SNAREing the basis of multicellularity: Consequences of protein family expansion during evolution

Tobias H. Kloepper[1,2], C. Nickias Kienle[1] and Dirk Fasshauer[1]

1 Research Group Structural Biochemistry, Department of Neurobiology, Max-Planck-Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany, e-mail D.F.: dfassha@gwdg.de , phone: +49-551-201 1637, Fax: +49-551-201 1499; C.N.K.: nkienle@gwdg.de, e-mail T.H.K.: kloepper@informatik.uni-tuebingen.de

2 Center for Bioinformatics, University of Tübingen, Sand 14, 72076 Tübingen, Germany

Correspondence should be addressed to D.F.

Submission as Research Article

Running head: Evolution of the SNARE protein family in metazoa

Key words: SNARE, membrane fusion, secretion, metazoa, multicellularity, vertebrata

Abstract

Vesicle trafficking between intracellular compartments of eukaryotic cells is mediated by conserved protein machineries. In each trafficking step, fusion of the vesicle with the acceptor membrane is driven by a set of distinctive SNARE proteins that assemble into tight four-helix bundle complexes between the fusing membranes. During evolution about twenty primordial SNARE types were modified independently in different eukaryotic lineages by episodes of duplication and diversification. Here we show that two major changes in the SNARE repertoire occurred in the evolution of animals, each reflecting a main overhaul of the endomembrane system. In addition, we found several lineage-specific losses of distinct SNAREs, particularly in nematodes and platyhelminthes. The first major transformation took place during the transition to multicellularity. The primary event that occurred during this transformation was an increase in the numbers of endosomal SNAREs, but the SNARE-related factor Lgl also emerged. Apparently, enhanced endosomal sorting capabilities were an advantage for early multicellular animals. The second major transformation during the rise of vertebrates resulted in a robust expansion of the secretory set of SNAREs, which may have helped develop a more versatile secretory apparatus.

Introduction

A defining feature of eukaryotic cells is their complex system of internal compartments. Directional transport between the compartments is mediated by vesicles that bud from a donor organelle and then fuse with an acceptor organelle. As each of the internal compartments serves a specific and vital cellular function, it must maintain its identity during vesicle trafficking. This is achieved by specific protein machineries that tightly regulate each transport step (Bonifacino and Glick 2004; Behnia and Munro 2005; Cai et al. 2007). It has been recognized that several of the key machineries involved in vesicle trafficking steps appear to be the result of gene duplications, followed by diversifications (Cavalier-Smith 2002; Dacks and Field 2007). For example, in each trafficking step the core of the protein machinery involved in the fusion process is composed of members of the so-called SNARE (Soluble NSF Attachment Protein REceptors) protein family (Hong 2005; Jahn and Scheller 2006). SNARE proteins are small cytoplasmatically orientated membrane-associated proteins with a relatively simple domain architecture. In most SNARE proteins, an extended coiled coil segment, the so-called SNARE motif, is directly adjacent to a single trans-membrane domain at the C-terminal end. As a common mechanism, the SNARE motifs of heterologous sets of SNARE proteins assemble in a zipper-like fashion into tight four-helix bundle units between two membranes, a process that clamps the membranes together and initiates fusion. The SNARE machinery has diversified only modestly in the course of eukaryotic evolution. In fact, we were able to show that a basic set of twenty types of SNARE proteins is conserved in all eukaryotes, suggesting that they constitute the set of SNARE genes of the proto-eukaryotic cell (Kloepper et al. 2007).

In the previous study, we recognized that the basic set of SNARE genes had been independently modified by paralogous expansion in several different eukaryotic lineages (Kloepper et al. 2007). So far, an increase in the number of SNARE genes has been observed mainly in animals and in land plants (Bock et al. 2001; Dacks and Doolittle 2004; Yoshizawa et al. 2006; Dacks and Field 2007; Sanderfoot 2007). A genome-wide inspection of the first four eukaryotic species sequenced came to the conclusion that mammals (e.g. *Homo sapiens*) contain an enlarged set of secretory SNAREs compared to invertebrates (e.g. the fruitfly *Drosophila melanogaster* and the round worm *Caenorhabditis elegans*). Since the number of SNARE genes in the two invertebrates was comparable to that of the yeast *Saccharomyces cerevisiae*, the authors concluded that multicellular organisms do not necessarily require an enlarged repertoire of SNARE proteins (Bock et al. 2001). Rather, they proposed, the enlargement of the secretory SNAREs took place in mammals. In other studies, by contrast, it was suggested that in both, animals and land plants, the increase in the number of SNAREs, in particular the ones involved in secretion, occurred much earlier in evolution, paralleling the leap from single to multicellular organisms (Sanderfoot et al. 2000; Dacks and Doolittle 2004). On that account, the idea has been put forward that the secretory pathway of multicellular organisms requires an additional level of complexity compared to that of their single-cell predecessors (Dacks and Field 2007; Sanderfoot 2007). Only with an increased number of secretory SNARE proteins, according to this idea, an individual cell within a tissue is able to deliver material to different regions of its plasma membrane. This line of reasoning, however, neglects the fact that the capability to transport and secrete material to specific regions of the plasma membrane is an essential part of the life cycle of every eukaryotic cell, multicellular or unicellular (Soldati and Schliwa 2006).

For example, polarized secretion is the basis of the well-studied process of asymmetric enlargement of the budding daughter cell during cytokinesis in the single-cell yeast *Saccharomyces cerevisae* (Brennwald and Rossi 2007). Furthermore, multicellular species are found throughout the eukaryotic kingdoms, for example in the lineage of fungi, yet all inspected fungi species appear to maintain an almost unchanged basic set of SNARE proteins throughout evolution (Kloepper et al. 2007). Moreover, considerably enlarged SNARE repertoires, including extra secretory SNAREs, have been detected in several unicellular eukaryotes, for example for the ciliate *Paramecium tetraurelia* with its multifaceted endomembrane system (Kloepper et al. 2007). Together, these observations cast doubt on a strict correlation between the increase in the number of secretory SNARE genes and the emergence of multicellularity.

So far, all accounts of SNAREs in animals were based on only few model organisms. In addition, in the previous studies, some more-derived SNAREs have not been included (Bock et al. 2001; Dacks and Doolittle 2004; Yoshizawa et al. 2006). Thus, it still remains unclear exactly which SNARE factors have been added at which point during animal evolution. Clearly, in order to gain deeper insights into adaptations of the metazoan endomembrane system, a much broader database containing the SNARE repertoires of species that represent most major branches of the complex evolutionary tree of animals must be used. In addition, not only the number of SNARE genes but, more importantly, the evolutionary history of the different functional SNARE types needs to be reconstructed more thoroughly. Only this knowledge will allow us to address a series of pressing questions: How did the transition to multicellularity affect the endomembrane system? Are the changes mostly observable in the SNAREs actually involved in the exocytotic step of the secretory pathway? Is it possible to associate the rise or loss of certain SNARE factors with specific lineages within animals? Are the routes of the secretory pathway indeed comparable between various animal species?

In order to answer these questions, we considerably expanded our previous collection (Kloepper et al. 2007) by integrating sequences from a multiplicity of different animal species, making use of the abundance of new genomic data that have been established lately. A point-by-point account of the SNARE repertoires was facilitated by our recent classification of the multigene SNARE family, with its largely improved ability to distinguish between orthologues and paralogues (Kloepper et al. 2007). Moreover, this approach allows us to integrate new SNARE sequence data from genome and EST projects quickly. Only this allowed us to pinpoint the changes in the SNARE sets throughout all metazoan lineages investigated. In addition, we constructed phylogenetic trees for each individual SNARE subgroup. All these factors enabled us to infer a probable sequence of events for the observed changes in the SNARE repertoire during animal evolution.

We discovered that during metazoan evolution, two major changes in the SNARE repertoire have occurred. The first major change took place during the transition from single to multicellular organisms. During this transition, mostly SNARE proteins involved in endosomal trafficking were duplicated and diversified, as well as some SNAREs involved in secretion. In addition, SNARE-related factors like the tumor suppressor lethal giant larvae (Lgl), which plays a key role in establishing epithelial cell polarity, originated. The second major change took place in the lineage of vertebrates, where the secretory SNARE repertoire in particular enlarged drastically.

Materials and Methods

Sequences

In a recent classification of SNARE proteins, we published a set of more than 2000 protein sequences from 145 eukaryotic species, including 811 sequences from 59 metazoan species (Kloepper et al. 2007). These sequences were classified into twenty distinct groups. Here, we used the established HMM models (Kloepper et al. 2007) to classify the SNAREs of newly available metazoan genomes and EST libraries. We gathered 1234 additional sequences from the nr-database at NCBI and various genome projects (DOE Joint Genome Institute (JGI); Baylor College of Medicine, J. Craig Venter Institute and Broad Institute) and various EST databases.
After removing sequences with more than one occurrence, sequences which seemed to be misassembled, or sequences which failed an a visual verification, we obtained a set of 1774 unique SNARE sequences from 177 metazoan species. In addition, we identified sequences of two SNARE-related protein families using an extensive blast search on the nr-database at NCBI, identifying 37 sequences of lethal giant larvae (Lgl) and 41 sequences of Sec22-like. Both proteins exhibit a very high N-terminal sequence similarity to tomosyn and Sec22, respectively, but, interestingly, have lost their SNARE domain. Since we were especially interested in pinpointing the emergence of innovations within the SNARE set of metazoans, we supplemented our collected sequences with a set of 31 sequences from two choanoflagellates and two sequences from two ichthyosporeans (JGI and EST data). All species and SNARE sequences used in this study are listed in Suppl Table 1. These newly classified sequences were integrated into the SNARE database (Kloepper et al. 2007) for public access, which can be accessed via our projects homepage (http://bioinformatics.mpibpc.mpg.de/snare/). Any sequence set used in the phylogenetic analysis was aligned using muscle (Edgar 2004) with standard settings. Sites with more than 25 percent gaps were removed from the alignments. After removing those sides, we removed aligned sequences which contained more than 50 percent gaps. The reduced alignments are available via the supplementary section of our projects homepage.

Phylogeny

The phylogenetic reconstruction was separated into two parts and was carried out essentially as described (Kloepper et al. 2007). Firstly, we reconstructed a tree from the conserved alignments using IQPNNI (Important Quartet Puzzling and Nearest Neighbor Interchange) (Vinh le and Von Haeseler 2004) with a gamma distribution as a model for rate heterogeneity. The estimation for the gamma distribution parameter used four rate categories. Additionally, the proportion of invariable sites was estimated from the data, an the Jones, Taylor and Thornton (JTT-) distance matrix was used as a substitution matrix, and the stopping rule of the algorithm was used but the algorithm had to run for at least the suggested number of iterations. All other settings of the application were set to default values. For each edge of the constructed tree, we estimated the confidence using Likelihood-Mapping. Secondly, we used the *phylip* package (Felsenstein 1998) to apply a distance-based bootstrap analysis with 1000 replicates. We used standard settings for seqboot, the JTT distance matrix once again and also a gamma distribution (with parameter approximation from tree-puzzle) for protdist and standard options for neighbour. Whenever necessary, we used a random seed of nine. Because bootstrap values have been shown to be systematically

biased, we used the almost unbiased (AU) test (Shimodaira 2002) to correct for this. The site wise log-likelihoods needed for the AU-test were obtained using a modified version of *phyml* (Guindon and Gascuel 2003) and the test was performed using *consel* (Shimodaira and Hasegawa 2001). We joined the results of both estimations using the tree of IQPNNI as a starting point and labeled the inner edges of the tree with their Likelihood-Mapping and corrected bootstrap support values. All trees can be downloaded via the supplemental section of our projects homepage. The trees are stored in Nexus format and can be interactively explored using, for example, *SplitsTree 4* (Huson and Bryant 2006).

Results

In order to catalog the SNARE inventory of animals, we complemented our previous collection (Kloepper et al. 2007) with newly published genomes and ESTs from different animal species, obtaining an overall set of 1839 SNARE sequences from 177 different metazoan species. This data set was complemented by 31 sequences from two choanoflagellates and 2 sequences from two ichthyosporeans, both of which are believed to be closely related to metazoans. All sequences and species used in this study are listed in Suppl. Table 1 and can also be found in our SNARE data base on our projects homepage. The data set comprises about 55 species with nearly complete sets of SNAREs (Suppl. Table 2). These sequences were classified into their corresponding basic subgroups according to our recently established HMM profiles. The classification scheme served as the framework to reconstruct the history of these changes. It is based on the highly conserved SNARE motif, as the other sequence regions are usually less conserved. However, we observed that within the animal kingdom, other sequence regions are also usually well conserved. Therefore, we constructed alignments of the entire sequences for each subgroup using *muscle* (Edgar 2004). Sites with more than 25 percent gaps were removed from the alignments. In addition, we excluded all sequences which contained more than 50 percent gaps. The length of these conserved alignments varied from 99 sites for the R.IV subgroup to 292 sites for the Qa.I subgroup.

Based on these alignments, we constructed phylogenetic trees for each of the fundamental SNARE types. When we used SNARE protein sequences from several other eukaryotic kingdoms as an outgroup, we observed that the different SNARE sequences from animals usually congregated within the phylogenetic trees and were well separated from the outgroup (data not shown), substantiating the view that animals constitute a monophyletic group.

The secretory SNARE set of Choanoflagellates is metazoan-like

The multicellular progenitor of animals, the urmetazoan, is believed to have evolved from an unicellular flagellated protist, probably closely related to extant heterotrophic, single-cell choanoflagellates (Brooke and Holland 2003; King 2004). In order to investigate whether SNARE genes were modified during this pivotal transition, we collected SNARE sequences from the genome of the choanoflagellate *Monosiga brevicollis* and supplemented these data with EST sequences from another choanoflagellate, *Monosiga ovata*. We detected a relatively simple repertoire of SNARE proteins in *M. brevicollis*, in principal consisting, except for a few duplications, of the basic set of SNARE proteins we had recently deduced for the proto-eukaryotic cell. Comparable basic sets of SNAREs can be found in green algae and fungi, for example. It should be noted, however, that we cannot rule out the possibility that the genome data of *M. brevicollis* is incomplete or that *M. brevicollis* has secondarily lost some SNARE factors.

In general, our phylogenetic analysis places *M. brevicollis* closer to metazoans than to other eukaryotes, supporting the notion that choanoflagellates are indeed closely related to the unicellular ancestor of animals. In particular, the secretory SNAREs of choanoflagellates are closely related to the set of animal SNAREs believed to mediate regulated secretion. For example, the SNAP-25 like sequences (Qbc-SNARE) from choanoflagellates contain a stretch of cysteines at the C-terminal end of their Qb-SNARE motif. These cysteines are known to be palmitoylated and serve as a membrane-anchor for animal SNAP-25 (Veit et al. 1996).

An enlarged set of endosomal SNAREs in multicellular animals

In order to reconstruct the SNARE repertoire of the metazoan ancestor, we compared the set of SNARE genes of lower metazoans with that of choanoflagellates. Remarkably, in comparison to the basal SNARE set of *M. brevicollis*, we detected a considerably enlarged repertoire of SNARE and SNARE-related genes in cnidarians, particularly in *Nematostella vectensis,* and also in the placozoan *Trichoplax adhaerens,* bringing to light a clear expansion of the SNARE repertoire during the transition from single-cell to multicellular organisms. Notably, both phyla exhibit a very similar SNARE repertoire, including several new molecular "inventions" (see Fig. 1a for the current view of the phylogenetic relationships of metazoans). Remarkably, we observed that the SNARE sets of the lower metazoans are very similar to those of various bilaterians, suggesting that the extended SNARE repertoire was already present in the last common ancestor of lower metazoans and bilaterians.

Our rigorous classification analysis was able to delineate which of the original functional SNARE types were duplicated during this first phase of expansion. In detail, we found at least nine evolutionary novel SNARE types in lower metazoans compared to single-cell choanoflagellates (Fig. 1b; for details, see also Supp. Table 2), suggesting the presence of a more multifaceted membrane trafficking system in metazoans in general. Notably, the exact function of most of these new factors and whether they were ubiquitously expressed is not entirely clear yet. Remarkably, most of these changes occurred in SNAREs involved in endosomal trafficking steps.

Besides the canonical R-SNARE Vamp7 (also referred to as Ti-Vamp; R.III-type), two additional endosomal R-SNAREs, Vamp7-like and Vamp4, were detected in lower metazoans (Fig. 1b). In animals, Vamp7 is involved in endosomal trafficking towards lysosomes (Advani et al. 1998; Advani et al. 1999), but it has also been implicated in secretion from lysosomal compartments (Martinez-Arca et al. 2003; Marcet-Palacios et al. 2007; Pocard et al. 2007). The function of the additional, R-SNARE Vamp7-like (not described to date), which is only present in lower metazoans, is unknown. However, its domain structure is almost identical to Vamp7, suggesting a similar function for both factors. Interestingly, in the phylogenetic tree (Fig. 2a), Vamp7-like appears to be closely related to endobrevin/Vamp8, which only exists in vertebrates. Endobrevin does not contain an N-terminal extension. It is therefore possible that Vamp7-like has lost the entire N-terminal profilin-like domain, giving rise to the novel R-SNARE endobrevin in the vertebrate lineage. Endobrevin was originally identified as an endosomal R-SNARE (Advani et al. 1998; Wong et al. 1998b), but later has been shown to be involved in regulated secretion of pancreatic zymogen granules (Wang et al. 2004; Wang et al. 2007). The third metazoan R.III-type SNARE, Vamp4, contains a shorter N-terminal extension comprising a highly conserved double-leucine motif followed by an acidic cluster (Advani et al. 1998). This region is important for its interaction with the cytosolic coat adaptor, AP-1, that targets Vamp4 to the TGN (Peden et al. 2001; Hinners et al. 2003; Zeng et al. 2003; Tran et al. 2007).

In choanoflagellates, we found only one Qa.III.b-SNARE, Syx7, whereas two additional endosomal syntaxins were found in most metazoans, Syx17 and Syx20 (Fig. 1b). Syx17 is the most deviated SNARE of the Qa.III.b-group in animals (Fig. 2b). Originally, Syx17 had been implicated in trafficking towards the smooth-surfaced ER (Steegmaier et al. 2000), but this function is debated (Zhang et al. 2005). Notably, and in contrast to all other known SNARE proteins, all Syx17 possess two adjacent transmembrane regions at the C-terminal end (Steegmaier et al. 2000). Syx20

is a SNARE gene we had first discovered in vertebrates only (Kloepper et al. 2007). A more detailed phylogenetic analysis of all Qa.III.b-type SNAREs (Fig. 2b) in animals now indicates that this factor is present in several other branches of metazoans as well. In insects, this factor originally has been termed Syntaxin 13. This nomenclature is somewhat confusing, since the name Syx13 (also called Syx12) (Advani et al. 1998; Tang et al. 1998d) is usually reserved for a Qa.III.b-SNARE closely related to Syx7 that is found only in vertebrates (Wong et al. 1998a) (see also below). We therefore propose to generally use the name Syx20 for this Qa.III.b-type.

Furthermore, two independent genes encoding for Qb.III.b-SNAREs, Vti1a and Vti1b (Advani et al. 1998; Xu et al. 1998), were found in most metazoans (Fig. 1b). Both are thought to function in early and late endosomal trafficking, respectively. A set of two independent Vti1 genes is present in *M. brevicollis* as well, although our analysis is not able to determine whether these genes correspond to the split of Vti1a and Vti1b found in metazoans (Suppl. Data) or if these two genes represent an independent duplication.

In addition to the enlargement of the endosomal SNARE set, some other (non-endosomal) SNAREs have been duplicated in this first expansion phase. Instead of the single secretory Qbc-SNARE (Qbc.IV) SNAP-25 discovered in *M. brevicollis*, the metazoan ancestor probably possessed three different Qbc-SNARE genes; SNAP-25 (Oyler et al. 1989), SNAP-29 (Steegmaier et al. 1998; Wong et al. 1999) and SNAP-47 (Holt et al. 2006) (Fig. 1b). In contrast to SNAP-25, both SNAP-29 and SNAP-47 do not own a stretch of cysteines in the center of the linker between the two SNARE motifs. In the phylogenetic tree (Fig. 2c), SNAP-29 and SNAP-47 are well separated from SNAP-25. As with several of the other novel SNAREs mentioned above, the exact function of SNAP-29 and SNAP-47 is not entirely clear. Both SNAP-29 and SNAP-47 show a widespread distribution on intracellular organelles. SNAP-29 is thought to be involved in trafficking within the Golgi apparatus, but has also been suggested to mediate constitutive secretion, whereas the function of SNAP-47 is not known so far.

Besides the bona-fide SNARE factors, i.e. proteins that posses a functional SNARE motif, we found two additional factors that are clearly derived from SNARE genes (Fig. 1b and Fig. 3a). One is Sec22-like (also referred to as Sec22a (Hay et al. 1996)), which is derived from the R.I-SNARE Sec22 (Sec22b) that functions in trafficking between the ER and the Golgi apparatus (Hay et al. 1997; Liu and Barlowe 2002). Sec22-like carries a profilin-like N-terminal domain closely related to Sec22, but instead of a C-terminal SNARE motif, it carries three consecutive C-terminal transmembrane domains. Sec22-like is localized to the ER, but its function is unknown (Hay et al. 1996; Tang et al. 1998b).

The second novel factor is Lethal giant larvae (Lgl), which has a function in establishing epithelial cell polarity in multicellular animals (Vasioukhin 2006; Wirtz-Peitz and Knoblich 2006). Lgl is derived from tomosyn (Pobbati et al. 2004; Fasshauer and Jahn 2007), a regulatory R-SNARE (R.Reg) that is thought to be involved in polarized secretion (Fig. 1b). As tomosyn does not contain a transmembrane anchor at the end of its R-SNARE motif, it cannot serve as a fusogenic R-SNARE; instead, it is thought to regulate the accessibility of the SNARE-acceptor complex for the R-SNARE residing on secretory vesicles. Tomosyn carries a large N-terminal domain composed of two consecutive seven-bladed β-propeller domains (Hattendorf et al. 2007). Lgl possesses a very similar N-terminal domain, but it has lost the C-terminal R-SNARE motif.

Furthermore, animals, like several other eukaryotic lineages, for example

fungi and plants, usually possess two different Qc.II-SNAREs, which are thought to participate in consecutive trafficking steps within the Golgi apparatus. Nevertheless, it remains unclear, whether a duplication of an original Qc.II-SNARE occurred only once in the eukaryotic ancestor or whether the pair of Qc.II-SNAREs in animals, Bet1 (Hay et al. 1996) and Gs15 (Xu et al. 1997), arose by an independent duplication.

Losses and gains of individual SNARE genes in different animal-lineages

The increased set of 29 to 30 different SNARE proteins in urmetazoans discovered in this study, has been modified differently in different lineages of animals. We discovered nested duplications, particularly in vertebrates, and probably independent losses of individual factors in different lineages of invertebrates. Typically, gene losses affected the novel metazoan SNAREs rather than the basic repertoire. A schematic overview of the observed lineage-specific changes in the SNARE repertoire is given in Suppl. Table 2.

The genes SNAP-47 and Sec22-like were often lost, for example in nematodes, platyhelminths, insects, and crustaceans, and also in tunicates. Furthermore, we were unable to detect the Qc.II SNARE Gs15 in nematodes and insects. Loss of this gene must have occurred independently in both lineages, which are often grouped together as ecdysozoans, as we found Gs15 to be present in crustaceans and arachnidans, both of which are grouped with insects to the arthropods.

Both nematodes and platyhelminths have lost, probably independently, the endosomal SNAREs Vti1b (Qb.III.b) and Syx8 (Qc.III), of which the latter belongs to the basic SNARE set. Both proteins are thought to cooperate in one SNARE unit during transport from early towards late endosomes. Interestingly, it has been observed that mice deficient in Vti1b were viable and fertile, but had reduced amounts of Syx8, whereas the quantity of the other putative complex components, endobrevin and Syx7, was not changed (Atlashkin et al. 2003). This suggests that Syx8 and Vti1b are functionally tightly linked. Yet upon first inspection, we noted that the fruitfly *Drosophila melanogaster* contained Syx8 but apparently only one Vti1 gene. However, as we observed that several other insects, e.g. *Anopheles gambiae*, possess two different Vti1 genes, we re-inspected the genome of *D. melanogaster* and discovered a second, so far not annotated Vti1b gene on chromosome 3R. In addition, we came across a loss of the R.III-SNARE Vamp4 within holometabolous insects: so far, we have discovered Vamp4 only in Orthoptera (grasshoppers) and Hymenoptera (bees, wasps and ants), but not in Coleoptera (beetles), Lepidoptera (butterflies and moths), or Diptera (flies). In nematodes as well as in platyhelminthes, only one copy of R.III-SNAREs was found, suggesting that they have secondarily lost the novel types. However, the assignment of their R.III-SNAREs to the three metazoan types, Vamp7, Vamp7-like/endobrevin and Vamp4, is challenging, because the sequences are rather deviated, something that is frequently observed for SNARE sequences from these two phyla.

Besides several gene losses in “invertebrates” discussed above, we also came across several nested duplications within these lineages, mostly of secretory SNAREs. For example, the fruitfly possesses a duplicated set of SNAREs involved in regulated secretion, the Qbc-SNAREs SNAP-24 and SNAP-25 (Niemeyer and Schwarz 2000; Vilinsky et al. 2002), and the R-SNAREs syb (Sudhof et al. 1989; Chin et al. 1993) and nsyb (DiAntonio et al. 1993). Furthermore, we noticed that even the simple animal *Trichoplax adhaerens* possesses an enlarged set of secretory SNAREs.

Two rounds of whole-genome duplications gave rise to the enlarged set of secretory SNAREs in vertebrates

A much more pronounced enlargement by nested duplications of the SNARE repertoire occurred in the vertebrate lineage: several SNARE genes can be found in multiple copies in vertebrates, while invertebrates and basal deuterostomes typically possess only one orthologue. This observation is consistent with two consecutive rounds of whole genome duplications that are believed to have occurred early in vertebrate evolution (Ohno 1970; Holland 2003; Panopoulou and Poustka 2005).

Notably, the gene duplications in vertebrates mostly affected the secretory SNAREs. For example, duplication of the SNAP-25 gene gave rise to a second, closely related Qbc-SNARE, SNAP-23 (Ravichandran et al. 1996; Wang et al. 1997). It is possible that the alternative splicing of SNAP-25, caused by duplication of exon 5 (Bark and Wilson 1994), into two different gene products, SNAP-25a and b, also occurred in the vertebrate lineage. Even more extensive were the duplications of the secretory syntaxin (Qa.IV) and synaptobrevin (R.IV) genes, yielding a multiplicity of new gene products (Fig. 1b). The large set of secretory syntaxins (Qa.IV) in vertebrates renders the phylogenetic tree more complex (Fig. 2d). The changes in the secretory apparatus also involved the regulatory R-SNARE tomosyn. Vertebrates usually contain two different tomosyn genes (Groffen et al. 2005) and the related factor amisyn (Scales et al. 2002), which, like tomosyn, does not possess a TMR. In contrast to tomosyn, however, amisyn carries a shorter N-terminal region that is homologous to the N-terminal region of the exocyst component Sec3 (Scales et al. 2002). Moreover, the tomosyn-related gene Lgl was duplicated in vertebrates. In addition, we unearthed a few more duplications outside of the secretory set. For example, a duplication of an ancestral Qc.III gene gave rise to Syx6 (Bock et al. 1996) and Syx10 (Tang et al. 1998c) in vertebrates. In addition, an ancestral Qa.III.b SNARE was duplicated, giving rise to Syx7 and Syx13 in vertebrates. These two Qa.III.b SNAREs are thought to mediate distinct endosomal trafficking steps (Antonin et al. 2000; Collins et al. 2002; Zwilling et al. 2007). Furthermore, the duplication yielded two Sec22-like genes in vertebrates, which had originally been termed Sec22a (Hay et al. 1996) and Sec22c (Tang et al. 1998b).

A third round of whole genome duplication in bony fishes

The second genome duplication in the vertebrate lineage probably occurred before the divergence of chondrichthyes (i.e. cartilaginous fishes). This is supported by EST sequences from *Leucoraja erinacea* and *Squalus acanthias*. The position of more basic vertebrates, like the sea lamprey *Petromyzon marinus,* in relation to the two rounds of genome duplication is not entirely clear. Up to now, only few EST sequences of SNARE proteins from the sea lamprey have been available. For example, we found a fragmentary sequence that is closely related to the secretory syntaxins Syx11 and Syx19. Both factors are closely related and probably arose in a second round of genome duplication. This suggests that at least one whole genome duplication must have taken place in this organism.

Furthermore, our data support the notion that a third whole genome duplication (3R) took place in the lineage of teleost fishes (Meyer and Van de Peer 2005); (Panopoulou and Poustka 2005), since we found several duplications of secretory SNAREs in teleosts. For example, the aforementioned Syx11 gene was duplicated. In addition, several factors involved in intracellular trafficking steps were affected. For example, two independent sequences for Sec20, Sec22, Gos28, endobrevin, and possibly also Syx5 were found (Suppl. Table 2). It should be noted,

however, that this assessment is preliminary, as the genome assemblies of fishes often appear to be partly fragmentary.

Discussion

The transition to multicellularity was a pivotal step during the evolution of animals from their protozoan ancestors. Staying together as a multicellular organism brought about radical changes in lifestyle and cellular organisation, yet the basic cellular functions needed to be preserved. The emergence of multicellularity also launched the development of different cell types. The way for multicellular traits like intercellular cohesion, communication, and differentiation was paved by molecular innovations. Indeed, several domain architectures found in proteins involved in these processes are largely unique to animals (Brooke and Holland 2003; King 2004).

Phylogenetic studies of the animal kingdom often focus on protein families involved in processes related to multicellularity. For example, the evolutionary history of animal-specific homeobox genes, particularly of the collection of Hox and ParaHox genes, which are transcription factors involved cell differentiation in animals, have been intensively studied (Garcia-Fernandez 2005; Lemons and McGinnis 2006). These genes participate in the developmental program that shapes animal morphology. Obviously, the body form is acted upon strongly by natural selection, whereas it is less clear whether vital cellular functions like vesicular trafficking are also affected in multicellular organisms.

In fact, the family of SNARE proteins has diversified only modestly in the course of eukaryotic evolution (Kloepper et al. 2007). Probably, the proto-eukaryotic cell contained about twenty different SNARE proteins to catalyze the vesicular trafficking steps between the main compartments found in contemporary cells (Kloepper et al. 2007). We have now found this type of basic repertoire of SNARE genes in several eukaryotic organisms including the choanoflagellate *M. brevicollis*, an organism that is thought to be closely related to the unicellular ancestor of animals (Brooke and Holland 2003; King 2004). In fact, it has been shown that the precursors of several of the molecular building blocks that were largely exploited during animal evolution already existed in choanoflagellates (King and Carroll 2001; King et al. 2003; Segawa et al. 2006; Abedin and King 2008; King et al. 2008; Ruiz-Trillo et al. 2008). In line with this notion, many of the SNAREs of *M. brevicollis* are closely related to the ones from metazoans. In particular, the secretory set of SNAREs (type IV) from *M. brevicollis* closely resembles the set of SNAREs from multicellular animals, i.e. Syx1, SNAP-25 and Syb1. This set of SNAREs, which is commonly attributed to regulated secretion, is highly conserved in all animals.

Remarkably, we uncovered a major increase in the set of SNARE proteins that probably occurred upon emergence of multicellularity in animals, i.e. before the split of basal metazoans and bilaterians. This expansion alludes to a major overhaul of the endomembrane system during this critical period in metazoan evolution. The expanded, primordial metazoan set is largely conserved in most animal phyla. In particular, placozoans, cnidarians, annelids, molluscs, and lower deuterostomes usually have only marginally modified SNARE repertoires. Moreover, the SNAREs from these phyla often form relatively short branches in our phylogenetic trees, reflecting their basal character. This can even be seen for the SNAREs from the faster evolving urochordates (Dehal et al. 2002), which, together with cephalochordates and vertebrates, make up the chordates.

The nature of the first expansion of the metazoan SNARE set alludes to a larger genome change, possibly whole genome duplication(s), at the base of animal evolution. Although the function of SNARE proteins does not allow for larger domain re-arrangements, we came across a few of these examples during the first phase of

expansion: Syx17, Lgl, Sec22-like and Vamp4. Some of these re-arrangements, in particular the acquisition of additional transmembrane domains in Syx17 and Sec22-like, are clear evolutionary novelties that corroborate the monophyletic origin of all animals inspected. However, it should be borne in mind that although the species used for our analysis provided a significant coverage of most main extant metazoan lineages, at the time, only limited sequence information for the basal metazoans, the phyla Porifera and Ctenophora, are available. Nevertheless, the few SNARE sequences collected from poriferans suggest that the first phase of expansion of the SNARE repertoire occurred already before the split of the poriferans.

In contrast to previous claims that animals generally possess an enlarged set of secretory SNARE proteins (Dacks and Doolittle 2004; Yoshizawa et al. 2006; Dacks and Field 2007), our more detailed analysis now demonstrates that most novel SNAREs that arose during the first expansion originated from SNAREs involved in endosomal trafficking, whereas the major expansion of secretory SNAREs only occurred later, during the rise of vertebrates. But why in particular did the number of endosomal SNAREs increase during the emergence of multicellularity in animals? Although unicellular eukaryotes already have the ability to segregate different types of secretory proteins and to deliver them to different subdivisions of the plasma membrane (Soldati and Schliwa 2006), our analysis suggests that this capability has been largely amplified in animals, where the majority of cells are polarized, facing two different environments. It is possible that these morphological changes provided a fertile soil for adaptations of the endosomal SNARE repertoire. The differentiation of cells is initiated by genetic programs under the control of secreted morphogenic cues. Indeed, in the last few years, it has become clear that the endocytic pathways play important roles not only in secretion of these cues, but also in modulation of their signal transduction in the receiving cell (Gonzalez-Gaitan 2003; Rodriguez-Boulan et al. 2005; Emery and Knoblich 2006; Leibfried and Bellaiche 2007). In fact, endosomal SNARE proteins have been shown to be important for establishing cell polarity (Lu and Bilder 2005; Balklava et al. 2007). For example, the loss of the early endosomal SNARE protein Syx7 (in *Drosophila,* also referred to as avalanche) strongly resembles the phenotypes of *Drosophila* neoplastic tumor-suppressor mutations such as *scrib, dlg,* and *Lgl* (Hariharan and Bilder 2006; Humbert et al. 2006; Vasioukhin 2006; Wirtz-Peitz and Knoblich 2006; Chia et al. 2008). Intriguingly, one of these tumor-suppressor genes involved in establishment of cell polarity, Lgl, also emerged during the transition to multicellularity in animals by duplication of the regulatory SNARE tomosyn. Like tomosyn, Lgl has been shown to interact with the secretory SNARE apparatus, suggesting that both factors still function via homologous molecular pathways (Brennwald and Rossi 2007).

The road map of the endosomal compartments is rather elaborate as they serve as a sorting hub for cargo arriving from the exocytic and the endocytic pathway. It is conceivable that the increased collection of SNAREs resulted in a more complex network of endosomal trafficking routes in animals compared to their unicellular progenitors. However, it should be kept in mind that the exact role of several of the novel metazoan SNAREs is not entirely clear yet. In addition, the endosomal R-SNAREs seem not to be confined to one transport step only. For instance, Vamp7 (Ti-Vamp) not only mediates trafficking towards the lysosomal compartment, but it is also involved in secretion of lysosome-related organelles (Luzio et al. 2007). Comparably, endobrevin mediates trafficking towards late endosomes, but is also involved in granule secretion in exocrine tissues (Wang et al. 2004; Ren et al. 2007; Wang et al. 2007). It needs to be clarified in the future whether the participation of

R.III-type SNAREs in lysosomal secretion represents a new adaptation of this proteins in animals. However, considering that several other eukaryotic lineages do not possess a specific secretory R-SNARE (R.IV-type) but use R.III-SNAREs for secretion, it is likely that this ability of R.III-type of SNAREs rather reflects an ancient function of this group of proteins. Furthermore, the novel Qbc-SNAREs SNAP-29, and SNAP-47, are "exocytotic" by origin. However, their localization on intracellular membranes might also be consistent with functions in intracellular trafficking. Altogether, systematic cell biological investigations are necessary to test our predictions.

The second major expansion of the metazoan SNARE repertoire is restricted to the lineage of vertebrates, leading to their multifaceted set of secretory SNARE proteins. This expansion in vertebrates is accompanied by few domain rearrangements, but no losses of the original metazoan set were observed. Hence, our finding is in line with an earlier notion that the gene content of lower metazoans and vertebrates is highly conserved (Kortschak et al. 2003; Kusserow et al. 2005; Raible et al. 2005; Technau et al. 2005; Putnam et al. 2007; Miller and Ball 2008). This expansion of the SNARE repertoire in vertebrates can probably be attributed to two rounds of whole genome duplications (Ohno 1970; Holland 2003; Panopoulou and Poustka 2005). In line with observations on other factors, our data suggest that many other duplicated SNARE genes in vertebrates were not retained after whole genome duplications. This phenomenon is usually explained by extensive gene loss of duplicates (Holland 2003; Panopoulou and Poustka 2005). The fact that mostly duplicates of the secretory SNARE were retained, suggests that these factors presented a selective advantage during the rise of vertebrates. Often the novel secretory SNAREs are largely confined to different subcellular locations like the apical or basolateral membrane, or their expression is restricted to particular tissues or different developmental stages (for a review, see (Hong 2005; Jahn and Scheller 2006; Stow et al. 2006)). Hence, the extended set of secretory SNAREs appears to have facilitated the development of new secretory cell types in vertebrates.

It is very likely that not only the major expansions but also the losses of individual SNARE factors in various animal lineages affected the itinerary of transport vesicles as the changes probably reflect lineage-specific adaptations of the endomembrane system. Hence, one should be cautious when comparing the vesicular trafficking pathways of different animal species. Losses occurred for example in the invertebrates *D. melanogaster* and *C. elegans*, which are well-studied model organisms. Particularly in the repertoire of nematodes and platyhelminthes, we noticed the absence of some novel metazoan SNARE types involved in endosomal trafficking. As both phyla possess a somewhat differently reduced set of novel SNAREs, these losses seem to reflect secondary gene losses, rather than a more basal status of these phyla. Hence, it seems possible that comparable losses of several endosomal SNARE genes in nematodes and platyhelminthes are a sign of convergent adaptations of their endosomal sorting pathways caused by comparable life-style and feeding requirements. In addition, and in line with observations on other genes, the SNARE sequences from both phyla often form long but well-separated branches, supporting their more derived status. In fact, the relative position of these phyla in the phylogeny of animals is still debated (for discussion see e.g. (Philippe and Telford 2006)).

It is worth mentioning that we also observed a few differences in the SNARE repertoire of different vertebrate lineages. For example, the Qc.III genes Syx6 and Syx10, which are the products of a gene-duplication in vertebrates, participate in

different SNARE units that mediate trafficking of different types of vesicles from endosomes towards the TGN. Remarkably, Syx10 is secondarily absent from mouse and rat genomes, suggesting that a subtle difference in the endosomal trafficking pathways between murine rodents and other vertebrates exists (Ganley et al. 2008). The existence of differences in their endosomal pathways is corroborated by the fact that Syx20 (Qa.III.b), an endosomal syntaxin with unknown function, appears to have been lost in mice and rats as well.

Overall, the animal genomes examined in this work comprise a remarkably consistent and homologous set of SNARE proteins. As outlined above, the two major expansions can probably be attributed to two distinct periods of larger genome remodeling, during which several individual SNARE factors were duplicated. Yet, duplications of SNAREs, in particular of the ones involved in secretion, followed by functional diversification, occurred recurrently in different lineages. For example, we came across duplicated sets of secretory SNAREs in several invertebrates but also in another choanoflagellate, *Monosiga ovata*. The identification of lineage-specific duplications was possible because we have integrated sequence data from a large quantity of species that represent a very broad spectrum of the animal kingdom. In addition, we rigorously removed duplicates from our dataset. Furthermore, we have used a highly accurate classification and a detailed phylogenetic analysis. We noted that these methodological key elements were paid less attention to in earlier studies, explaining why these studies were unable to distinguish the two major expansions of the metazoan SNARE repertoire found by us (Dacks and Doolittle 2004; Yoshizawa et al. 2006; Dacks and Field 2007).

Although it was not the primary goal of this study, we noticed that the trees built from orthologous SNAREs generally recapitulate their presumed evolutionary relationships. In fact, the sequences of species that belong to the same phylum usually congregated. This is seen best in the trees of SNAREs involved in trafficking within the ER and the Golgi, since these factors are present usually as singletons in all animal genomes. These factors apparently underwent slow lineage-specific adaptations (Fig. 3). More drastic changes occurred in SNARE proteins involved in endosomal and secretory trafficking steps, probably, as discussed above, evoked by episodes of larger genome re-arrangements. For example, expansion of the set of secretory syntaxins (Qa.IV) in vertebrates led to a highly diversified set that obscures the underlying information of the phylogenetic relationships (Fig. 2d). Generally, the trees based on single sequences are not ideally suited to reflect the relationships between the larger phylogenetic groups of animals. In fact, the paralogous expansions render it challenging to assemble and align concatenated sequences to increase the resolution of the phylogenetic tree. Also, several animal lineages are still under-represented in the genome sequences that are currently available. Nevertheless, as new genome data can be quickly integrated, our detailed analysis of the evolutionary history of SNARE genes in animals is a promising supplementary tool for future phylogenetic studies.

Supplementary Material

The supplementary data of this study can be downloaded from the SNARE projects homepage (http://bioinformatics.mpibpc.mpg.de/snare/). The SNARE projects homepage allows access to the SNARE database (Kloepper et al. 2007), into which all newly classified sequences of this study have been integrated.

Acknowledgements

We are greatly indebted to D. H. Huson for generously supporting this project. In addition, we thank F. Raible for kindly providing us with unpublished EST sequences of *Platynereis dumerilii*. We also want to thank the various genome research institutes for making available the sequenced genomes of the different species used here. Moreover, we thank U. Winter for critical reading of the manuscript.

Literature Cited

Abedin, M., and N. King. 2008. The Premetazoan Ancestry of Cadherins. Science **319**:946-948.

Advani, R. J., H. R. Bae, J. B. Bock, D. S. Chao, Y. C. Doung, R. Prekeris, J. S. Yoo, and R. H. Scheller. 1998. Seven novel mammalian SNARE proteins localize to distinct membrane compartments. J Biol Chem **273**:10317-10324.

Advani, R. J., B. Yang, R. Prekeris, K. C. Lee, J. Klumperman, and R. H. Scheller. 1999. VAMP-7 mediates vesicular transport from endosomes to lysosomes. J Cell Biol **146**:765-776.

Antonin, W., C. Holroyd, D. Fasshauer, S. Pabst, G. F. Von Mollard, and R. Jahn. 2000. A SNARE complex mediating fusion of late endosomes defines conserved properties of SNARE structure and function. EMBO J. **19**:6453-6464.

Atlashkin, V., V. Kreykenbohm, E. L. Eskelinen, D. Wenzel, A. Fayyazi, and G. Fischer von Mollard. 2003. Deletion of the SNARE vti1b in mice results in the loss of a single SNARE partner, syntaxin 8. Mol Cell Biol **23**:5198-5207.

Balklava, Z., S. Pant, H. Fares, and B. D. Grant. 2007. Genome-wide analysis identifies a general requirement for polarity proteins in endocytic traffic. Nat Cell Biol **9**:1066-1073.

Bark, I. C., and M. C. Wilson. 1994. Human cDNA clones encoding two different isoforms of the nerve terminal protein SNAP-25. Gene **139**:291-292.

Behnia, R., and S. Munro. 2005. Organelle identity and the signposts for membrane traffic. Nature **438**:597-604.

Bock, J. B., R. C. Lin, and R. H. Scheller. 1996. A new syntaxin family member implicated in targeting of intracellular transport vesicles. J Biol Chem **271**:17961-17965.

Bock, J. B., H. T. Matern, A. A. Peden, and R. H. Scheller. 2001. A genomic perspective on membrane compartment organization. Nature **409**:839-841.

Bonifacino, J. S., and B. S. Glick. 2004. The mechanisms of vesicle budding and fusion. Cell **116**:153-166.

Brennwald, P., and G. Rossi. 2007. Spatial regulation of exocytosis and cell polarity: yeast as a model for animal cells. FEBS Lett **581**:2119-2124.

Brooke, N. M., and P. W. Holland. 2003. The evolution of multicellularity and early animal genomes. Curr Opin Genet Dev **13**:599-603.

Cai, H., K. Reinisch, and S. Ferro-Novick. 2007. Coats, tethers, Rabs, and SNAREs work together to mediate the intracellular destination of a transport vesicle. Dev Cell **12**:671-682.

Cavalier-Smith, T. 2002. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int J Syst Evol Microbiol **52**:297-354.

Chia, W., W. G. Somers, and H. Wang. 2008. Drosophila neuroblast asymmetric divisions: cell cycle regulators, asymmetric protein localization, and tumorigenesis. J Cell Biol **180**:267-272.

Chin, A. C., R. W. Burgess, B. R. Wong, T. L. Schwarz, and R. H. Scheller. 1993. Differential expression of transcripts from syb, a Drosophila melanogaster gene encoding VAMP (synaptobrevin) that is abundant in non-neuronal cells. Gene **131**:175-181.

Collins, R. F., A. D. Schreiber, S. Grinstein, and W. S. Trimble. 2002. Syntaxins 13 and 7 function at distinct steps during phagocytosis. J Immunol **169**:3250-3256.

Dacks, J. B., and W. F. Doolittle. 2004. Molecular and phylogenetic characterization of syntaxin genes from parasitic protozoa. Mol Biochem Parasitol **136**:123-136.

Dacks, J. B., and M. C. Field. 2007. Evolution of the eukaryotic membrane-trafficking system: origin, tempo and mode. J Cell Sci **120**:2977-2985.

Dehal, P., Y. Satou, R. K. Campbell, J. Chapman, B. Degnan, A. De Tomaso, B. Davidson, A. Di Gregorio, M. Gelpke, D. M. Goodstein, N. Harafuji, K. E. Hastings, I. Ho, K. Hotta, W. Huang, T. Kawashima, P. Lemaire, D. Martinez, I. A. Meinertzhagen, S. Necula, M. Nonaka, N. Putnam, S. Rash, H. Saiga, M. Satake, A. Terry, L. Yamada, H. G. Wang, S. Awazu, K. Azumi, J. Boore, M. Branno, S. Chin-Bow, R. DeSantis, S. Doyle, P. Francino, D. N. Keys, S. Haga, H. Hayashi, K. Hino, K. S. Imai, K. Inaba, S. Kano, K. Kobayashi, M. Kobayashi, B. I. Lee, K. W. Makabe, C. Manohar, G. Matassi, M. Medina, Y. Mochizuki, S. Mount, T. Morishita, S. Miura, A. Nakayama, S. Nishizaka, H. Nomoto, F. Ohta, K. Oishi, I. Rigoutsos, M. Sano, A. Sasaki, Y. Sasakura, E. Shoguchi, T. Shin-i, A. Spagnuolo, D. Stainier, M. M. Suzuki, O. Tassy, N. Takatori, M. Tokuoka, K. Yagi, F. Yoshizaki, S. Wada, C. Zhang, P. D. Hyatt, F. Larimer, C. Detter, N. Doggett, T. Glavina, T. Hawkins, P. Richardson, S. Lucas, Y. Kohara, M. Levine, N. Satoh, and D. S. Rokhsar. 2002. The draft genome of Ciona intestinalis: insights into chordate and vertebrate origins. Science **298**:2157-2167.

Dellaporta, S. L., A. Xu, S. Sagasser, W. Jakob, M. A. Moreno, L. W. Buss, and B. Schierwater. 2006. Mitochondrial genome of Trichoplax adhaerens supports placozoa as the basal lower metazoan phylum. Proc Natl Acad Sci U S A **103**:8751-8756.

DiAntonio, A., R. W. Burgess, A. C. Chin, D. L. Deitcher, R. H. Scheller, and T. L. Schwarz. 1993. Identification and characterization of Drosophila genes for synaptic vesicle proteins. J. Neurosci. **13**:4924-4935.

Edgar, R. C. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics **5**:113.

Emery, G., and J. A. Knoblich. 2006. Endosome dynamics during development. Curr

Opin Cell Biol **18**:407-415.
Fasshauer, D., and R. Jahn. 2007. Budding insights on cell polarity. Nat Struct Mol Biol **14**:360-362.
Felsenstein, J. 1998. PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics:164-166.
Ganley, I. G., E. Espinosa, and S. R. Pfeffer. 2008. A syntaxin 10-SNARE complex distinguishes two distinct transport routes from endosomes to the trans-Golgi in human cells. J Cell Biol **180**:159-172.
Garcia-Fernandez, J. 2005. The genesis and evolution of homeobox gene clusters. Nat Rev Genet **6**:881-892.
Gonzalez-Gaitan, M. 2003. Signal dispersal and transduction through the endocytic pathway. Nat Rev Mol Cell Biol **4**:213-224.
Groffen, A. J., L. Jacobsen, D. Schut, and M. Verhage. 2005. Two distinct genes drive expression of seven tomosyn isoforms in the mammalian brain, sharing a conserved structure with a unique variable domain. J Neurochem **92**:554-568.
Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol **52**:696-704.
Hariharan, I. K., and D. Bilder. 2006. Regulation of imaginal disc growth by tumor-suppressor genes in Drosophila. Annu Rev Genet **40**:335-361.
Hattendorf, D. A., A. Andreeva, A. Gangar, P. J. Brennwald, and W. I. Weis. 2007. Structure of the yeast polarity protein Sro7 reveals a SNARE regulatory mechanism. Nature **446**:567-571.
Hay, J. C., D. S. Chao, C. S. Kuo, and R. H. Scheller. 1997. Protein interactions regulating vesicle transport between the endoplasmic reticulum and Golgi apparatus in mammalian cells. Cell **89**:149-158.
Hay, J. C., H. Hirling, and R. H. Scheller. 1996. Mammalian vesicle trafficking proteins of the endoplasmic reticulum and Golgi apparatus. J Biol Chem **271**:5671-5679.
Hinners, I., F. Wendler, H. Fei, L. Thomas, G. Thomas, and S. A. Tooze. 2003. AP-1 recruitment to VAMP4 is modulated by phosphorylation-dependent binding of PACS-1. EMBO Rep **4**:1182-1189.
Holland, P. W. 2003. More genes in vertebrates? J Struct Funct Genomics **3**:75-84.
Holt, M., F. Varoqueaux, K. Wiederhold, S. Takamori, H. Urlaub, D. Fasshauer, and R. Jahn. 2006. Identification of SNAP-47, a novel Qbc-SNARE with ubiquitous expression. J Biol Chem **281**:17076-17083.
Hong, W. 2005. SNAREs and traffic. Biochim Biophys Acta **1744**:493-517.
Humbert, P. O., L. E. Dow, and S. M. Russell. 2006. The Scribble and Par complexes in polarity and migration: friends or foes? Trends Cell Biol **16**:622-630.
Huson, D. H., and D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol **23**:254-267.
Jahn, R., and R. H. Scheller. 2006. SNAREs--engines for membrane fusion. Nat Rev Mol Cell Biol **7**:631-643.
King, N. 2004. The unicellular ancestry of animal development. Dev Cell **7**:313-325.
King, N., and S. B. Carroll. 2001. A receptor tyrosine kinase from choanoflagellates: molecular insights into early animal evolution. Proc Natl Acad Sci U S A **98**:15032-15037.
King, N., C. T. Hittinger, and S. B. Carroll. 2003. Evolution of key cell signaling and adhesion protein families predates animal origins. Science **301**:361-363.
King, N., M. J. Westbrook, S. L. Young, A. Kuo, M. Abedin, J. Chapman, S. Fairclough, U. Hellsten, Y. Isogai, I. Letunic, M. Marr, D. Pincus, N. Putnam,

A. Rokas, K. J. Wright, R. Zuzow, W. Dirks, M. Good, D. Goodstein, D. Lemons, W. Li, J. B. Lyons, A. Morris, S. Nichols, D. J. Richter, A. Salamov, J. G. Sequencing, P. Bork, W. A. Lim, G. Manning, W. T. Miller, W. McGinnis, H. Shapiro, R. Tjian, I. V. Grigoriev, and D. Rokhsar. 2008. The genome of the choanoflagellate Monosiga brevicollis and the origin of metazoans. Nature **451**:783-788.

Kloepper, T. H., C. N. Kienle, and D. Fasshauer. 2007. An elaborate classification of SNARE proteins sheds light on the conservation of the eukaryotic endomembrane system. Mol Biol Cell **18**:3463-3471.

Kortschak, R. D., G. Samuel, R. Saint, and D. J. Miller. 2003. EST analysis of the cnidarian Acropora millepora reveals extensive gene loss and rapid sequence divergence in the model invertebrates. Curr Biol **13**:2190-2195.

Kusserow, A., K. Pang, C. Sturm, M. Hrouda, J. Lentfer, H. A. Schmidt, U. Technau, A. von Haeseler, B. Hobmayer, M. Q. Martindale, and T. W. Holstein. 2005. Unexpected complexity of the Wnt gene family in a sea anemone. Nature **433**:156-160.

Leibfried, A., and Y. Bellaiche. 2007. Functions of endosomal trafficking in Drosophila epithelial cells. Curr Opin Cell Biol **19**:446-452.

Lemons, D., and W. McGinnis. 2006. Genomic evolution of Hox gene clusters. Science **313**:1918-1922.

Liu, Y., and C. Barlowe. 2002. Analysis of Sec22p in Endoplasmic Reticulum/Golgi Transport Reveals Cellular Redundancy in SNARE Protein Function. Mol Biol Cell **13**:3314-3324.

Lu, H., and D. Bilder. 2005. Endocytic control of epithelial polarity and proliferation in Drosophila. Nat Cell Biol **7**:1232-1239.

Luzio, J. P., P. R. Pryor, and N. A. Bright. 2007. Lysosomes: fusion and function. Nat Rev Mol Cell Biol **8**:622-632.

Marcet-Palacios, M., S. O. Odemuyiwa, J. J. Coughlin, D. Garofoli, C. Ewen, C. E. Davidson, M. Ghaffari, K. P. Kane, P. Lacy, M. R. Logan, A. D. Befus, R. C. Bleackley, and R. Moqbel. 2007. Vesicle-associated membrane protein 7 (VAMP-7) is essential for target cell killing in a natural killer cell line. Biochem Biophys Res Commun.

Martinez-Arca, S., R. Rudge, M. Vacca, G. Raposo, J. Camonis, V. Proux-Gillardeaux, L. Daviet, E. Formstecher, A. Hamburger, F. Filippini, M. D'Esposito, and T. Galli. 2003. A dual mechanism controlling the localization and function of exocytic v-SNAREs. Proc Natl Acad Sci U S A **100**:9011-9016.

Meyer, A., and Y. Van de Peer. 2005. From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). Bioessays **27**:937-945.

Miller, D. J., and E. E. Ball. 2008. Cryptic complexity captured: the Nematostella genome reveals its secrets. Trends Genet **24**:1-4.

Miller, D. J., and E. E. Ball. 2005. Animal evolution: the enigmatic phylum placozoa revisited. Curr Biol **15**:R26-28.

Niemeyer, B. A., and T. L. Schwarz. 2000. SNAP-24, a Drosophila SNAP-25 homologue on granule membranes, is a putative mediator of secretion and granule-granule fusion in salivary glands. J Cell Sci **113 (Pt 22)**:4055-4064.

Ohno, S. 1970. Evolution by Gene Duplication. Springer Verlag, New York.

Oyler, G. A., G. A. Higgins, R. A. Hart, E. Battenberg, M. Billingsley, F. E. Bloom, and M. C. Wilson. 1989. The identification of a novel synaptosomal-associated protein, SNAP-25, differentially expressed by neuronal

subpopulations. J Cell Biol **109**:3039-3052.
Panopoulou, G., and A. J. Poustka. 2005. Timing and mechanism of ancient vertebrate genome duplications -- the adventure of a hypothesis. Trends Genet **21**:559-567.
Peden, A. A., G. Y. Park, and R. H. Scheller. 2001. The Di-leucine motif of vesicle-associated membrane protein 4 is required for its localization and AP-1 binding. J Biol Chem **276**:49183-49187.
Philippe, H., and M. J. Telford. 2006. Large-scale sequencing and the new animal phylogeny. Trends Ecol Evol **21**:614-620.
Pobbati, A. V., A. Razeto, M. Boddener, S. Becker, and D. Fasshauer. 2004. Structural basis for the inhibitory role of tomosyn in exocytosis. J Biol Chem **279**:47192-47200.
Pocard, T., A. Le Bivic, T. Galli, and C. Zurzolo. 2007. Distinct v-SNAREs regulate direct and indirect apical delivery in polarized epithelial cells. J Cell Sci **120**:3309-3320.
Prekeris, R., J. Klumperman, and R. H. Scheller. 2000. Syntaxin 11 is an atypical SNARE abundant in the immune system. Eur J Cell Biol **79**:771-780.
Putnam, N. H., M. Srivastava, U. Hellsten, B. Dirks, J. Chapman, A. Salamov, A. Terry, H. Shapiro, E. Lindquist, V. V. Kapitonov, J. Jurka, G. Genikhovich, I. V. Grigoriev, S. M. Lucas, R. E. Steele, J. R. Finnerty, U. Technau, M. Q. Martindale, and D. S. Rokhsar. 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science **317**:86-94.
Raible, F., K. Tessmar-Raible, K. Osoegawa, P. Wincker, C. Jubin, G. Balavoine, D. Ferrier, V. Benes, P. de Jong, J. Weissenbach, P. Bork, and D. Arendt. 2005. Vertebrate-type intron-rich genes in the marine annelid Platynereis dumerilii. Science **310**:1325-1326.
Ravichandran, V., A. Chawla, and P. A. Roche. 1996. Identification of a novel syntaxin- and synaptobrevin/VAMP-binding protein, SNAP-23, expressed in non-neuronal tissues. J. Biol. Chem. **271**:13300-13303.
Ren, Q., H. K. Barber, G. L. Crawford, Z. A. Karim, C. Zhao, W. Choi, C. C. Wang, W. Hong, and S. W. Whiteheart. 2007. Endobrevin/VAMP-8 is the primary v-SNARE for the platelet release reaction. Mol Biol Cell **18**:24-33.
Rodriguez-Boulan, E., G. Kreitzer, and A. Musch. 2005. Organization of vesicular trafficking in epithelia. Nat Rev Mol Cell Biol **6**:233-247.
Ruiz-Trillo, I., A. J. Roger, G. Burger, M. W. Gray, and B. F. Lang. 2008. A phylogenomic investigation into the origin of metazoa. Mol Biol Evol **25**:664-672.
Sanderfoot, A. 2007. Increases in the number of SNARE genes parallels the rise of multicellularity among the green plants. Plant Physiol **144**:6-17.
Sanderfoot, A. A., F. F. Assaad, and N. V. Raikhel. 2000. The Arabidopsis genome. An abundance of soluble N-ethylmaleimide-sensitive factor adaptor protein receptors. Plant Physiol **124**:1558-1569.
Scales, S. J., B. A. Hesser, E. S. Masuda, and R. H. Scheller. 2002. Amisyn, a Novel Syntaxin-binding Protein That May Regulate SNARE Complex Assembly. J Biol Chem **277**:28271-28279.
Schierwater, B. 2005. My favorite animal, Trichoplax adhaerens. Bioessays **27**:1294-1302.
Segawa, Y., H. Suga, N. Iwabe, C. Oneyama, T. Akagi, T. Miyata, and M. Okada. 2006. Functional development of Src tyrosine kinases during evolution from a unicellular ancestor to multicellular animals. Proc Natl Acad Sci U S A

103:12021-12026.
Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst Biol **51**:492-508.
Shimodaira, H., and M. Hasegawa. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics **17**:1246-1247.
Signorovitch, A. Y., L. W. Buss, and S. L. Dellaporta. 2007. Comparative genomics of large mitochondria in placozoans. PLoS Genet **3**:e13.
Soldati, T., and M. Schliwa. 2006. Powering membrane traffic in endocytosis and recycling. Nat Rev Mol Cell Biol **7**:897-908.
Steegmaier, M., V. Oorschot, J. Klumperman, and R. H. Scheller. 2000. Syntaxin 17 is abundant in steroidogenic cells and implicated in smooth endoplasmic reticulum membrane dynamics [In Process Citation]. Mol Biol Cell **11**:2719-2731.
Steegmaier, M., B. Yang, J. S. Yoo, B. Huang, M. Shen, S. Yu, Y. Luo, and R. H. Scheller. 1998. Three novel proteins of the Syntaxin/SNAP-25 family. J Biol Chem **273**:34171-34179.
Stow, J. L., A. P. Manderson, and R. Z. Murray. 2006. SNAREing immunity: the role of SNAREs in the immune system. Nat Rev Immunol **6**:919-929.
Sudhof, T. C., M. Baumert, M. S. Perin, and R. Jahn. 1989. A synaptic vesicle membrane protein is conserved from mammals to Drosophila. Neuron **2**:1475-1481.
Tang, B. L., D. Y. Low, and W. Hong. 1998a. Syntaxin 11: a member of the syntaxin family without a carboxyl terminal transmembrane domain. Biochem Biophys Res Commun **245**:627-632.
Tang, B. L., D. Y. Low, and W. Hong. 1998b. Hsec22c: a homolog of yeast Sec22p and mammalian rsec22a and msec22b/ERS-24. Biochem Biophys Res Commun **243**:885-891.
Tang, B. L., D. Y. Low, A. E. Tan, and W. Hong. 1998c. Syntaxin 10: a member of the syntaxin family localized to the trans- Golgi network. Biochem Biophys Res Commun **242**:345-350.
Tang, B. L., A. E. Tan, L. K. Lim, S. S. Lee, D. Y. Low, and W. Hong. 1998d. Syntaxin 12, a member of the syntaxin family localized to the endosome. J Biol Chem **273**:6944-6950.
Technau, U., S. Rudd, P. Maxwell, P. M. Gordon, M. Saina, L. C. Grasso, D. C. Hayward, C. W. Sensen, R. Saint, T. W. Holstein, E. E. Ball, and D. J. Miller. 2005. Maintenance of ancestral complexity and non-metazoan genes in two basal cnidarians. Trends Genet **21**:633-639.
Tran, T. H., Q. Zeng, and W. Hong. 2007. VAMP4 cycles from the cell surface to the trans-Golgi network via sorting and recycling endosomes. J Cell Sci **120**:1028-1041.
Vasioukhin, V. 2006. Lethal giant puzzle of Lgl. Dev Neurosci **28**:13-24.
Veit, M., T. H. Soellner, and J. E. Rothman. 1996. Multiple palmitoylation of synaptotagmin and the t-SNARE SNAP-25. FEBS Lett. **385**:119-123.
Vilinsky, I., B. A. Stewart, J. Drummond, I. Robinson, and D. L. Deitcher. 2002. A Drosophila SNAP-25 null mutant reveals context-dependent redundancy with SNAP-24 in neurotransmission. Genetics **162**:259-271.
Vinh le, S., and A. Von Haeseler. 2004. IQPNNI: moving fast through tree space and stopping in time. Mol Biol Evol **21**:1565-1571.
Wang, C. C., C. P. Ng, L. Lu, V. Atlashkin, W. Zhang, L. F. Seet, and W. Hong. 2004. A role of VAMP8/endobrevin in regulated exocytosis of pancreatic

acinar cells. Dev Cell **7**:359-371.

Wang, C. C., H. Shi, K. Guo, C. P. Ng, J. Li, B. Q. Gan, H. Chien Liew, J. Leinonen, H. Rajaniemi, Z. H. Zhou, Q. Zeng, and W. Hong. 2007. VAMP8/endobrevin as a general vesicular SNARE for regulated exocytosis of the exocrine system. Mol Biol Cell **18**:1056-1063.

Wang, G., J. W. Witkin, G. Hao, V. A. Bankaitis, P. E. Scherer, and G. Baldini. 1997. Syndet is a novel SNAP-25 related protein expressed in many tissues. J Cell Sci **110 (Pt 4)**:505-513.

Wang, Y., L. Y. Foo, K. Guo, B. Q. Gan, Q. Zeng, W. Hong, and B. L. Tang. 2006. Syntaxin 9 is enriched in skin hair follicle epithelium and interacts with the epidermal growth factor receptor. Traffic **7**:216-226.

Wirtz-Peitz, F., and J. A. Knoblich. 2006. Lethal giant larvae take on a life of their own. Trends Cell Biol **16**:234-241.

Wong, S. H., Y. Xu, T. Zhang, G. Griffiths, S. L. Lowe, V. N. Subramaniam, K. T. Seow, and W. Hong. 1999. GS32, a novel Golgi SNARE of 32 kDa, interacts preferentially with syntaxin 6. Mol Biol Cell **10**:119-134.

Wong, S. H., Y. Xu, T. Zhang, and W. Hong. 1998a. Syntaxin 7, a novel syntaxin member associated with the early endosomal compartment. J Biol Chem **273**:375-380.

Wong, S. H., T. Zhang, Y. Xu, V. N. Subramaniam, G. Griffiths, and W. Hong. 1998b. Endobrevin, a novel synaptobrevin/VAMP-like protein preferentially associated with the early endosome. Mol Biol Cell **9**:1549-1563.

Xu, Y., S. H. Wong, B. L. Tang, V. N. Subramaniam, T. Zhang, and W. Hong. 1998. A 29-kilodalton Golgi soluble N-ethylmaleimide-sensitive factor attachment protein receptor (Vti1-rp2) implicated in protein trafficking in the secretory pathway. J Biol Chem **273**:21783-21789.

Xu, Y., S. H. Wong, T. Zhang, V. N. Subramaniam, and W. Hong. 1997. GS15, a 15-kilodalton Golgi soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) homologous to rbet1. J Biol Chem **272**:20162-20166.

Yoshizawa, A. C., S. Kawashima, S. Okuda, M. Fujita, M. Itoh, Y. Moriya, M. Hattori, and M. Kanehisa. 2006. Extracting sequence motifs and the phylogenetic features of SNARE-dependent membrane traffic. Traffic **7**:1104-1118.

Zeng, Q., T. T. Tran, H. X. Tan, and W. Hong. 2003. The cytoplasmic domain of Vamp4 and Vamp5 is responsible for their correct subcellular targeting: the N-terminal extenSion of VAMP4 contains a dominant autonomous targeting signal for the trans-Golgi network. J Biol Chem **278**:23046-23054.

Zhang, Q., J. Li, M. Deavers, J. L. Abbruzzese, and L. Ho. 2005. The subcellular localization of syntaxin 17 varies among different cell types and is altered in some malignant cells. J Histochem Cytochem **53**:1371-1382.

Zwilling, D., A. Cypionka, W. H. Pohl, D. Fasshauer, P. J. Walla, M. C. Wahl, and R. Jahn. 2007. Early endosomal SNAREs form a structurally conserved SNARE complex and fuse liposomes with multiple topologies. Embo J **26**:9-18.

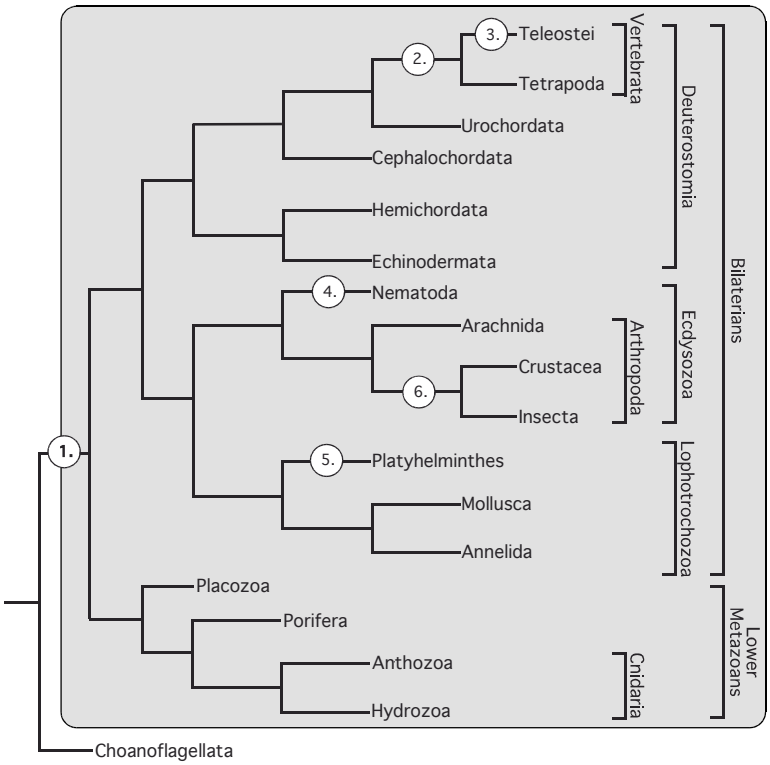
Figure Legends

Fig. 1. Schematic depiction of the major changes in the SNARE repertoire during animal evolution.
a) Phylogenetic relationships between the major animal groups represented in this study. The tree is based on the current view of the phylogeny of metazoans, but several diversifications are still heavily debated (Philippe and Telford 2006). For example, the phylogenetic placement of placozoan *Trichoplax adhaerens* among the lower metazoans is not clear (Miller and Ball 2005; Schierwater 2005), although recent data strongly suggest that their position should be as the most basal extant lower metazoan (Dellaporta et al. 2006; Signorovitch et al. 2007). The numbers indicate the major changes in the SNARE repertoire uncovered in this study. The details of the changes indicated by the numbers 1 and 2 are illustrated in b); the other lineage-specific changes (3 through 7) are displayed in Suppl. Table. 2.
b) Schematic depiction of the major changes in the SNARE repertoire during the rise of multicellular animals and of vertebrates.
To illustrate the changes of the SNARE sets representative species with well-maintained genomes were used: *Monosiga brevicollis* for choanoflagellates, the cnidarian *Nematostella vectensis* for basal metazoans, and *Mus musculus* for vertebrates/tetrapods. The domain architecture of the SNARE proteins is drawn to scale. The highly conserved SNARE motif is depicted by a gray box; the adjacent transmembrane region by a black box. The degenerated SNARE motif of Lgl is depicted as a white box. The two β-propeller domains of tomosyn and Lgl are indicated as white ovals. Note that the C-terminal region of Syx6 from *M. brevicollis* lacks a transmembrane region and is shown as a dashed line to indicate that this region of the sequence most likely is not correctly annotated. The expansion of the SNARE set in vertebrates is probably caused by two rounds of whole genome duplications. Consequently, tetrapods can have up to four orthologous copies – although two or three copies are more often seen. Note that an enlarged set of regulatory R-SNAREs in vertebrates is not shown for space limitations. Notably, one has to assume that at least two Qa.IV-genes existed in the initial repertoire of the vertebrate ancestor to be able to account for the numerous novel Qa.IV genes in vertebrates, e.g. Syx1-4, Syx11 (Tang et al. 1998a; Prekeris et al. 2000), Syx19 (Wang et al. 2006), and Syx21. To account for this, the putative second Qa.IV-SNARE of the vertebrate ancestor is shown in gray. Interestingly, in few animal species (e.g. *Nematostella vectensis*), we found only a single set of secretory Qa- and R-SNAREs. This is comparable to many unicellular eukaryotes, for which it is thought that the single secretory set generally is capable of mediating constitutive and regulated secretion. Possibly, discrimination between both types of secretion can be imposed on this SNARE unit by additional factors. It is thus conceivable that even in multicellular organisms on the level of the SNARE factors, no strict separation between constitutive and regulated secretion exists.

Fig. 2. Outlines of unrooted phylogenetic trees of individual SNARE subgroups that have undergone major expansions in the evolution of animals.
Each tree is shown as a schematic outline. For each tree, a detailed figure is given at the SNARE projects homepage (http://bioinformatics.mpibpc.mpg.de/snare/). In addition, all individual phylogenetic trees, and the corresponding sequence alignments are available there. In each tree, the diverged metazoan SNARE types are shown by different colors. In addition, the major groups of animals are indicated. The labels at the major branches represent the likelihood mapping (left) and AU support values (right).
a) The endosomal R-SNAREs (R.III-type) of animals split into three major branches Vamp7, Vamp4 and endobrevin. The gene, Vamp7-like, which is only present in a few basal metazoans, is not especially indicated in the schematic tree, but can be seen in the enlarged version.
b) The endosomal Qa-SNAREs (Qa.III.b-type) of animals split into the three elementary branches, Syx7, Syx20, and Syx17. In the lineage of vertebrates, another type, Syx13, arose.
c) The Qbc-SNAREs (Qbc.IV-type) of animals basically split into three branches, SNAP-25, SNAP-29 and SNAP-47. Note that SNAP-47 contains an extended linker region compared to the other metazoan Qbc-SNAREs. In the lineage of vertebrates, another type, SNAP-23, emerged, but independent duplications of the SNAP-25-type occurred in other lineages as well.
d) The phylogenetic tree of secretory syntaxins (Qa.IV-type) is dominated by the major expansion in the vertebrate lineage that gave rise to various new types. For example, both Syx11 and Syx19 belong to a more deviated vertebrate specific group of secretory syntaxins that, instead of a transmembrane region (TMR), possess a C-terminal cysteine rich sequence, which is probably palmitoylated.
Originally, metazoans appear to have contained only one or two isoforms of secretory syntaxins. For example, besides the highly preserved Syx1, several lineages possess a much more diverged secretory syntaxin (e.g. Syx4 of insects).

Fig. 3. The phylogenetic trees of individual SNAREs reflect the relationship of animal groups.
In metazoans, a new factor, Sec22-like, arose by gene duplication of the R-SNARE Sec22 (R.I-type) followed by domain rearrangements. Sec22-like has lost its R-SNARE motif, but possess, three consequtive TMRs. **(a)** The schematic overview of the unrooted phylogenetic tree of metazoan Sec22 subfamily shows that the two factors are well separated. Sec22-like appears to have been lost again in different animal lineages, whereas Sec22, besides a duplication in teleost fishes, has been retained as a singleton. The phylogenetic tree of Sec22 reflects the evolutionary relationships among the larger animal groups well. For example, Sec22 from fishes lies on the vertebrate branch, but is separated from tetrapod sequences. Comparably, Sec22 sequences from insects are well confined, despite the enormous morphological diversity of this animal group. **(b).** A close-up view of this branch shows that various insect groups are also well separated.

A



B

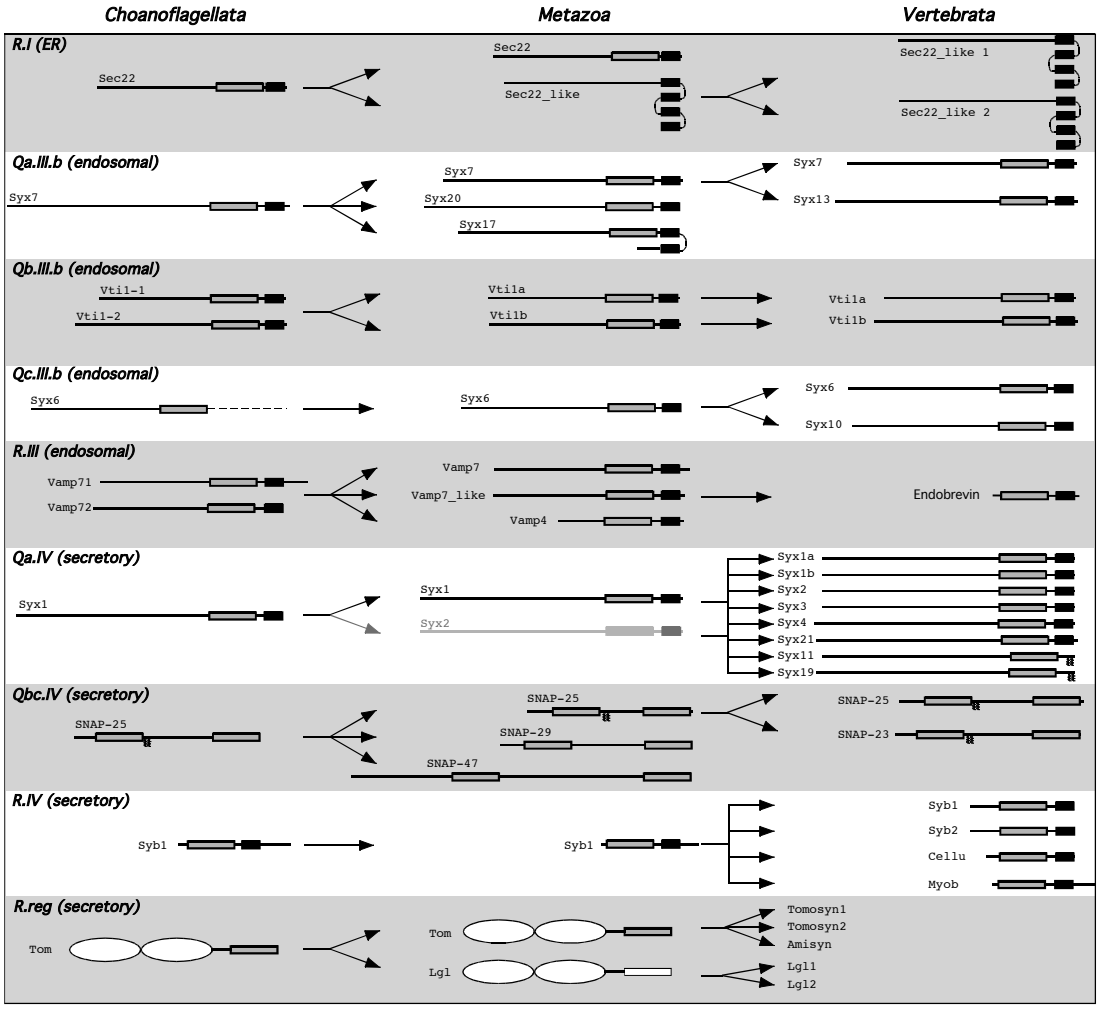


Fig1- Kloepper et al

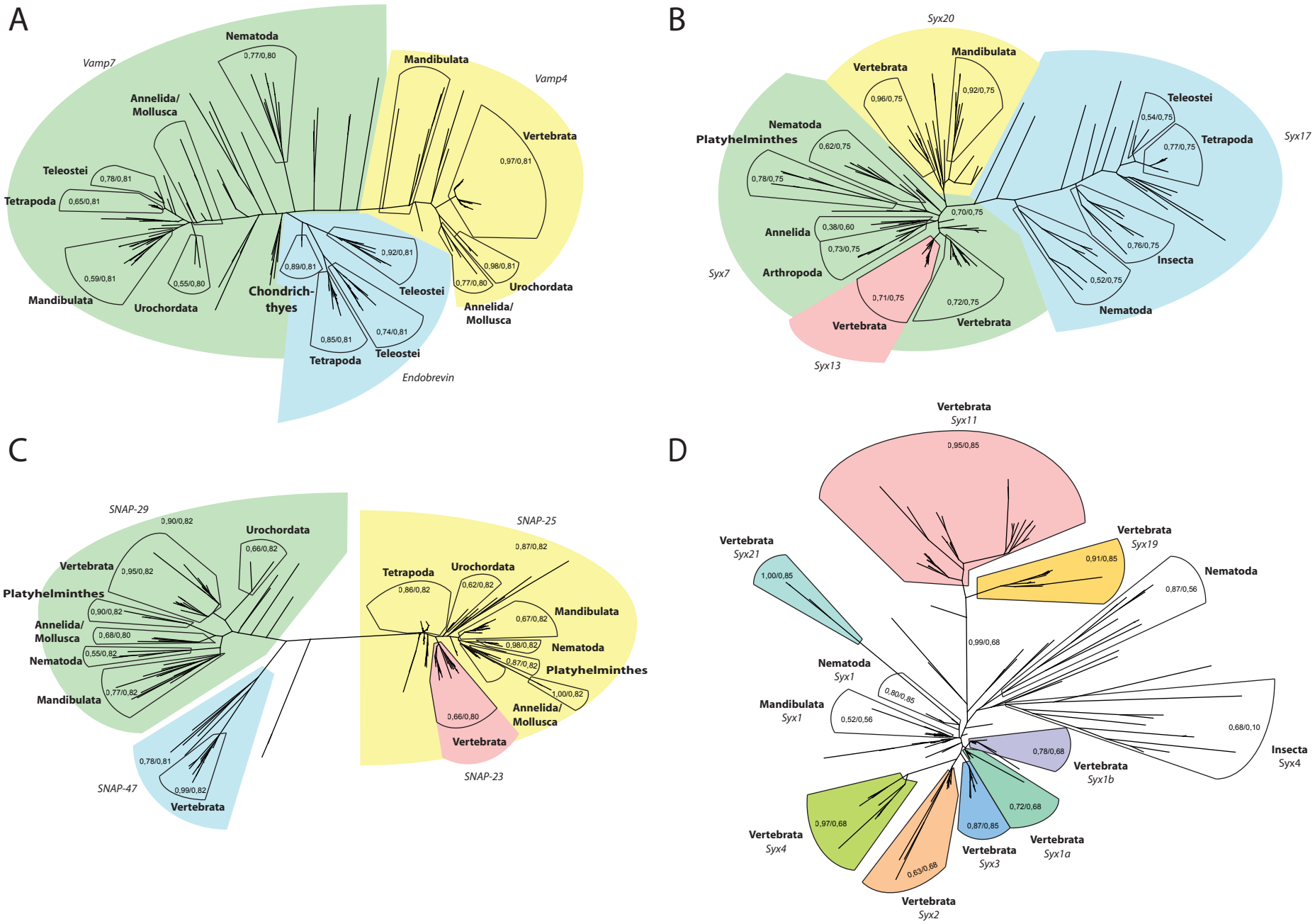
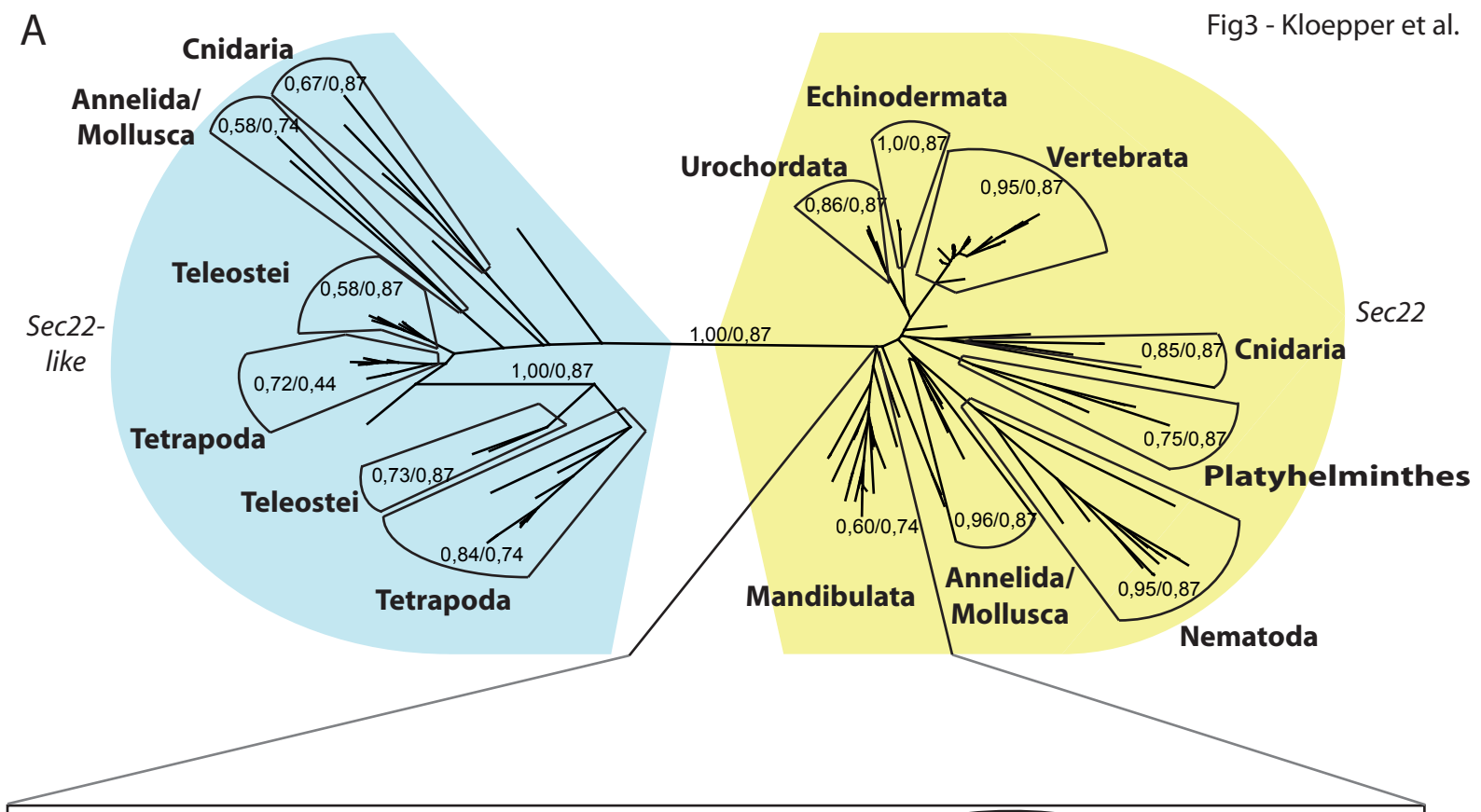
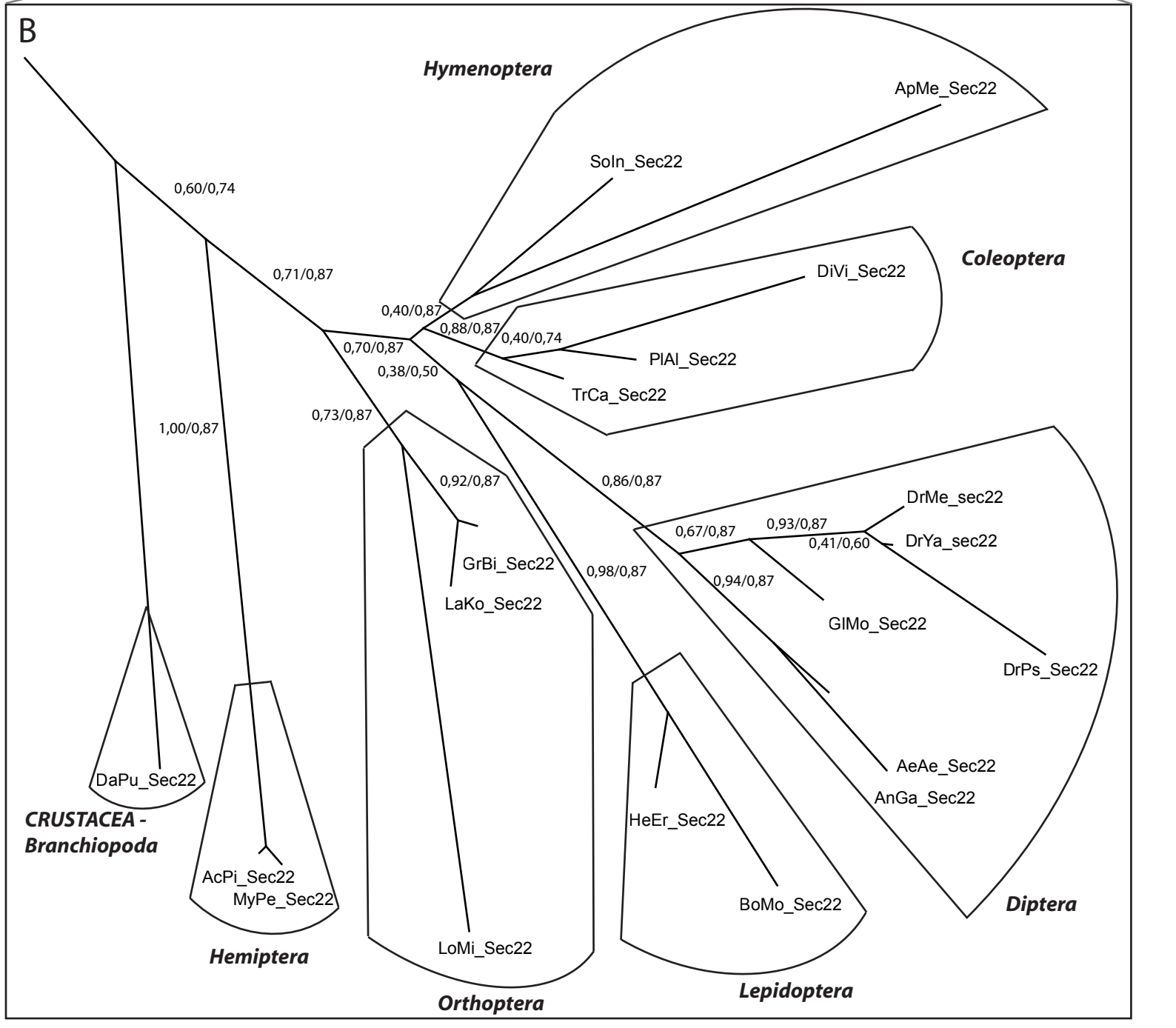


Fig2 - Kloepper et al.


A
Cnidaria
Annelida/
Mollusca
0,67/0,87
0,58/0,74
Teleostei
0,58/0,87
Sec22-
like
0,72/0,44
1,00/0,87
Tetrapoda
0,73/0,87
Teleostei
0,84/0,74
Tetrapoda
Echinodermata
1,0/0,87
Urochordata
Vertebrata
0,86/0,87
0,95/0,87
Sec22
1,00/0,87
0,85/0,87
Cnidaria
0,75/0,87
Platyhelminthes
0,60/0,74
0,96/0,87
Mandibulata
Annelida/
Mollusca
0,95/0,87
Nematoda


B
Hymenoptera
ApMe_Sec22
SoIn_Sec22
0,60/0,74
0,71/0,87
Coleoptera
DiVi_Sec22
0,40/0,87
0,88/0,87
0,40/0,74
0,70/0,87
PlAl_Sec22
0,38/0,50
TrCa_Sec22
0,73/0,87
1,00/0,87
0,92/0,87
0,86/0,87
DrMe_sec22
0,67/0,87
0,93/0,87
0,41/0,60
DrYa_sec22
GrBi_Sec22
0,98/0,87
0,94/0,87
LaKo_Sec22
GlMo_Sec22
DrPs_Sec22
AeAe_Sec22
AnGa_Sec22
DaPu_Sec22
CRUSTACEA -
Branchiopoda
HeEr_Sec22
AcPi_Sec22
MyPe_Sec22
BoMo_Sec22
Diptera
Hemiptera
LoMi_Sec22
Orthoptera
Lepidoptera

Supplemental Data

SNAREing the basis of multicellularity: Consequences of protein family expansion during evolution

Tobias H. Klöpper, C. Nickias Kienle, and Dirk Fasshauer

The supplementary data of this study can be downloaded from the SNARE projects homepage (http://bioinformatics.mpibpc.mpg.de/snare/). The SNARE projects homepage allows access to the SNARE database (Kloepper et al., 2007), into which all newly classified sequences of this study have been integrated.

The supplementary data contain detailed version of the trees depicted in Figures 2a-d and 3a (i), the supplementary tables 1 and 2 (ii), and the generated Nexus files and Alignments of all SNARE types (iii).

i) The following detailed schematic trees of Figures 2a-d and 3a can be downloaded:

Suppl. Fig. 1. Detailed schematic depiction of the unrooted phylogenetic trees of the endosomal R-SNAREs (R.III-type) shown in Fig 2a.

Suppl. Fig. 2. Detailed schematic depiction of the unrooted phylogenetic trees of the endosomal Qa-SNAREs (Qa.III.b-type) shown in Fig 2b.

Suppl. Fig. 3. Detailed schematic depiction of the unrooted phylogenetic trees of the Qbc-SNAREs (Qbc.IV-type) shown in Fig 2c.

Suppl. Fig. 4. Detailed schematic depiction of the unrooted phylogenetic trees of the secretory syntaxins (Qa.IV-type) shown in Fig 2d.

Suppl. Fig. 5. Detailed schematic depiction of the unrooted phylogenetic trees of the secretory syntaxins (R.I-type) shown in Fig 3a.

ii) The following supplementary tables can be downloaded:

Suppl. Table 1. List of all SNARE sequences used.

Suppl. Table 2. List of the SNARE sets of 55 different species, highlighting important changes during the evolution of distinct animal lineages.
The two major expansion, of the SNARE repertoire (i.e. during the transition to multicellularity in animals and during the rise of vertebrates) are indicated in an

additional row, in which the new SNAREs are highlighted in light green. In addition, putative losses of distinct SNARE genes are highlighted in light orange. The additional round of duplication of distinct SNARE genes in bony fishes is highlighted in light yellow.

iii) The following 17 phylogenetic trees in Nexus file format and their corresponding alignments can be downloaded:

SNARE subgroups	***Common names of the basic metazoan SNARE types included in analysis***
Qa.I	Syx18
Qa.II	Syx5
Qa.III.a	Syx16
Qa.III.b	Syx7, Syx17, Syx20
Qa.IV	Syx1
Qb.I	Sec20
Qb.II	Bos1, Gos1
Qb.III	Vti1a, Vti1a
Qc.I	Use1
Qc.II	Bet1, Gs15
Qc.III	Syx6, Syx8
SNAP (Qbc.IV)	SNAP-25, SNAP-29, SNAP-47
R.I	Sec22, Sec22-like
R.II	Ykt6
R.III	Vamp7, Vamp7-like, Vamp4
R.IV	Syb
R.Reg	Tomosyn, Lgl

Kloepper et al. - SupplFig. 1

Vamp7

Vamp4

Endobrevin

Nematoda

AsSu_Vamp7
BrMa_Vamp7
HeGl_Vamp7p
CaEl_Vamp7
CaEl_Vamp7sh
CaBr_Vamp7
CaRe_Vamp7
CaBr_Vamp7sh
AnCe_Vamp7

TrAd_Vamp72
LoGi_Vamp72
ScMe_Vamp7
ScMa_Vamp7
ScJa_Vamp7
PaWe_Vamp7
ScMe_Vamp7sh
DuRy_Vampx
DaPu_Vamp72

Annelida/Mollusca

HeRo_Vamp7
HeRo_Vamp72
CaSp_Vamp7
LoGi_Vamp71
ApCa_Vamp7
MyCa_Vamp7
MoBr_Vamp72
CaOw_Vamp7
CiEl_Vamp7

Teleostei

XiIn_Vamp7
PeMa_Vamp7
PoRe_Vamp7p
FuRu_Vamp7
OsMo_Vamp7
HiHi_Vamp7p
TeNi_Vamp7
OrLa_Vamp7
GaAc_Vamp7

Tetrapoda

XeTr_Vamp7
PoPy_Vamp7
MaMu_Vamp7
GaGa_Vamp7
BoTa_Vamp7
CaFa_Vamp7
HoSa_Vamp7
XeLa_Vamp7
OrAn_Vamp7
RaNo_Vamp7
MoDo_Vamp7
MuMu_Vamp7

TrAd_Vamp7
HyMa_Vamp7
StPu_Vamp7
AsPe_Vamp7
BrFl_Vamp7
NeVe_Vamp7
BoMi_Vamp7
RhAp_Vamp7

Mandibulata

DaPu_Vamp7
AcPi_Vamp7
ApGo_Vamp7
ToCi_Vamp7
HeEr_Vamp7
SpFr_Vamp7
BoMo_Vamp7
LoMi_Vamp7
ApMe_Vamp7
NaVi_Vamp7
TrCa_Vamp7
ChTe_Vamp7
LuLo_Vamp7
AnGa_Vamp7
DrPs_Vamp7
DrWi_Vamp7
GlMo_Vamp7
DrMe_Vamp7

Urochordata

CiIn_Vamp7
MoTe_Vamp7

MoBr_Vamp7
MoOv_Vamp7

TrAd_Vamp7like
NeVe_Vamp7-like

StPu_Vamp7like
CiIn_Vamp7-like
DiLi_Vamp7-like
LoGi_Vamp7-like
BrFl_Vamp7like
CaSp_Vamp7like

0,77/0,80
0,99/0,81
0,78/0,81
0,65/0,81
0,89/0,81
0,59/0,81
0,55/0,80
0,52/0,69
0,56/0,69

Mandibulata

ApMe_Vamp4
SoIn_Vamp4
NaVi_Vamp4
GrBi_Vamp4
BoMi_Vamp4
BrFl_Vamp4
HyMa_Vamp4
NeVe_Vamp4
DaPu_Vamp4

0,65/0,81
0,71/0,62
0,54/0,21
0,24/0,80
StPu_Vamp4p

Vertebrata

OrLa_Vamp4
MiAn_Vamp4
PoRe_Vamp4
TeNi_Vamp4
RuRu_Vamp4
HiHi_Vamp4
DaRe_Vamp4
OsMo_Vamp4
IcPu_Vamp4
FuRu_Vamp4
ScMx_Vamp4p

0,59/0,81
0,67/0,81
0,97/0,81

XeTr_Vamp4
SqAc_Vamp4
AmTi_Vamp4
RaNo_Vamp4
BoTa_Vamp4p
EgCa_Vamp4
LeEr_Vamp4
MuMu_Vamp4
HoSa_Vamp4
OrCu_Vamp4
GaGa_Vamp4
MaMu_Vamp4
MoDo_Vamp4
PaTr_Vamp4
TaGu_Vamp4

CiIn_Vamp4
CiSa_Vamp4

Urochordata

CaSp_Vamp4
MyCa_Vamp4
LoGi_Vamp4

Annelida/Mollusca

Chondrichthyes

LeEr_Endob
SqAc_Endob

0,92/0,81
0,69/0,81
0,74/0,81
0,85/0,81

PlFl_Endob2
OrLa_Endob2
OnMy_Endob2
SaSa_Endob2
HiHi_Endob2
FuHe_Endob2
GaMo_Endob2
IcPu_Endob-2
CyCa_Endob2
CyCa_Endob3
FuRu_Endo2p

Teleostei

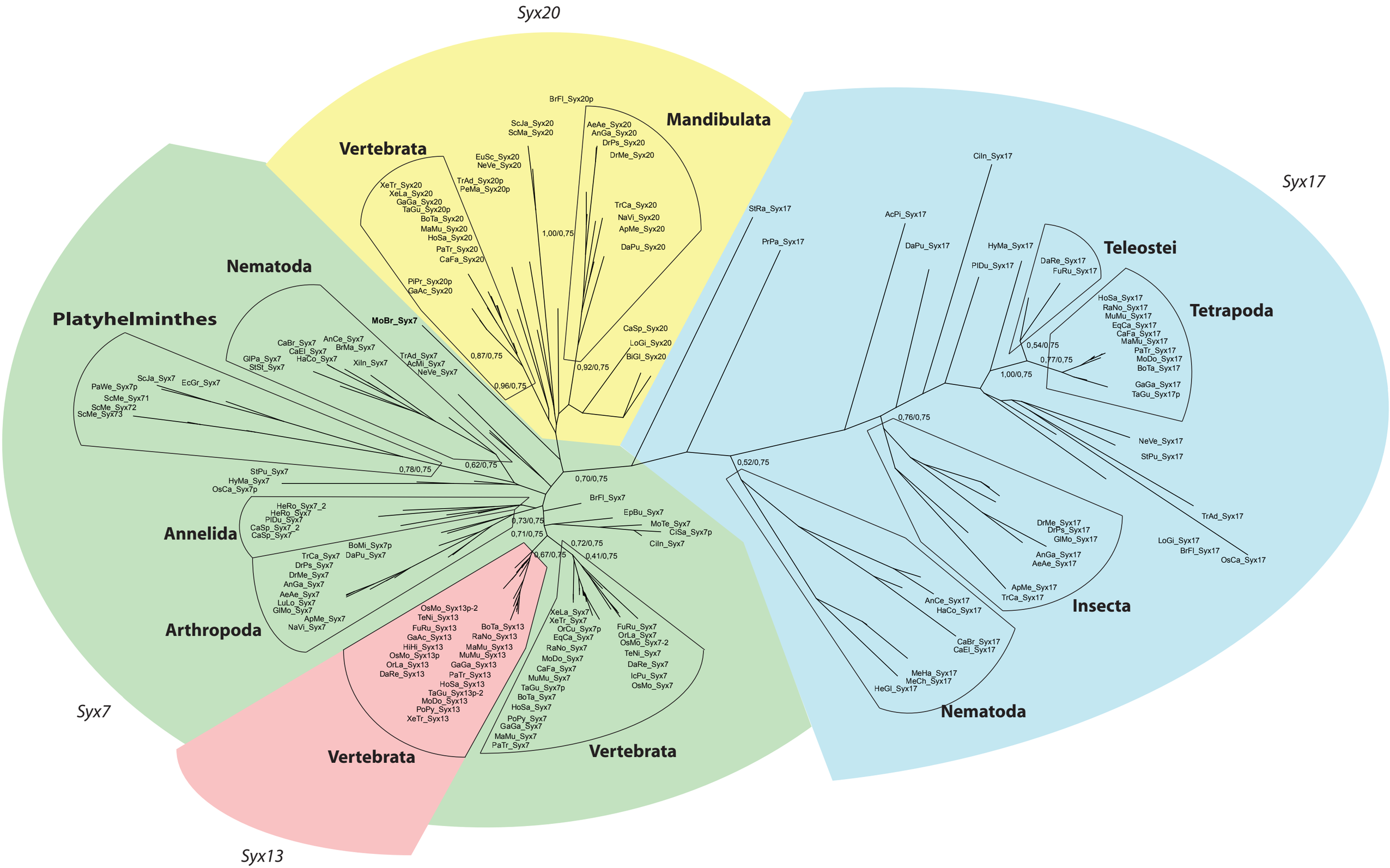
GaAs_Endob
HiHi_Endob
OrLa_Endob
FuRu_Endob
PoRe_Endob
PiPr_Endob
CyCa_Endob
DaRe_Endob
MiAn_Endob
IcPu_Endob
SaSa_Endob
OnMy_Endob

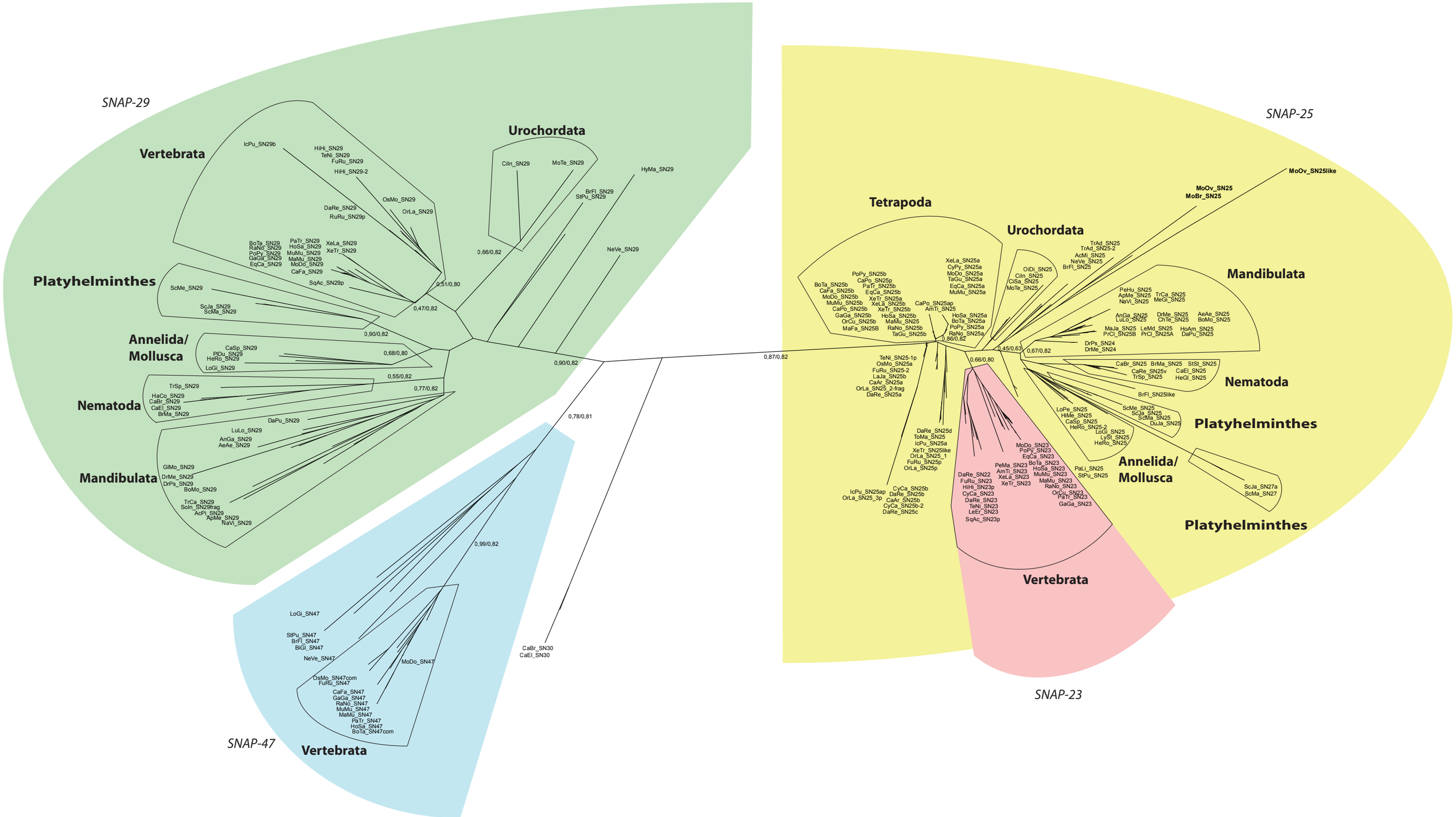
Teleostei

HoSa_Endob
MoDo_Endob
BoTa_Endob
PaTr_Endob
PoPy_endob
CaFa_Endob
XeLa_Endob
XeTr_Endob
MuMu_Endob
MaMu_Endob
RaNo_Endob
CaPo_Endob
EqCa_Endob

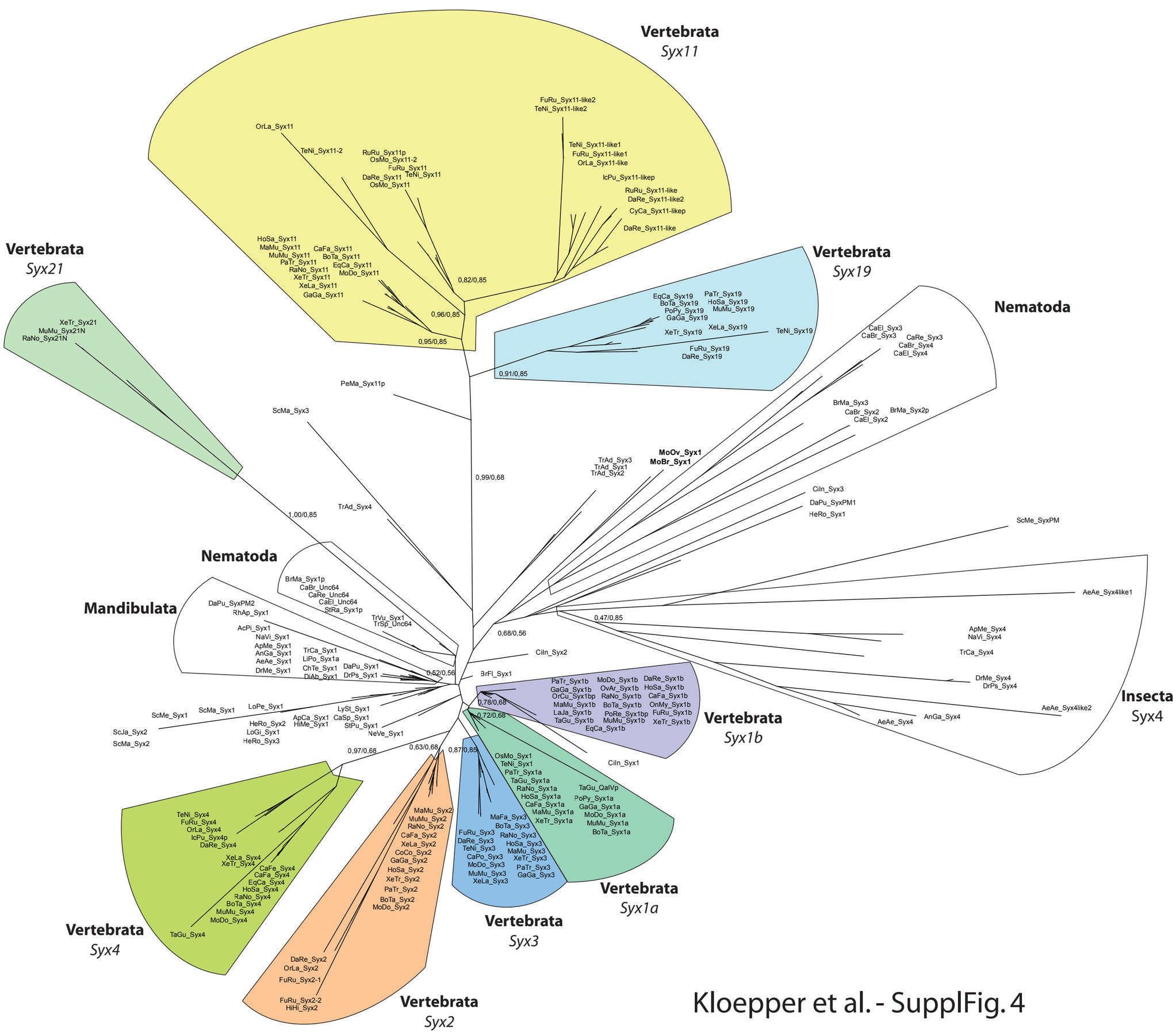
Tetrapoda

Kloepper et al. - SupplFig. 2

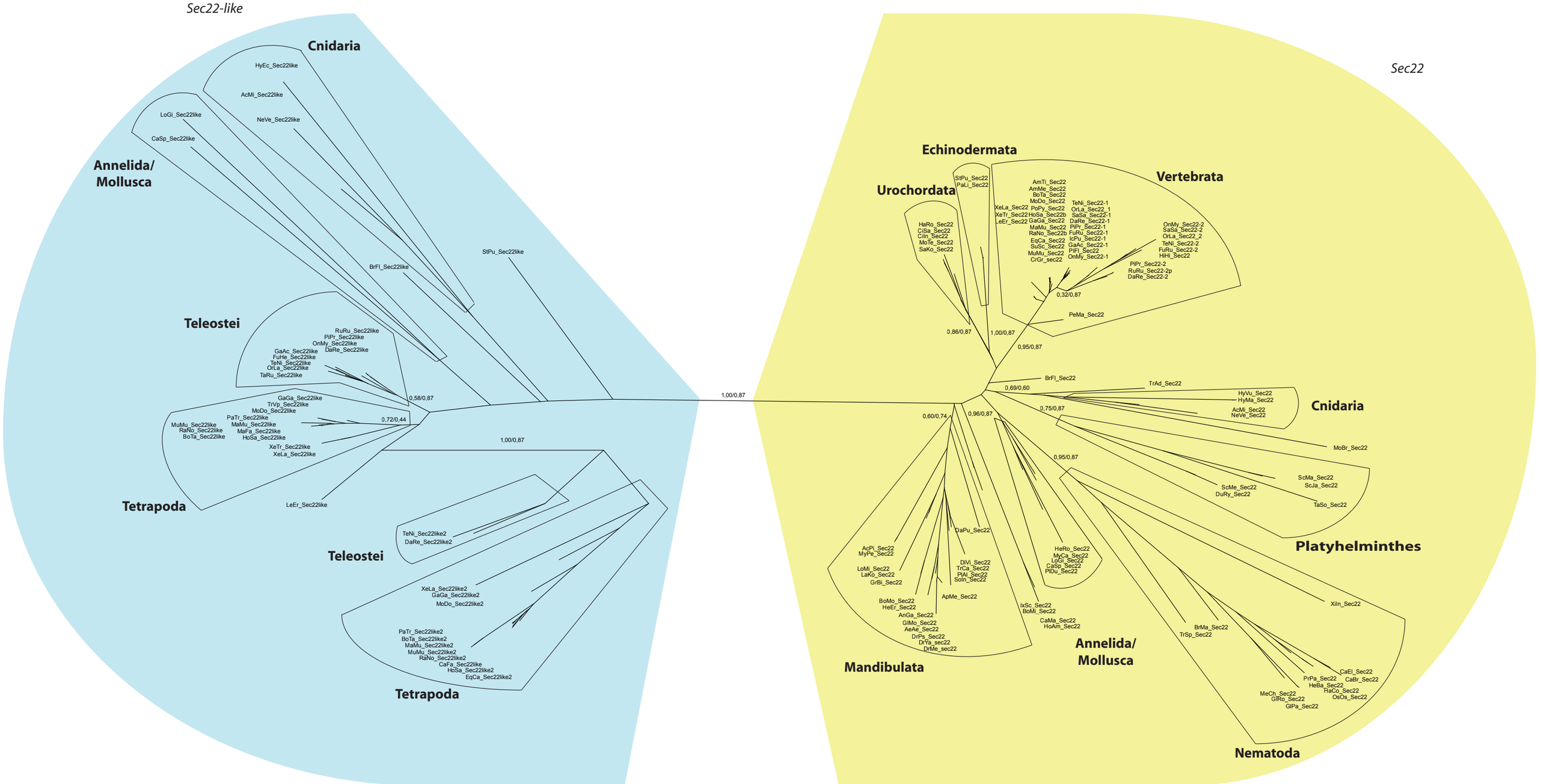




Kloepper et al. - SupplFig. 3



Kloepper et al. - SupplFig. 4



Kloepper et al. - SupplFig. 5

Species-group	Species	Endoplasmatic Reticulum					Golgi Apparatus						TGN	Endosomes/Lysosomes										Secretion						Polarized		
		Qa.I	Qb.I	Qc.I	R.I		Qa.II	Qb.II		Qc.II		R.II	Qa.III.a	Qa.III.b				Qb.III	Qc.III		R.III			Qa.IV	Qbc.IV			R.IV		R.Reg		
		Syx18	Sec20	Use1	Sec22	Sec22-like	Syx5	Memb	Gos28	Bet1	Gs15	Ykt6	Syx16	Syx7	Syx13	Syx17	Syx20	Vti1	Syx6	Syx8	Vamp7	Vamp7-like/Endo	Vamp4	Syx1	SN25	SN29	SN47	Syb	Myob	Tom	Ami	Lgl
Ichthyo-sporea	Amoebidium parasiticum																											*				
	Capsaspora owczarzaki																				*							*				
Choanofl.	Monosiga ovata											*	*					*	*	**	*			*	**			**				
	Monosiga brevicollis	*			*		*	*	*		*	*	*	*				**	*	*	**			*	*			*		*		
Placozoa	Trichoplax adhaerens	*	*	*	*		*	*	*		*	*	*	*		*	*	**	*	*	**	*		****	**		*	**		*		*
Porifera	Oscarella carmela															*	*		*													
Cnidaria	Nematostella vectensis	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*	**	*	*	*	*	*	*	*	*	*	*		*		*
	Acropora aspera								*																							
	Acropora millepora				*	*			*		*		*	*											*			*				
	Acropora palmata								*				*																			
	Montastraea faveolata			*																												
	Hydra magnipapillata	*		*	*		*	*			*	*		*		*		**	*	*	*		*			*	*	*				
	Hydractinia echinata					*					*																					
	Hydra vulgaris				*																											
Annelida/ Mollusca	Platynereis dumerilii		*	*	*			*	*	*		*		*		*		*		*						*		*				
	Hirudo medicinalis																							*	*			*				
	Lumbricus rubellus	*										*																				
	Capitella sp. I	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*	**	*	*	*	*	*	*	*	*	*	*		*		*
	Helobdella robusta	*	*	*	*		*	*	*	*	*	**	*	*			*	*	*		**			***	**	*		***		**		*
	Aplysia californica		*									*						*			*			*		*		*				
	Euprymna scolopes	*		*			*				*	*					*	*	*	*							*					
	Mytilus californianus				*													*			*		*									
	Lymnaea stagnalis																							*	*			*				
	Biomphalaria glabrata		*														*										*					
	Lottia gigantea	*	*	*	*	*	*	*	*	*	*	*	*			*	*	**	*	*	*	*	**	*	*	*	*	*		*		*
	Loligo pealei																							*	*			*				
Echino-dermata	Strongylocentrotus purpuratus	*	*	*	*	*	*	*	*		*	*	*	*		*		*	*	*	*	*	*	*	*	*	*	*		**		*
	Lytechinus variegatus																											*				
	Paracentrotus lividus		*		*			*		*	*							**	*	*					*							
	Asterina pectinifera									*	*										*											
Uro-chordata	Halocynthia roretzi		*		*																			*				*				
	Oikopleura dioica																								*							
	Ciona intestinalis		*	*	*		*	*	*	*	*	*		*		*		**	*	*	*	*	*	***	*	*		*		*		*
	Molgula tectiformis			*	*		*	*	*		*	*		*				*			*				*	*		*		*		
	Ciona savignyi	*	*	*	*		*		*	*	*	*	*	*				*					*		*					*		
	Diplosoma listerianum																				*											
	Saccoglossus kowalevskii				*																											
Cephaloch. Craniata	Branchiostoma floridae	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*	**	*	*	**		*	*	**	*	*	*		**		*
	Eptatretus burgeri														*																	
Vertebrata	Petromyzon marinus				*		*			*	*	*					*	*			*			**	**			*			*	
	Xenopus laevis	*	*	*	*	**	*	*	*	*	*	*	*	*			*	*	*	*	*	*		*****	***	*		**			*	**
	Xenopus tropicalis	*	*	*	*	*	*		*	*		*	*	*	*		*	**	**	*	*	*	*	********	****	*		**	*	**	*	**
	Cynops pyrrhogaster																								*							
	Ambystoma tigrinum			*	*			*															*		***			*	*			
	Ambystoma mexicanum				*																											
	Homo sapiens	*	*	*	*	**	*	*	*	*	*	*	*	*	*	*	*	**	**	*	*	*	*	*******	***	*	*	***	*	**	*	**
	Mus musculus	*	*	*	*	**	*	*	*	*	*	*	*	*	*	*		**	*	*	*	*	*	********	***	*	*	***	*	**	*	**
	Canis familiaris	*	*	*		*	*	*	*		*		*	*		*	*	*	**	*	*	*		*****	*	*	*	**	*	*	*	**
	Rattus norvegicus	*	*	*	*	**	*	*	*	*	*	*	*	*	*	*		*	*	*	*	*	*	*******	***	*	*	***	*	**	*	**
	Pan troglodytes			*		**	*	*	*		*	*	*	*	*	*	*	**	*	*		*	*	******	**	*	*	*	*	**	*	**
	Pongo pygmaeus	*			*				*		*		*	*	*				*		*	*		**	***	*						*
	Macaca mulatta	*	*	*	*	**	*	*	*	*	*	*	*	*	*	*	*	*	**	*	*	*	*	*****	**	*	*	***	*		*	
	Oryctolagus cuniculus			*	*		*	*		*				*									*	*	**			*				
	Macaca fascicularis					*							*											*	**			*				
	Bos taurus	*	*	*	*	**	*	*	*	*	*	*	*	*	*	*	*	*	**	*	*	*	*	*******	***	*	*	***	*	**	*	**
	Ovis aries																							*								
	Cricetulus griseus				*				*																							
	Sus scrofa				*													*														
	Ornithorhynchus anatinus																				*									*	*	
	Monodelphis domestica	*	*	*	*	**		*	*			*	*	*	*	*		**	*	*	*	*	*	******	***	*	*	***		**	*	
	Equus caballus		*		*	*	*	*	*	*		*	*	*		*		**		*		*	*	****	***	*		***	*	*	*	
	Felis catus																							*								
	Cavia porcellus						*				*							*				*		*	***		*					
	Tetraodon nigroviridis	*	*	*	**	**	*	*	**	*	*	*	*	*	*			*	**		*		*	********	**	*	*	**		**	*	**
	Danio rerio	*	*	*	**	**	**	*	*	*	*	*	*	*	*	*		**	**	*		*	*	********	*******	*	*	****	*	**	**	*
	Takifugu rubripes					*																						*				
	Lateolabrax japonicus																							*	*			*				
	Oncorhynchus mykiss				**	*																**		*								
	Carassius auratus																								**							
	Scophthalmus maximus																						*					*				
	Torpedo marmorata																								*							
	Fugu Rubripes	*	**	*	**		*	*	**	*	*	*	*	*	*	*		**	*		*	**	*	*********	****	*	*	****	*	**	*	*
	Gasterosteus aculeatus			*	*	*									*		*				*	*										
	Oryzias latipes		*	*	**	*		*	**	*	*		*	*	*			**	**		*	**	*	*****	*****	*		***	*	***		
	Leucoraja erinacea				*	*				*	*	*								*		*	*		*		*	*				
	Ictalurus punctatus			*	*			*					*	*				*				**	*	**	***	*		**	*			
	Pimephales promelas				**	*							*				*					*										
	Rutilus rutilus				*	*		*	*	*	*							*		*			*	**		*		*			*	
	Cyprinus carpio																	*				***		*	****							
	Misgurnus anguillicaudatus																					*	*						*			
	Salmo salar			*	**																	**										
	Poecilia reticulata	*						*	*	*		*						*			*	*	*	*	**	*		*				
	Osmerus mordax		*	*			*	*	*		*	*		**	**			*			*		*	***	**	*	*	*			*	
	Fundulus heteroclitus					*																*										
	Platichthys flesus				*		*															*										
	Hippoglossus hippoglossus		*		*			*		*		**			*					*	*	**	*	**	*	**						
	Gadus morhua																					*										
	Squalus acanthias								*		*							**		*		*	*		*	*		*			*	
	Gobiocypris rarus																								*							
	Gallus gallus	*	*	*	*	**		*	*	*	*	*	*	*	*	*	*	**	*	*	*		*	******	**	*	*	***		**	*	*
	Taeniopygia guttata		*						*	*			*	*	*	*	*	*					*	****	**			**		**		
	Coturnix coturnix																							*								
	Meleagris gallopavo		*	*																												
Tardigrada	Hypsibius dujardini																											*				
Chelicerata	Acanthoscurria gomesiana										*																					
	Boophilus microplus	*	*	*	*		*	*			*	*		*				*	*	*	*		*									
	Argas monolakensis									*																						
	Blomia tropicalis																											*				
	Ixodes scapularis				*		*			*		*							*													
	Rhipicephalus appendiculatus						*														*			*				*				
	Aleuroglyphus ovatus											*																				
	Mesobuthus gibbosus																								*							
Crustacea	Procambarus clarkii																								**							
	Carcinus maenas				*																							*				
	Homarus americanus			*	*																				*							
	Marsupenaeus japonicus																								*							
	Daphnia magna		*																	*								*				
	Daphnia pulex	*	*	*	*		*	*	*		*	*	**	*		*	*	**	*	*	**		*	***	*	*		**		*		
	Limulus polyphemus																							*								
Insecta	Anopheles gambiae	*	*	*	*		*	*	*	*		*	*	*		*	*	**	*	*	*			**	*	*		**		*		*
	Drosophila melanogaster	*	*	*	*		*	*	*	*		*	*	*		*	*	**	*	*	*			**	**	*		**		*		*
	Drosophila pseudoobscura	*	*	*	*		*	*	*	*		*	*	*		*	*	*		*	*			**	*	*		*				*
	Apis mellifera	*	*		*		*	*	*	*		*	*	*		*	*	*	*	*	*		*	**	*	*		**		*		*
	Leucophaea maderae																								*							
	Bombyx mori		*	*	*		*	*		*									*		*				*	*						
	Culex pipiens																	*														
	Drosophila yakuba				*																							*				
	Tribolium castaneum	*	*	*	*		*	*	*	*		*	*	*		*	*	**	*		*			**	*	*		**		*		*
	Aedes aegypti	*	*	*	*			**		*			*	*		*	*	*	*	*				****	*	*		*		*		*
	Pediculus humanus																								*							
	Lutzomyia longipalpis		*				*	*				*		*						*	*				*	*						
	Chironomus tentans																				*			*	*							
	Acyrthosiphon pisum			*	*		*	*	*							*					*			*		*						
	Glossina morsitans		*		*		*		*					*		*		*		*	*					*						
	Locusta migratoria				*					*											*											
	Oncometopia nigricans																	*														
	Diaprepes abbreviatus			*																				*								
	Platystomus albinus				*																											
	Heliconius erato				*																*											
	Aphis gossypii								*												*											
	Papilio dardanus									*																						
	Rhynchosciara americana									*		*																				
	Myzus persicae	*			*		*	*										*														
	Toxoptera citricida																				*											
	Spodoptera frugiperda																				*											
	Drosophila ananassae																	*														
	Solenopsis invicta				*			*		*		*						*					*			*						
	Gryllus bimaculatus				*																		*									
	Drosophila willistoni																				*											
	Laupala kohalensis				*							*																				
	Nasonia vitripennis	*	*				*	*				*	*	*			*				*		*	**	*	*		**		*		
	Diabrotica virgifera virgifera				*																											
Platyhel-minthes	Schistosoma japonicum	*	*	*	*			*	*	*	*	*	*	*			*	*	*		*			*	**	*		***		*		
	Schistosoma mansoni	*	*	*	*		*	*	*	*	*	*					*	*	*		*			***	**	*		***		*		
	Schmidtea mediterranea			*	*			*		*			**	***				*			**			**	*	*		******		*		
	Dugesia japonica																								*							
	Dugesia ryukyuensis				*														*		*							*				
	Echinococcus granulosus													*					*						*							
	Clonorchis sinensis																											*				
	Paragonimus westermani													*							*											
	Taenia solium				*																											
Nematoda	Caenorhabditis briggsae	*	*	*	*		*	**	*	*		*	*	*		*		*	*		**			****	*	*		***		*		*
	Caenorhabditis elegans	*	*	*	*		*	**	*	*		*	*	*		*		*	*		**			****	*	*		****		*		*
	Caenorhabditis remanei						*	**					*					*			*			**	*			**				
	Heterodera glycines						*	*	*							*		*			*				*							
	Strongyloides stercoralis							*				*		*											*							
	Trichinella spiralis		*		*			*		*		*						*	*					*	*	*		**				
	Haemonchus contortus				*				*					*		*		*								*						
	Bursaphelenchus mucronatus									*																						
	Meloidogyne arenaria									*																						
	Meloidogyne hapla			*						*						*																
	Parastrongyloides trichosuri		**																													
	Trichuris muris												*																			
	Strongyloides ratti															*								*								
	Ancylostoma ceylanicum														*				*		*											
	Globodera pallida		*		*										*			*														
	Pristionchus pacificus				*					*		*				*		*													*	
	Meloidogyne chitwoodi			*	*					*						*		*	*													
	Meloidogyne javanica		*							*																						
	Bursaphelenchus xylophilus		*																									*				
	Necator americanus		*																													
	Xiphinema index				*		*		*						*						*											
	Meloidogyne paranaensis			*						*																						
	Pratylenchus penetrans			*																												
	Globodera rostochiensis			*	*							*																				
	Meloidogyne incognita			*																												
	Trichuris vulpis																							*								
	Teladorsagia circumcincta																											*				
	Ascaris suum																				*											
	Ostertagia ostertagi				*																											
	Brugia malayi	*	*	*	*		*	*	*	*		*	*	*		*					*			***	*	*		**		*		
	Heterorhabditis bacteriophora				*																											

Suppl. Table 1 - Kloepper et al.

Species-group	Species	Endoplasmatic Reticulum					Golgi Apparatus						TGN	Endosomes/Lysosomes									Secretion						Polarized					
		Qa.I	Qb.I	Qc.I	R.I		Qa.II	Qb.II		Qc.II		R.II	Qa.III.a	Qa.III.b			Qb.III	Qc.III			R.III			Qa.IV	Qbc.IV			R.IV	R.Reg					
		Syx18	Sec20	Use1	Sec22		Syx5	Memb	Gos28	Bet1	Gs15	Ykt6	Syx16	Syx7			Vti1	Syx6	Syx8		Vamp7			Syx1	SNAP-25			Syb	Tom					
Choano-flagellata	*Monosiga ovata*											*					*	*	**		*			*		**								
	Monosiga brevicollis	*			*		*	*	*		*	*	*	*			**	*	*		**			*	*			*	*					
Transition to multicellularity					Sec22	Sec22-like								Syx7		Syx17	Syx20	Vti1a	Vti1b			Vamp7	Vamp7-like	Vamp4		SN25		SN29	SN47		Tomosyn		Lgl		
Placozoa	*Trichoplax adhaerens*	*	*	*	*		*	*	*		*	*	*	*		*	*	**		*	*	**	*		****	**			*	**	*		*		
Cnidaria	*Nematostella vectensis*	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*	**		*	*	*	*	*	*	*			*	*	*		*		
	Hydra magnipapillata	*		*	*		*	*			*	*		*		*		**		*	*	*		*				*	*	*				
Mollusca/ Annelida	*Platynereis dumerilii*		*	*	*			*	*	*		*		*		*		*			*							*					
	Capitella sp. I	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*	*	*	*	*	*	*	*	*					*		*		*	
	Helobdella robusta	*	*	*	*		*	*	*	*	*	**	*	*			*	*		*		**			***		**			***	**		*		
	Euprymna scolopes	*		*			*				*	*					*	*	*									*						
	Lottia gigantea	*	*	*	*	*	*	*	*	*	*	*	*			*	*	*	*	*	*	*	**	*		*	*	*	*		*		*		
Echino-dermata	*Strongylocentrotus purpuratus*	*	*	*	*	*	*	*	*		*	*	*	*		*		*	*	*	*	*	*	*		*	*	*	*		**		*		
	Paracentrotus lividus		*		*			*		*	*						*	*	*							*								
Urochordata	*Ciona intestinalis*		*	*	*		*	*	*	*	*	*		*		*		*		*	*	*	*	***		*	*		*		*		*		
	Molgula tectiformis			*	*		*	*	*		*	*		*			*				*					*	*		*		*				
	Ciona savignyi	*	*	*	*		*		*	*	*	*	*	*			*						*			*					*				
Cephaloch.	*Branchiostoma floridae*	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*	*	*	*	*	*	*	*		**	*	*	*		**		*		
Chelicerata	*Boophilus microplus*	*	*	*	*		*	*			*	*		*				*	*	*	*		*											
Crustacea	*Daphnia pulex*	*	*	*	*		*	*	*		*	*	**	*		*	*	*	*	*	*	**		*	***		*	*		**		*			
Insecta	*Anopheles gambiae*	*	*	*	*		*	*	*	*		*	*	*		*	*	*	*	*	*	*			**		*	*		**		*			
	Drosophila melanogaster	*	*	*	*		*	*	*	*		*	*	*		*	*	*	*	*	*	*			**		**	*		**		*		*	
	Drosophila pseudoobscura	*	*	*	*		*	*	*	*		*	*	*		*	*	*	*		*	*			**		*	*		*				*	
	Apis mellifera	*	*		*		*	*	*	*		*	*	*		*	*	*	*		*	*		*	**		*	*		**		*		*	
	Bombyx mori		*	*	*		*	*		*								*			*					*	*							
	Tribolium castaneum	*	*	*	*		*	*	*	*		*	*	*		*	*	*	*		*	*			**		*	*		**		*		*	
	Aedes aegypti	*	*	*	*			**		*			*	*		*	*		*	*		*			****		*	*		*		*		*	
	Lutzomyia longipalpis		*				*	*				*		*						*	*					*	*							
	Glossina morsitans		*		*		*		*					*		*	*			*	*						*							
	Nasonia vitripennis	*	*				*	*				*	*	*							*			*	**		*	*		**		*			
Platyhel-minthes	*Schistosoma japonicum*	*	*	*	*			*	*	*	*	*	*	*			*	*			*				*		**	*		***		*			
	Schistosoma mansoni	*	*	*	*		*	*	*	*	*	*					*	*			*				***		**	*		***		*			
	Schmidtea mediterranea			*	*			*		*			**	***			*					**			**		*	*		******		*			
Nematoda	*Caenorhabditis briggsae*	*	*	*	*		*	**	*	*		*	*	*		*		*				**			****		*	*		***		*		*	
	Caenorhabditis elegans	*	*	*	*		*	**	*	*		*	*	*		*		*				**			****		*	*		****		*		*	
	Trichinella spiralis		*		*			*		*		*					*								*		*	*		**					
	Brugia malayi	*	*	*	*		*	*	*	*		*	*	*		*						*			***		*	*		**		*			
Rise of vertebrates						Sec22-like	Sec22-like2							Syx7	Syx13					Syx6	Syx10			Endo-brevin		Syx1-4, 21	Syx11	SN25a/b	SN23			Syb1-3	Myo-brevin	Tom1	Tom2	Ami	Lgl1	Lgl2
Vertebrata	*Petromyzon marinus*				*			*			*	*	*				*		*				*		*	*	*	*			*					
	Xenopus laevis	*	*	*	*	*	*	*	*	*	*	*	*	*			*	*		*	*	*	*	*		***	**	*	*	*		**			*	*	*
	Xenopus tropicalis	*	*	*	*	*		*		*	*		*	*	*		*	*	*	*	*	*	*	*	*	******	**	**	*	*		**	*	*	*	*	*
	Canis familiaris	*	*	*		*		*	*	*		*		*		*	*	*	*	*	*	*	*	*	*	****	*	*		*	*	**	*	*	*	*	*
	Homo sapiens	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*****	**	*	*	*	*	***	*	*	*	*	*
	Mus musculus	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*	*		*	*	*	*	******	**	*	*	*	*	***	*	*	*	*	*
	Rattus norvegicus	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*		*		*	*	*	*	******	*	*	*	*	*	***	*	*	*	*	*
	Pan troglodytes			*		*	*	*	*			*	*	*	*	*	*	*	*	*		*		*	*	****	**	*	*	*	*	*	*	*	*	*	*
	Macaca mulatta	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	****	*	*	*	*	*	***	*		*		
	Bos taurus	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*****	**	*	*	*	*	***	*	*	*	*	*
	Monodelphis domestica	*	*	*	*	*	*		*	*			*	*	*	*		*	*	*		*	*	*	*	*****	*	*	*	*	*	***		*	*	*	
	Equus caballus		*		*	*		*	*	*			*	*		*		*	*			*		*	**	**	*	*	*		***	*	*		*	
	Gallus gallus	*	*	*	*	*	*		*	*	*	*	*	*	*	*	*	*	*	*		*		*	*	****	**	*	*	*	*	***		*		*	*
	Taeniopygia guttata		*							*	*		*	*	*	*	*	*						*	****		*				**		*	*			
	Tetraodon nigroviridis	*	*	*	**	*	*	*	*	**	*	*	*	*	*	*			*	*	*		*		*	***	*****	*	*	*	*	**		*	*	*	*	*
	Danio rerio	*	*	*	**	*	*	**	*	*	*	*	*	*	*	*		**		*	*	*		*	*	****	****	******	*	*	*	****	*	*	*	**		*
	Fugu Rubripes	*	**	*	**			*	*	**	*	*	*	*	*	*		**			*		*	**	*	*****	****	***	*	*	*	****	*	*	*	*		*
	Oryzias latipes		*	*	**	*			*	**	*	*		*	*			**		*	*		*	**	*	***	**	*****		*		***	*	**	*			
	Ictalurus punctatus			*	*				*					*				*						**	*	**			*	*		**	*				
	Osmerus mordax		*	*				*	*	*			*	**	**				*					*		*	**	**		*	*				*			

Suppl. Table 2 - Kloepper et al.