100	100	1	100	0	100	0	0	100	0	100	0	100	100	62	5	0	27	0
Pediculus	***humanus***	20	0	100	0	0	0	8	16	0	5	2	0	0	5	0	0	41	12	1	0	0	0	1	0	3	1	56	1	0	2	0	0	4	8
Petromyzon	***marinus***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2	0	0	1	0
Platynereis	***dumerilii***	100	100	100	100	0	49	100	100	100	0	100	100	0	1	1	100	100	100	100	100	0	0	0	2	100	100	100	100	100	36	100	16	12	0
Reniera	***sp.***	0	0	2	0	0	0	2	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	3	7	0	14	2	1
Rhipicephalus	***microplus***	100	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	2	0	0	0	0	1	0	0	3	100	1	100	0	0	0	1	26
Saccoglossus	***kowalevskii***	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	3	0	0	0	0	0	0	0	29	8
Solaster	***stimpsonii***	100	100	100	19	0	1	100	0	32	8	5	0	100	100	100	100	100	100	0	100	0	100	100	100	100	100	0	100	100	100	0	100	100	100
Spadella	***cephaloptera***	0	0	0	0	100	22	1	1	1	19	14	3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	0	8	100	0	100	100	7
Squalus	***acanthias***	100	0	0	0	0	0	26	100	100	14	100	0	24	100	100	0	0	100	21	0	0	100	100	0	0	0	100	0	0	56	0	0	36	2
Strongylocentrotus	***purpuratus***	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	8
Suberites	***domuncula***	0	0	100	0	0	0	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	100	0	100	100	100
Tetraodon	***nigroviridis***	6	0	100	27	16	0	64	0	0	0	0	0	0	0	100	100	0	0	26	13	100	0	100	16	1	1	0	10	0	0	100	0	0	0
Tribolium	***castaneum***	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Xenopus	***tropicalis***	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	69	0	0	0	0	0	0
Xenoturbella	***bocki***	0	0	2	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	2	0	0	3	0	0	0	2	0	1	100	100
% missing positions		**19**	**20**	**24**	**13**	**11**	**8**	**15**	**10**	**11**	**5**	**12**	**7**	**11**	**10**	**19**	**15**	**14**	**14**	**16**	**15**	**11**	**12**	**12**	**13**	**14**	**17**	**18**	**9**	**15**	**23**	**9**	**25**	**32**	**24**
% missing OTUs		**14**	**14**	**20**	**10**	**8**	**4**	**10**	**8**	**10**	**0**	**10**	**4**	**8**	**8**	**18**	**14**	**10**	**12**	**14**	**14**	**10**	**10**	**12**	**8**	**12**	**16**	**16**	**6**	**14**	**12**	**8**	**18**	**20**	**18**
% chimeras		**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**2**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**2**	**0**	**0**
# amino-acid sites		**85**	**81**	**65**	**88**	**326**	**275**	**186**	**207**	**183**	**289**	**244**	**119**	**140**	**151**	**149**	**139**	**138**	**119**	**152**	**132**	**106**	**130**	**143**	**122**	**92**	**101**	**84**	**155**	**61**	**424**	**212**	**291**	**304**	**179**

		vatc	vate	vatpased	vdac2	% missing positions	% missing genes	% chimeras
Acropora	***millepora***	28	5	10	0	**25**	**7**	**2**
Ambystoma	***mexicanum***	45	10	3	0	**10**	**3**	**3**
Apis	***mellifera***	0	0	0	0	**1**	**1**	**0**
Aplysia	***californica***	9	8	2	0	**4**	**0**	**0**
Asterina	***pectinifera***	100	100	100	0	**43**	**19**	**0**
Bombyx	***mori***	0	0	0	7	**2**	**0**	**0**
Bos	***taurus***	0	0	0	0	**2**	**1**	**0**
Branchiostoma	***floridae***	0	0	1	0	**0**	**0**	**0**
Callorhinchus	***milii***	52	100	63	47	**55**	**30**	**0**
Capitella	***sp.-2004***	18	5	24	0	**7**	**0**	**0**
Ciona	***intestinalis***	9	4	0	0	**5**	**2**	**0**
Ciona	***savignyi***	0	0	0	0	**1**	**0**	**0**
Crassostrea	***virginica***	51	100	13	1	**26**	**8**	**8**
Cyanea	***capillata***	100	28	100	82	**72**	**46**	**0**
Danio	***rerio***	0	0	0	0	**1**	**2**	**0**
Daphnia	***pulex***	0	0	0	0	**0**	**0**	**0**
Diplosoma	***listerianum***	100	30	100	0	**66**	**44**	**0**
Eptatretus	***burgeri***	68	2	100	1	**18**	**6**	**0**
Euprymna	***scolopes***	100	1	100	0	**27**	**10**	**3**
Gallus	***gallus***	0	0	0	0	**8**	**8**	**0**
Halocynthia	***roretzi***	1	13	4	0	**36**	**42**	**0**
Helobdella	***robusta***	60	100	27	55	**18**	**3**	**8**
Holothuria	***glaberrima***	100	100	100	100	**60**	**36**	**0**
Hydra	***magnipapillata***	0	0	0	0	**0**	**0**	**0**
Hydractinia	***echinata***	100	100	100	1	**25**	**9**	**8**
Ixodes	***scapularis***	0	6	0	0	**3**	**0**	**0**
Litopenaeus	***vannamei***	29	8	61	42	**18**	**3**	**13**
Lottia	***gigantea***	81	0	29	38	**32**	**7**	**1**
Lumbricus	***rubellus***	64	0	0	1	**26**	**8**	**6**
Molgula	***tectiformis***	1	0	0	0	**4**	**3**	**0**
Monodelphis	***domestica***	0	0	0	0	**7**	**11**	**0**
Mytilus	***galloprovincial***	11	0	60	5	**25**	**6**	**3**
Nematostella	***vectensis***	1	0	1	0	**0**	**0**	**0**
Oikopleura	***dioica***	12	14	1	100	**8**	**1**	**0**
Oscarella	***carmela***	0	100	1	0	**39**	**40**	**1**
Paracentrotus	***lividus***	32	0	0	0	**29**	**34**	**0**
Pediculus	***humanus***	0	0	7	0	**3**	**1**	**0**
Petromyzon	***marinus***	29	0	4	0	**10**	**2**	**0**
Platynereis	***dumerilii***	34	100	100	0	**48**	**48**	**0**
Reniera	***sp.***	28	0	0	0	**2**	**0**	**0**
Rhipicephalus	***microplus***	0	20	0	0	**4**	**5**	**1**
Saccoglossus	***kowalevskii***	0	0	45	0	**3**	**0**	**0**
Solaster	***stimpsonii***	100	0	100	100	**66**	**58**	**0**
Spadella	***cephaloptera***	100	38	100	19	**56**	**29**	**1**
Squalus	***acanthias***	46	16	77	18	**40**	**23**	**0**
Strongylocentrotus	***purpuratus***	0	0	0	100	**2**	**3**	**0**
Suberites	***domuncula***	100	100	0	100	**56**	**37**	**0**
Tetraodon	***nigroviridis***	10	0	0	1	**14**	**13**	**0**
Tribolium	***castaneum***	0	0	0	0	**0**	**0**	**0**
Xenopus	***tropicalis***	0	0	0	0	**0**	**0**	**0**
Xenoturbella	***bocki***	80	100	0	0	**36**	**21**	**0**
% missing positions		**33**	**24**	**28**	**16**			
% missing OTUs		**18**	**20**	**20**	**10**			
% chimeras		**0**	**0**	**0**	**2**			
# amino-acid sites		**336**	**200**	**210**	**276**			