Supplemental Data

A single ancient origin for prototypical SR splicing factors

Sophie Califice, Denis Baurain[¶], Marc Hanikenne and Patrick Motte[¶]

[¶]To whom correspondence should be addressed.
E-mail: denis.baurain@ulg.ac.be, patrick.motte@ulg.ac.be

Table of contents

Datasets

Text S1. Detailed description of computational procedures

1. RRM distribution across the Tree of Life

1.1. Data downloads

All complete prokaryotic and eukaryotic proteomes available on Ensembl and NCBI FTP servers were downloaded in November 2007 (http://www.ncbi.nlm.nih.gov/sites/genome/). Additional complete proteomes from photosynthetic eukaryotes were downloaded from JGI (*Aureococcus anophagefferens*, *Chlamydomonas reinhardtii*, *Ostreococcus lucimarinus*, *Ostreococcus tauri*, *Phaeodactylum tricornutum*, *Physcomitrella patens*, *Phytophthora ramorum*, *Phytophthora sojae*, *Thalassiosira pseudonana*, and *Volvox carteri*; http://genome.jgi-psf.org/) and TIGR (*Ricinus communis* and *Zea mays*; now located at http://cmr.jcvi.org/) servers. *Cyanidioschyzon merolae* and *Medicago truncatula* complete proteomes were downloaded from their respective project websites (http://merolae.biol.s.u-tokyo.ac.jp/ and http://medicago.org/).

1.2. Prediction of RRM domains

The PFAM (Finn et al., 2010) alignment for the RRM domain (pfam00076; 79 domain sequences from 57 proteins) was downloaded from the NCBI Conserved Domain Database (CDD) (Marchler-Bauer et al., 2009) and used to build a HMM profile with the HMMBUILD component of the HMMER software package (http://hmmer.org/, see also Durbin et al., 1998). As mentioned on the PFAM website (http://pfam.sanger.ac.uk/ family?acc=PF00076), the C-terminal β-strand and final helix are hard to align and have been omitted in the seed alignment, thus yielding a 72-amino-acid (AA) profile slightly truncated relative to the complete RRM structure. RRM domains were predicted in each complete proteome with the program HMMSEARCH (HMMER) using the HMM profile and a loose domain E-value threshold of 1e-5.

1.3. Extraction of RRM domains and RRM-containing proteins

HMMER reports were parsed to extract both the predicted RRM domains and the corresponding full-length proteins in FASTA format. To determine the best domain E-value threshold, sequences were repeatedly extracted using five increasingly stringent thresholds (1e-5, 1e-10, 1e-15, 1e-20, and 1e-25). Eventually, the threshold was set to 1e-10 as this value appeared to give the best compromise between sensitivity and specificity. RRM domains were numbered according to their N-term relative position within each protein (e.g., the second RRM domain of any given protein is numbered RRM2). Consequently, RRM domains located downstream of at least one unpredicted RRM domain (owing to an evolutionarily divergent sequence) were

assigned an incorrect number. This shortcoming was taken into account during annotation and manual analyses of RRM trees (see Sections 2.4-2.5). Finally, proteins containing more than one predicted RRM were extracted only once to ensure accurate statistics.

1.4. Prediction of putative SR proteins

Extracted RRM-containing proteins were tagged as 'putative SR' when they met the rather crude criterion proposed in reference (Boucher et al., 2001), i.e., having at least one occurrence of either quadripeptide 'RSRS' or 'SRSR'. Tagging was carried out for annotation purposes only; thus the absence of the 'SR' tag was never used to exclude a protein from subsequent analyses.
For Figure 1C, RS/SR dipeptides were counted either in the complete dataset of non-redundant 8,042 RRM-containing proteins or in a subset of 197 proteins displaying at least 95% identity to reference sequences confirmed as SR or SRrp proteins (see Table S2 below for details).

1.5. Taxonomic distribution of RRM-containing proteins

For each domain E-value threshold, the numbers of RRM domains, RRM-containing proteins and putative SR proteins predicted in each complete proteome were compiled, along with proteome size and lineage, as fetched from the NCBI Taxonomy server (Benson et al., 2009; Sayers et al., 2009). Proteomes were further categorized according to domain (Archaea, Bacteria, or Eukaryota), complexity (unicellular vs. multicellular) and metabolism (standard vs. capable of oxygenic photosynthesis) and the influence of these factors was assessed using F-tests performed with the R statistical software (R Development Core Team, 2010). Though all eukaryotic and 126 prokaryotic proteomes yielded at least a few RRM-containing proteins, most prokaryotic and viral proteomes appeared devoid of RRM domains (Figure 1A-B and Table S1). These were nevertheless included for completeness.

2. Origin and evolution of SR proteins

2.1. Clustering of RRM domains

To allow phylogenetic analyses, the large number of RRM domains extracted in Section 1.3 (12,023 at the 1e-10 threshold) was reduced to a more tractable dataset through a two-step clustering based on sequence similarity.
First, exactly identical domains were consolidated and renamed after the most inclusive taxonomic group of each cluster members (e.g., a RRM domain found in both *Mus musculus* and *Rattus norvegicus* proteomes would be renamed 'Murinae'). Most of the time, merged domains corresponded to orthologues of closely related organisms or to in-paralogues (i.e.,

terminal gene duplications specific to an organism or to a small clade) or to protein isoforms. This step roughly reduced the number of RRM domains by 50% (6,633 at the 1e-10 threshold).
Second, these unique RRM domains were more aggressively clustered by single-linkage using the (now deprecated) BLASTCLUST component of the NCBI C toolkit (Altschul et al., 1997). Two main parameters control the granularity of the clustering: these are the similarity threshold (fraction of identical residues in the local pairwise alignment) and the coverage threshold (fraction of the longest sequence covered by the alignment) that are to be reached to consider two domains as neighbors. To ensure the best clustering (C_k), we explored the parameter space for both thresholds (by steps of 0.05) and chose the combination of values that maximized a composite score rewarding normalized entropy (Cover and Thomas, 2006) but penalizing the number of singleton clusters (i.e., having only one domain) :

$$Sc = \frac{H(C_k)}{\log K} - 0.8\frac{N_1}{N}$$

where

$$H(C_k) = -\sum_{k=1}^{K} \frac{N_k}{N} \log \frac{N_k}{N}$$

with N = total number of domains
K = number of clusters
N_k = number of domains in cluster k
N_1 = number of singleton clusters

As shown in Figure S2, optimal parameters were as follow: similarity threshold = 0.60 and coverage threshold = 0.85. These values held for domain E-value thresholds ranging from 1e-5 to 1e-20 and this step further reduced the cluster number by 80% (down to 1266 at the 1e-10 threshold). In the clustered datasets, each cluster was represented by the sequence of one of its domains and names were attributed as in the first step, albeit often resulting into much higher level taxa (e.g., Viridiplantae, Eukaryota, cellular organisms).
By default, BLASTCLUST selects the longest sequence as the cluster representative. Since longer sequences are also more prone to be evolutionarily divergent (due to the accumulation of non-conserved insertions), we modified BLASTCLUST to instead allow the selection of the most slowly evolving cluster sequence. In practice, for each cluster, the domain showing the highest average similarity with non-cluster domains was selected as the cluster representative. Our algorithm is an independent re-implementation of the approach used in the SCAFOS software package (Roure et al., 2007). The usefulness of this modification was evaluated by comparing

the length of the phylogenetic trees inferred by parsimony (see Section 2.3 for details) from three alternative datasets based on the selection strategy applied: (i) longest sequence; (ii) most slowly evolving sequence; (iii) fastest-evolving sequence (control experiment). Out of 1266 clusters, there were 352 differences when comparing cluster representatives between strategy (i) and (ii); 299 between (i) and (iii); and 479 between (ii) and (iii). In terms of performance, Figure S3 shows that strategy (ii) generally yielded the shortest trees, followed by strategies (i) and (iii). These results indicate that our algorithmic modification to BLASTCLUST should help to reduce the occurrence of phylogenetic artifacts by actively avoiding the selection of evolutionarily divergent sequences.

In the following, the 1266-cluster dataset built by selection of the most slowly evolving RRM domains was used as the basis of all analyses.

2.2. Alignment of RRM domains

Due to weak conservation combined to a large number of RRM domains to consider, standard alignment methods (e.g., CLUSTALW) led to overly wide alignments with so many gaps that they were practically unusable (data not shown). Alternatively, RRM domains were aligned with HMMALIGN (HMMER) against the HMM profile already used for their prediction (see Section 1.2). To control alignment width, amino acids that did not align to match states of the profile were discarded (option -m), thus literally 'splicing' autapomorphic (and unalignable) insertions out of longer domain sequences. Consequently, the resulting alignment had only 72 AA positions, as many as the HMM profile. Visual inspection of the 1266 domain sequences indicated that this alignment was suitable for large-scale analyses.

2.3. Phylogenetic analyses of RRM domains

The large alignment of RRM domains was analysed by Maximum Parsimony (MP) using PAUP* (Swofford, 2002) (Figure S5) and by Maximum Likelihood (ML) using either TREEFINDER (Jobb et al., 2004) with a WAG+Γ_4 model (Yang, 1993; Whelan and Goldman, 2001) (Figure S7) or RAxML with the very similar PROTMIXWAG model (Stamatakis et al., 2005) (Figure S6) or with the LG+F+Γ_4 model (Le and Gascuel, 2008) (Figure S8). SEQBOOT and CONSENSE from the PHYLIP software package (Felsenstein, 2005) were used respectively to generate 100 bootstrap replicates (Felsenstein, 1985) for each analysis and to compute support values. Bayesian inference with the CAT model known to reduce phylogenetic artifacts (Lartillot and Philippe, 2004; Lartillot et al., 2009; Philippe et al., 2011) was also investigated. However, this sophisticated model designed for phylogenomics did not yield useful results on our sequence-

rich but position-poor dataset, even after having run two independent PHYLOBAYES chains for more than 85,000 cycles (data not shown).

To evaluate our selection algorithm for representative cluster sequences (see Section 2.1) in a reasonable time, PAUP* was invoked with the following options: NREPS=1; ADDSEQ=RANDOM; SWAP=TBR; MAXTREE=10. In contrast, the tree built for phylogenetic purposes was inferred using more time-consuming options: NREPS=10; ADDSEQ=RANDOM; SWAP=TBR; MAXTREE=1000. As a consequence, the latter MP tree (Figure S5; 36,439 parsimony steps) was slightly shorter than the former (Figure S3; 36,617 parsimony steps).

Trees were rooted and left-ladderized using the TREEPLOT component of the MUST software package (Philippe, 1993). Since prokaryotic (and prokaryotic-containing) clusters were mostly concentrated in a small region of the tree, the root was set on the branch separating the bulk of these predominantly prokaryotic clusters from eukaryote-specific clusters. Automatic ladderization based on subtree content helped to order branches so that similarly composed subtrees were easily recognizable by eye across analyses.

2.4. Annotation of RRM trees

Prior to manual examination, RRM trees were automatically annotated using several data sources. To maximize the sensitivity, annotation was actually performed on the individual proteins corresponding to the RRM domains clustered in the trees. Thus, the main challenge was to keep track of the relationships between individual RRM domains (12,023) and RRM-containing proteins (8,042) on one side and cluster representative RRM domains (1266) on the other side, especially considering the two-step clustering and the taxonomic renaming of cluster representatives (see Section 2.1). Another challenge was to efficiently summarize the annotation of potentially numerous domains on the single tree leaf corresponding to each cluster representative. This was addressed by adding to leaf labels all annotation pieces that qualified at least 5% of cluster members, ordered by relevance within each data source. In addition to this relative measurement, stretches of stars ('*') growing with the natural logarithm of the absolute number of domains qualified by each annotation piece were included. Finally, the color code of the most relevant (i.e., with the highest percentage) KOG annotation (see next paragraph and Figure S4) was applied to all labels for which the information was available. Altogether, this considerably facilitated visual comparison of the trees.

Two main data sources were used for annotation: (i) the KOG (Tatusov et al., 2003) component of NCBI CDD and (ii) a hand-curated corpus of about 500 RRM-containing proteins collected in the primary literature (Table S2). In preliminary experiments, other components (SMART, PFAM, COG, and PRK) of NCBI CDD were evaluated, but KOG was ultimately selected as the

most complete source. KOG was mined using RPS-BLAST (a variant of PSI-BLAST) (Altschul et al., 1997) with an E-value threshold of 1e-10, whereas reference proteins were searched using BLASTP (or TBLASTN for reference mRNAs) with an identity threshold of 75%. More stringent thresholds (respectively 1e-25 and 95%) were also considered but eventually abandoned due to a general lack of sensitivity that left most leafs without annotation. The flip side of adopting such a sensitive approach was a potential lack of specificity. That is why we took the resulting annotation with a grain of salt when analyzing the trees.

Other annotation bits added to leaf labels included: (iii) the number of domains in the cluster; (iv) the distribution of RRM numbers in the cluster (see Section 1.3 for the caveat about unpredicted RRM domains); (v) the amount of prokaryotic proteins; and (vi) the amount of 'putative SR' proteins. Notably, all annotation also applied to singleton clusters, except that numbers were dropped for simplicity. Furthermore, the following typographical conventions were enforced: (vii) prokaryotic-containing clusters in italic and (viii) clusters with domains from more than one species in boldface.

2.5. Analyses of an enlarged RRM dataset

Phylogenetic accuracy is known to be difficult to achieve. Consequently, robustness of phylogenetic inference is often assessed by analyzing the same alignment with multiple reconstruction methods and/or evolutionary models. Orthogonal to this approach, it is possible to enlarge or reduce the set of sequences included in the alignment. Experience has indeed shown that varying 'taxon sampling' is very efficient at uncovering phylogenetic artifacts (for reviews, see Delsuc et al., 2005; Philippe et al., 2005; Philippe et al., 2011). Here, both strategies were used to generate five different large RRM trees that were carefully compared to test the hypothesis of a single origin for SR splicing factors (see Section 2.6). As stated in Section 2.3, the first four trees (Figures S5-8) were inferred from the same 1266-cluster alignment. The purpose of this section is to explain how we assembled the huge alignment used for the fifth (even larger) tree (Figure S9).

As an alternative to complete proteomes used throughout this study, two other protein databases were mined for RRM domains: (i) NCBI RefSeq release 26 (Pruitt et al., 2007) and (ii) SMART 6 (Letunic et al., 2009). While RefSeq proteins were processed exactly as in Sections 1.2-1.3 (using the HMM profile for prediction), RRM domains and RRM-containing proteins were directly downloaded through the batch interface of the SMART website (http://smart.embl-heidelberg.de/). In both cases, the usual 1e-10 domain E-value threshold was used for extraction, which respectively yielded 10,739 and 9,978 RRM domains (see Table S1 for taxonomic distributions).

RRM domains from all three sources (including complete proteomes) were then merged into one huge (yet overlapping) dataset and clustered using the same parameter values as in Section 2.1 to give a non-redundant dataset of 1831 clusters (to be compared to the 1266 clusters of complete proteomes alone). Automated alignment with HMMALIGN resulted in an alignment that had again 72 AA positions and that was analyzed using RAxML with the WAG+Γ_4 model. Finally, the tree was rooted on most prokaryotic domains, left-ladderized and annotated, even if its assembly from multiple sources made the process more difficult.

2.6. Comparison of large RRM trees and phylogenetic analysis of the largest RRM clusters

The overall congruence of the large RRM trees was manually assessed by taking advantage of the multiple annotation pieces described in Section 2.4. The primary criterion was the KOG color coding, along with matches to RRM-containing reference proteins. The RRM number was also heavily used for proteins with multiple (up to six) RRM domains. For practical reasons, we mainly focused on leafs (clusters) representing a large number of RRM domains and/or including matches to reference RRM-containing proteins. When comparing the five different trees, two types of issues pertaining to RRM associations were examined: (i) *do proteins from a given KOG consistently group together?* and (ii) *do different KOG consistently group together?* These analyses allowed us to identify 88 groupings that were generally stable across the trees (Table S3). Associations between these groupings are summarized in a qualitative consensus tree (Figure S10). Altogether, mutual affitinities among groups of SR-associated RRM domains suggested that most SR proteins share a single origin (Figure S11 and Table S4).

To check that SR proteins indeed derive from a common ancestor, an additional ML tree focusing on the 152 largest multi-species RRM clusters was computed using RAxML with the LG+F+Γ_4 model (Figures 2 and S12). Briefly, the 1266-cluster RRM alignment was filtered to discard sequences that did not represent at least 8 RRM domains. We also removed 5 more sequences that represented larger RRM clusters, of which sequences all came from a single species, as these corresponded to multiple in-paralogues and/or mis-predicted proteins: Caenorhabditis@F42A6_7a_1_RRM_2, (8 domains), Caenorhabditis@Y46G5A_13_RRM_3 (10), Caenorhabditis@Y73B6BL_6a_3_RRM_2 (10), Caenorhabditis@Y73B6BL_6b_3_RRM_1 (10), and Danio @ENSDARP00000091192_RRM_1 (14). Therefore, the 152 retained sequences represented 10,101 of the 12,023 originally retrieved RRM domains (84%). Strikingly, in this sparser yet still unbiased tree, the RRM1 domains of most SR proteins were resolved as a single clade, while the global phylogenetic resolution was much improved (Figures 2 and S12). This result greatly strengthened our confidence in a single origin of main SR protein architectures.

2.7. Phylogenetic analyses of SR-associated RRM domains

To analyze the subsequent evolution of SR proteins, we decided to focus on three subtrees (shown in pink in Figure S6). For each subtree, RRM domains hidden behind each leaf were extracted from the 6,633-cluster dataset obtained after the first clustering step (see Section 2.1). Starting from this non-redundant dataset allowed us to greatly refine our sampling and to discard any potential mis-clustering caused by the greedy single linkage algorithm of BLASTCLUST, while ensuring that no duplicate domains were included. RRM domains were anew aligned using the HMM profile but this time without discarding insertions to preserve as much phylogenetic information as possible. The resulting alignments were cursorily optimized by hand using the interactive editor ED (MUST) then analyzed by ML using both TREEFINDER and RAXML with the WAG+Γ_4 model, yielding six small RRM trees (Figures S13-18) out of three datasets. After automatic annotation as already described, the trees computed from the RRM1 of SR45/RNPS1 proteins (42 domains x 73 AA; Figures S15-16) and from the RRM2 of dual RRM SR proteins (349 domains x 97 AA; Figures S17-18) were directly interpretable, whereas the poor phylogenetic resolution of those inferred from the RRM1 of SR proteins (434 domains x 93 AA; Figures S13-14) prompted us to experiment with variations in taxon sampling.

In a first variant, we removed 130 fast-evolving domains (i.e., having long branches in the trees) to reduce the occurrence of phylogenetic artifacts, which helps to improve resolution. The two trees inferred from this curated dataset (304 domains x 87 AA) are shown in Figures S19-20. Then, we produced three more variants of this curated dataset by removing domains either from RS proteins (292 domains x 85 AA) or from dual ZnK SR proteins (297 domains x 87 AA) or from both subfamilies at the same time (285 domains x 85 AA). The trees obtained from these additional datasets are not shown but were used to monitor the effect of alternative samplings on the phylogenetic resolution of the relationships among the remaining SR-protein families and subfamilies (Table 1).

Along with RRM1/RRM2 and ZnK sequence logos (see Section 3.2), thorough comparison and integration of 5 (+1) large and 14 small RRM trees (see Section 4.2 for 4 more small trees) together laid the basis of our scenario for the single origin and subsequent diversification of SR proteins into 4 natural families (Figures 2-3 and S10-11 and Tables 1 and S3-4).

3. Profiling of sequence features in SR families

3.1. Extraction of RRM domains of SR-protein subfamilies

Based on the curated RAxML tree of SR-associated RRM domains (Figure S19), RRM1 domain sequences corresponding to each SR-protein subfamily were extracted from the initial 12,023-domain dataset (see Figure 3 for an up-to-date nomenclature). Subfamilies defined in this tree were (i) SC35; (ii); SCL; (iii) SRrp; (iv) 9G8; (v) SRp20; (vi) RSZ; (vii) RS2Z; (viii) animal ASF-like; (ix) plant ASF-like; (x) SRp40-55-75; and (xi) RS. Similarly, the two remaining small RAxML trees were used to extract the RRM1 of (xii) SR45 and of (xiii) RNPS1 (Figure S15), as well as the RRM2 of animal ASF-like and of SRp40-55-75 (Figure S17). Owing to its evolutionary divergence, the RRM2 of the RS family had to be extracted from the large RAxML/WAG tree inferred from complete proteomes (Figure S6). Furthermore, to complete our subfamily-specific RRM domain datasets, the missed RRM2 of the plant ASF-like subfamily was predicted anew using a very loose domain E-value threshold (0.1) directly on the corresponding proteins.

As explained in the main text, the RRM of human SRp54 was also too divergent to be detected by the universal HMM profile. This SR protein was thus not investigated in our study as it would have featured a very long branch prone to phylogenetic artifacts. To complicate matters, the official SRP54 symbol actually corresponds to another protein, the signal recognition particle 54 kDa (ENSG00000100883), which is a ribonucleoprotein belonging to the GTP-binding SRP family and devoid of RRM domain... yet interacting with RNPS1! See http://www.uniprot.org/uniprot/P61011 for details.

3.2. Sequence logos of aligned RRM and ZnK domains of SR-protein subfamilies

Within each subfamily, RRM1 domains were carefully aligned with ED (MUST) using the secondary structure as a guide. To improve alignments, minor domain-specific insertions (1-2 AA) were occasionally spliced out, while a handful of divergent domains were excluded. Then, RRM2 domains from dual RRM SR subfamilies were separately aligned. Finally, subfamilies were aligned with each other based on their consensus sequences to yield a high-quality structural alignment including RRM1 and RRM2 domains from all SR-protein subfamilies. Using different excerpts of this alignment, sequence logos (Schneider and Stephens, 1990) for each subfamily (Figure 4; see also Figure S21) were computed with WEBLOGO (Crooks et al., 2004) after enabling the correction for small samples.

Two overlapping strategies were used to sample prokaryotic RRM domains: (i) either all prokaryotic domains found in the initial dataset or (ii) only those located in the mainly prokaryotic subtrees used to root the large RAxML/WAG tree (Figure S6). The resulting

prokaryotic RRM datasets (cleared of potential eukaryotic contaminants) were processed as SR-protein subfamilies to generate the corresponding sequence logos (Figure S22).

To predict ZnK domains, a HMM profile was built from the PFAM alignment for the zf-CCHC domain (pfam00098; 182 domain sequences from 109 proteins). As for RRM domains before, ZnK domains were predicted in all complete proteomes. However, only those corresponding to single RRM ZnK-like SR subfamilies were actually extracted, thus resulting in four small datasets: 9G8, RSZ, RS2Z (left) and RS2Z (right). To increase the sampling for RSZ and RS2Z, ZnK domains from SR proteins identified in the inventory of additional proteomes (see Sections 4.1-4.4) were also included. Alignments and logos were easy to obtain due to the apparent constraint on size characterizing these domains (Figure S23).

3.3. Profiling of compositional features in SR-protein subfamilies

Full-length proteins corresponding to subfamily datasets used for RRM logos were collected and submitted to blind compositional analyses. First, a 24-AA sliding window was moved (1 AA at a time) along every individual protein. At each step, the numbers of occurrences for all possible overlapping unordered words of size 1 to 3 AA were recorded. Unordered means that occurrences of, e.g., SR and RS dipeptides were consolidated, while overlapping means that, e.g., the SRR tripeptide was counted as either (i) 1xS and 2xR or (ii) 1xRS/SR and 1xRR or (iii) 1xRRS/RSR/SRR, depending on word size. These word densities were then plotted for all proteins of each subfamily (data not shown, but see Dataset S3b-d for an illustration of a similar procedure). To avoid unnecessary clutter, density curves that did not reach a predefined threshold (1 AA: 8; 2-3 AA: 6) for any protein within a subfamily were discarded. After visual inspection of the resulting graphs, a series of words were identified as potentially useful for discriminating the different SR subfamilies (Table S5). Second, this preselection was refined by comparing the distribution (across all proteins of a given subfamily) of each word between subfamilies. Since we were mostly interested in inter-domain compositional features, promising words were tracked in particular protein regions: (i) before the first domain; (ii) between the first and the second domain; and (iii) after the last domain. This process led to one graph per word/region combination, each displaying the distribution of the compositional feature in every SR subfamily (Figure S24). Third, one representative protein was selected out of each subfamily, except for SRp40-55-75 and animal ASF-like, from which were selected three (SRp40, SRp55, SRp75) and two (ASF/SF2 and SRp30c) representatives, respectively. These proteins were used to graphically summarize the results of the compositional profiling (Figures 5 and S25).

3.4. Profiling of conservation in SR-protein subfamilies

The same full-length protein datasets were automatically aligned with CLUSTALW (Thompson et al., 1994). Alignments were then analyzed with the program PLOTCON (EMBOSS software package, Rice et al., 2000) to profile sequence conservation in every subfamily. Briefly, conservation along aligned proteins was computed by averaging all possible pairwise scores (using the EBLOSUM62 similarity matrix) at each position within a 24-AA sliding window. To allow a perfect juxtaposition with composition profiles, subfamily alignments were first cleared of sequence insertions relative to the representative proteins selected in Section 3.3. In practice, columns corresponding to positions missing in these proteins were removed from the alignments using the program NET (MUST), which implies that, for each subfamily, PLOTCON actually analyzed an alignment of exactly the same length as the protein representing the subfamily (Figures 5 and S25).

3.5. Development of a determination key for classifying SR proteins

To help classifying uncharacterized SR proteins, a determination key was developed by integrating most of our analyses: (i) protein architecture; (ii) conserved motifs in RRM domains; and (iii) compositional features (Table S6). This key was validated on candidate SR proteins identified in Sections 4.1-4.4.

4. Inventory of SR proteins in selected organisms

4.1. Optimized prediction of SR-associated RRM1 domains

Our original HMM profile for the RRM domain was deliberately universal in order to sample as much functional diversity as possible. This allowed us to investigate the origin and evolution of SR proteins using the ubiquitous RRM domain as a proxy for whole proteins. However, for uncharacterized proteomes, directly predicting domains that belong to SR proteins is more advantageous. To this end, six new HMM profiles optimized for the prediction of SR-associated RRM1 domains were computed from our alignments. The first three profiles were built from two sequence-rich but automatically HMM-aligned RRM1 subtrees (Figures S12,14; outgroups excluded), whereas the three remaining profiles were derived from RRM1 sequence logos (Figure S21), which had been structurally aligned by hand but lacked fast-evolving sequences.
These six new profiles were then tested on 20 selected proteomes: 13 model organisms already used in the previous analyses and 7 recently sequenced photosynthetic species (downloaded in June 2009; see Tables 2 and S7 for species list). HMMSEARCH was run with the same loose domain E-value threshold as in Section 1.2 (1e-5) but each HMMER report (6 x 20 = 120) was

individually examined to determine the optimal threshold. In practice, we set the threshold in the valley between the best hits (normally derived from SR proteins) and the remaining (non-SR) hits. For most reports, this task was rather easy to perform as the valley was unambiguous (except for *Aspergillus fumigatus* that did not yield any obvious SR-associated RRM1 domain). After consolidation of duplicates (i.e., predicted by more than one profile), a total of 247 candidate SR proteins were retained out of 19 proteomes.

4.2. Phylogenetic analyses of candidate SR protein RRM domains

RRM1 (and RRM2) domains from candidate SR proteins were extracted and automatically aligned on our structural alignment of SR-protein subfamily consensus sequences (see Section 3.2). This was carried out with the BABA component of the MUST software package, which also affiliated each candidate RRM domain to the most appropriate SR-protein subfamily.

The resulting high-quality alignment was then analyzed as in Section 2.3 (with both TREEFINDER and RAXML heuristics, as well as with PHYLOBAYES). Again, different variants of the dataset were considered: (i) with or without RRM2 domains; (ii) including or excluding positions corresponding to subfamily-specific insertions. Finally, we settled on a dataset containing all SR-associated RRM1 domains, along with RRM2 domains from plant RS proteins to test the duplication hypothesis (see main text), but no outgroup domains to minimize phylogenetic artifacts (Philippe et al., 2011). The trees, which eventually excluded subfamily-specific insertions, were rooted on the SR45/RNPS1 RRM domains (for convenience) and automatically annotated (Figures S26-29; see also Table 1 for the phylogenetic resolution of relationships among SR-protein families and subfamilies).

4.3. Architectural and compositional analyses of candidate SR proteins

The architecture of every candidate SR protein was elucidated by scripting the NCBI Conserved Domain Database web server (http://www.ncbi.nlm.nih.gov/cdd/). Consequently, domain identification was not restricted to RRM and ZnK domains but took advantage of the full breadth of domains available in CDD (Dataset S3a).

In parallel, candidate SR proteins were submitted to the blind compositional analyses already described, except that all candidate SR proteins were processed at the same time, rather than subfamily by subfamily as when selecting discriminating features (Dataset S3b-d).

4.4. Curation of candidate SR proteins

To assemble our curated inventory of SR proteins, candidates were validated by integrating the results of all aforementioned analyses. At the domain level, we considered the subfamily attributed by the alignment procedure as well as the phylogenetic position of the RRM1 domain in the last trees (Figures S26-29). At the protein level, we applied our determination key (Table S6) based on the architectural and compositional profiling of candidates (Dataset S3a-d). For dubious cases, confirmatory searches were conducted on the NCBI BLAST website (http://blast.ncbi.nlm.nih.gov/). This stepwise validation led to the confirmation of 192 proteins as genuine SR proteins along with 52 isoforms, of which the overwhelming majority could be classified into one of the previously identified subfamilies (Table 2; see also Table S7).

Table S1. Taxonomic distribution of RRM-containing proteins. The three Excel sheets summarize the statistics for RRM domain, RRM-containing and putative SR protein content per proteome for each data source: (i) complete proteomes; (ii) NCBI RefSeq; and (iii) SMART database. Statistics are given for the five domain E-value thresholds (1e-5, 1e-10, 1e-15, 1e-20 and 1e-25) used in the prediction of RRM domains. Proteomes are sorted by taxonomic lineage, except that proteomes devoid of RRM domains are listed at the bottom.

Table S2. Hand-curated corpus of RRM-containing proteins collected in the primary literature. Reference proteins are organized by family. For convenience, an independent bibliography is provided at the end of the corresponding file.

Table S3. Qualitative comparison of the five large RRM domain trees. The Table is a companion to Figure S10. Node numbers match those given in tree annotations (Figures S5-9) and in other comparative display elements (Figures S10-11 and Table S4). Unstable subgroups are enclosed in brackets. RRM#: RRM number; KOG#: abridged KOG number and name; MP: 1266-cluster parsimony tree; RW: 1266-cluster RAxML (WAG+Γ_4) tree; TW: 1266-cluster TreeFinder (WAG+Γ_4) tree; RL: 1266-cluster RAxML (LG+F+Γ_4) tree; mRX: 1831-cluster RAxML tree; y/n: does/does not exist in the corresponding tree.

Table S4. Summary of the groupings observed for SR protein RRM1 domains. The Table is a companion to Figure S11. Datasets are described in the text. Node numbers are as in Figure S10 and Table S3: (16) RRM1 domains of dual RRM SR proteins; (17) RRM1 domains of single RRM ZnK-like SR proteins; (22-23) RRM1 domains of single RRM SR proteins; (26-27) RRM1 domains of the atypical RNPS1/SR45 proteins.

Table S5. Compositional features specific of one or more SR subfamilies. For each subfamily, discriminating features used in the SR determination key and secondary features are listed. Features correspond to single, unordered 2-AA or 3-AA words, respectively. The SRp40-55-75 (SRSF5-6-4) and animal ASF-like (SRSF1-ASF/SF2 and SRSF9/SRp30c) subfamilies were split in their individual component proteins thus resulting in 16 rows in the Table.

Table S6. Determination key for SR families and subfamilies. The key was constructed by integrating architectural and compositional features that are specific either of SR natural families or subfamilies.

Table S7. Statistics for the curated inventory of SR proteins in 20 selected proteomes. The Table is a companion to Table 2 found in the main text. It provides detailed numbers for the stepwise validation of candidate SR proteins from RRM prediction to phylogenetic analysis and determination key classification. cand: candidate SR proteins; rej: candidates rejected as non-SR proteins; iso: isoforms of confirmed SR proteins (mis-modeling or alternative splicing); conf: candidates confirmed as SR proteins; key: confirmed candidates correctly classified by the determination key; trunc: truncated candidates (mis-modeling or alternative splicing) that were incorrectly classified by the determination key. The 7 recently sequenced photosynthetic added for the inventory are shown in boldface.

Figure S1. Scheme of the analysis pipeline. The four sections of the main text results are shown by discontinuous shaded areas (labeled 1 to 4). Within an area, each colored box corresponds to a logical analysis step (e.g., program execution followed by processing of its output). For simplicity, data and execution flows are confounded in the same arrows. Input data sources and pipeline outputs are symbolized by distinct icons. Both main and Supplemental Figure and Table numbers are indicated on the corresponding icons. See Text S1 for detailed methods and technical results.

Figure S2. Optimization of clustering parameters for RRM domains. Non-redundant RRM domains were clustered by single-linkage based on sequence similarity (fraction of identical residues). The parameter space for both similarity and coverage thresholds was explored to determine the combination that maximized a composite score rewarding normalized entropy but penalizing the number of singleton clusters. The analysis was conducted for each of the five non-redundant RRM datasets that were obtained at the following domain E-value thresholds: *(A)* 1e-5, *(B)* 1e-10, *(C)* 1e-15, *(D)* 1e-20 and *(E)* 1e-25.

Figure S3. Evaluation of the selection algorithm for cluster representatives. Three clustered datasets were assembled using either of the following selection criteria for cluster representatives: (i) longest sequence (orange); (ii) most slowly evolving sequence (green); (iii) fastest-evolving sequence (red). For each dataset, the corresponding tree was inferred by Maximum Parsimony along with 100 bootstrap replicates. The resulting tree lengths (in parsimony steps) are plotted either as a vertical line for the original tree or as a cumulative distribution for bootstrap replicates.

Figure S4. Color key for KOG annotation. A unique combination of background and text color has been attributed to each KOG, which are listed by descending occurrence count in the RRM-containing protein dataset.

Figure S5. Maximum Parsimony tree of representative RRM domains. The tree was obtained using PAUP* from the analysis of 72 aligned amino acid positions across 1266 sequences and rooted using prokaryotic RRM domains as outgroups. The length of the tree is 36,439 parsimony steps. Leaves were automatically annotated and colored based on cluster content using (i) KOG annotation and (ii) reference RRM-containing proteins (Figure S4 and Table S2). Furthermore, the following information is provided: (iii) the number of domains in the cluster; (iv) the distribution of RRM numbers in the clusters; (v) the amount of prokaryotic proteins; and

(vi) the amount of putative SR proteins. All numbers are given as percentages as well as logarithmic estimates represented as stretches of stars (*). Clusters that contain domains from more than one species or at least one prokaryotic RRM are shown in bold and italic, respectively. Stable groupings are indicated and follow the numbering of Figure S10 and Table S3. Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S6. RAxML tree of representative RRM domains obtained under the WAG+Γ_4 model. The tree was actually obtained under the very similar PROTMIXWAG model from the analysis of 72 aligned amino acid positions across 1266 sequences and rooted using prokaryotic RRM domains as outgroups. Leaves were automatically annotated and colored based on cluster content using (i) KOG annotation and (ii) reference RRM-containing proteins (Figure S4 and Table S2). Furthermore, the following information is provided: (iii) the number of domains in the cluster; (iv) the distribution of RRM numbers in the clusters; (v) the amount of prokaryotic proteins; and (vi) the amount of putative SR proteins. All numbers are given as percentages as well as logarithmic estimates represented as stretches of stars (*). Clusters that contain domains from more than one species or at least one prokaryotic RRM are shown in bold and italic, respectively. Stable groupings are indicated and follow the numbering of Figure S10 and Table S3. SR-associated subtrees that were further analyzed are shown in pink. Subtrees prok-1 to prok-3 used to compute prokaryotic RRM logos are also indicated. Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S7. TREEFINDER tree of representative RRM domains obtained under the WAG+Γ_4 model. The tree was obtained from the analysis of 72 aligned amino acid positions across 1266 sequences and rooted using prokaryotic RRM domains as outgroups. Leaves were automatically annotated and colored based on cluster content using (i) KOG annotation and (ii) reference RRM-containing proteins (Figure S4 and Table S2). Furthermore, the following information is provided: (iii) the number of domains in the cluster; (iv) the distribution of RRM numbers in the clusters; (v) the amount of prokaryotic proteins; and (vi) the amount of putative SR proteins. All numbers are given as percentages as well as logarithmic estimates represented as stretches of stars (*). Clusters that contain domains from more than one species or at least one prokaryotic RRM are shown in bold and italic, respectively. Stable groupings are indicated and follow the numbering of Figure S10 and Table S3. Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S8. RAxML tree of representative RRM domains obtained under the LG+F+Γ_4 model. The tree was obtained from the analysis of 72 aligned amino acid positions across 1266 sequences and rooted using prokaryotic RRM domains as outgroups. Leaves were automatically annotated and colored based on cluster content using (i) KOG annotation and (ii) reference RRM-containing proteins (Figure S4 and Table S2). Furthermore, the following information is provided: (iii) the number of domains in the cluster; (iv) the distribution of RRM numbers in the clusters; (v) the amount of prokaryotic proteins; and (vi) the amount of putative SR proteins. All numbers are given as percentages as well as logarithmic estimates represented as stretches of stars (*). Clusters that contain domains from more than one species or at least one prokaryotic RRM are shown in bold and italic, respectively. Stable groupings are indicated and follow the numbering of Figure S10 and Table S3. Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S9. RAxML tree of the enlarged RRM dataset obtained under the WAG+Γ_4 model. The tree was actually obtained under the very similar PROTMIXWAG model from the analysis of 72 aligned amino acid positions across 1831 sequences and rooted using prokaryotic RRM domains as outgroups. Leaves were automatically annotated and colored based on cluster content using (i) KOG annotation and (ii) reference RRM-containing proteins (Figure S4 and Table S2). Furthermore, the following information is provided: (iii) the number of domains in the cluster; (iv) the distribution of RRM numbers in the clusters; (v) the amount of prokaryotic proteins; and (vi) the amount of putative SR proteins. All numbers are given as percentages as well as logarithmic estimates represented as stretches of stars (*). Clusters that contain domains from more than one species or at least one prokaryotic RRM are shown in bold and italic, respectively. Stable groupings are indicated and follow the numbering of Figure S10 and Table S3. Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S10. Qualitative consensus of phylogenetic analyses of representative RRM domains. The consensus is based on a manual comparison of the 5 large trees (Figures S5-9). Recurrent nodes across analyses were uniquely numbered by descending grouping level. Unstable subgroups are enclosed in brackets. Notes about individual nodes and trees can be found in companion Table S3. Simple terminal branches correspond to functionally homogenous groupings, whereas open triangles represent mixed functionality. SR-associated subtrees are indicated by shaded boxes and the corresponding branches are color-coded to match Figures 2, 3 and S11: RRM1 of single RRM SR proteins (green); RRM1 of single RRM ZnK-like SR proteins

(blue); RRM1 of dual RRM SR proteins (red); RRM1 of RNPS1/SR45 proteins (orange); RRM2 of non-plant dual RRM SR proteins (violet); RRM2 of the plant-specific 'RS' group of SR proteins (brown). Note that node names are based on automated KOG annotation without manual curation, hence the presence of occasional inaccuracies (e.g., node 83 labelled as 'splicing factor RNPS1' while it actually includes PABP proteins; see Table S3).

Figure S11. Mutual affinities of SR protein RRM domains in the five large trees. Cartoons *(A)* to *(E)* summarize the trees given in Figures S5 to S9, respectively, and leaf node numbers are described in Figure S10 and Table S3. Simple terminal branches correspond to functionally homogenous groupings, whereas open triangles represent mixed functionality. SR-associated RRM domains are color-coded as in Figures 2, 3 and S10, while outgroup RRM domains are shown in black. Companion Table S4 summarizes the groupings observed in each tree for the RRM1 domains of SR splicing factors.

Figure S12. RAxML tree of the largest RRM clusters obtained under the LG+F+Γ_4 model. The tree was obtained from the analysis of 72 aligned amino acid positions across 152 slowly evolving RRM sequences representative of all multi-species RRM clusters with ≥ 8 member domains (i.e., 10,101 out of 12,023 retrieved domains), and rooted using prokaryotic RRM clusters as outgroups (in italic). Leaves were automatically annotated and colored based on cluster content using (i) KOG annotation and (ii) reference RRM-containing proteins (Figure S4 and Table S2). Furthermore, the following information is provided: (iii) the number of domains in the cluster; (iv) the distribution of RRM numbers in the clusters; (v) the amount of prokaryotic proteins; and (vi) the amount of putative SR proteins. All numbers are given as percentages as well as logarithmic estimates represented as stretches of stars (*). Stable groupings are indicated and follow the numbering of Figure S10 and Table S3. All non-zero bootstrap proportions are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S13. Exhaustive RAxML tree of SR-associated RRM1 domains (WAG+Γ_4 model). The tree was actually obtained under the very similar PROTMIXWAG model from the analysis of 93 aligned amino acid positions across 434 sequences and was left unrooted as 'outgroups' are not monophyletic. Leaves correspond to unique domain sequences belonging to clusters present in the 'RRM1 of SR proteins' subtree (in pink in Figure S6). Annotation is as described for Figures S5-9. Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S14. Exhaustive TreeFinder tree of SR-associated RRM1 domains (WAG+Γ_4 model). The unrooted tree was obtained from the analysis of 93 aligned amino acid positions across 434 sequences and was left unrooted as 'outgroups' are not monophyletic. Leaves correspond to unique domain sequences belonging to clusters present in the 'RRM1 of SR proteins' subtree (in pink in Figure S6). Annotation is as described for Figures S5-9. Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S15. RAxML tree of RRM1 domains from RNPS1-like proteins (WAG+Γ_4 model). The tree was actually obtained under the very similar PROTMIXWAG model from the analysis of 73 aligned amino acid positions across 42 sequences and rooted between RNPS1 and SR45 domains. Leaves correspond to unique domain sequences belonging to clusters present in the 'RRM1 of SR45/RNPS1' subtree (in pink in Figure S6). Annotation is as described for Figures S5-9. Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S16. TreeFinder tree of RRM1 domains from RNPS1-like proteins (WAG+Γ_4 model). The unrooted tree was obtained from the analysis of 73 aligned amino acid positions across 42 sequences and rooted between RNPS1 and SR45 domains. Leaves correspond to unique domain sequences belonging to clusters present in the 'RRM1 of SR45/RNPS1' subtree (in pink in Figure S6). Annotation is as described for Figures S5-9. Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S17. RAxML tree of RRM2 domains from dual RRM SR proteins (WAG+Γ_4 model). The tree was actually obtained under the very similar PROTMIXWAG model from the analysis of 97 aligned amino acid positions across 349 sequences and rooted between SR-associated and outgroup domains. Leaves correspond to unique domain sequences belonging to clusters present in the 'RRM2 of dual RRM SR proteins' subtree (in pink in Figure S6). Annotation is as described for Figures S5-9. Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S18. TreeFinder tree of RRM2 domains from dual RRM SR proteins (WAG+Γ_4 model). The unrooted tree was obtained from the analysis of 97 aligned amino acid positions across 349 sequences and rooted between SR-associated and outgroup domains. Leaves correspond to unique domain sequences belonging to clusters present in the 'RRM2 of dual RRM SR proteins'

subtree (in pink in Figure S6). Annotation is as described for Figures S5-9. Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S19. RAxML tree of slowly evolving SR-associated RRM1 domains (WAG+Γ_4 model). The tree was actually obtained under the very similar PROTMIXWAG model from the analysis of 87 aligned amino acid positions across 304 sequences and was left unrooted as 'outgroups' are not monophyletic. Leaves correspond to unique domain sequences belonging to clusters present in the 'RRM1 of SR proteins' subtree (in pink in Figure S6), of which 130 fast-evolving (i.e., long-branched) sequences have been removed. Annotation is as described for Figures S5-9. Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S20. TreeFinder tree of slowly evolving SR-associated RRM1 domains (WAG+Γ_4 model). The tree was obtained from the analysis of 87 aligned amino acid positions across 304 sequences and was left unrooted as 'outgroups' are not monophyletic. Leaves correspond to unique domain sequences belonging to clusters present in the 'RRM1 of SR proteins' subtree (in pink in Figure S6), of which 130 fast-evolving (i.e., long-branched) sequences have been removed. Annotation is as described for Figures S5-9. Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S21. Sequence conservation in the first RRM domain of SR proteins. RRM sequence logos are aligned based on secondary structure (bottom). At a given position, the height of any residue is proportional to its frequency while overall stack height corresponds to sequence conservation. Error bars reflects the uncertainty of conservation estimates. Small sample sizes lead to conservatively downscaled letters (e.g., RS2Z subfamily and SR45). RNP1 and RNP2 motifs are shaded, as are conserved aromatic residues in these motifs. β1-4, α1-2 and L1-5 respectively stand for β-sheets, α-helices and linkers of the RRM secondary structure. The Figure contains four panels, one for each natural family of SR proteins: *(A)* single RRM; *(B)* single RRM ZnK-like; *(C)* dual RRM; and *(D)* RNPS1-like. In *(C)*, RS2Z proteins were grouped with other ZnK-containing proteins even if their RRM is quite different, including RNP motifs.

Figure S22. Sequence conservation across prokaryotic RRM domains. RRM sequence logos are aligned based on secondary structure (bottom). At a given position, the height of any residue is proportional to its frequency while overall stack height corresponds to sequence conservation. Error bars reflects the uncertainty of conservation estimates. RNP1 and RNP2 motifs are shaded,

as are conserved aromatic residues in these motifs. β1-4, α1-2 and L1-5 respectively stand for β-sheets, α-helices and linkers of the RRM secondary structure. The 2 upper logos respectively include all prokaryotic or cyanobacterial RRM domains independently of their actual position in the phylogenetic tree (Figure S6). The 3 lower logos each correspond to one of the mainly prokaryotic subtrees used as outgroups, of which non-prokaryotic domains were excluded. Grey boxes highlight motifs specific of the discrete cyanobacterial subtree (middle).

Figure S23. Sequence conservation across ZnK domains of SR proteins. Thanks to their very constraint size, ZnK sequence logos were easily aligned by hand. At a given position, the height of any residue is proportional to its frequency while overall stack height corresponds to sequence conservation. Error bars reflects the uncertainty of conservation estimates. To increase sampling for RSZ and RS2Z logos, ZnK domains from SR proteins identified in the inventory of the seven additional photosynthetic proteomes were also included. Left and right ZnK domains from RS2Z proteins were considered separately in an attempt to determine their respective ancestry by comparison with SRSF7/9G8 and RSZ homologues (grey boxes). However, this analysis proved inconclusive due to the short size of ZnKs. Universally conserved CCHC motifs (yellow) and glycine residues (orange) are also shaded. ZnKs of the CCHC family have amino acid arrangements with the consensus sequence 'C-φ-X-C-G-±-X-G-H-X3-δ-C,' where X is a variable amino acid, φ is an aromatic amino acid, (±) is a charged amino acid, and δ is a carbonyl-containing residue (Green and Berg, 1989; Cavaloc et al., 1994; Cavaloc et al., 1999; Armas et al., 2008).

Figure S24. Distributions of relevant compositional features within SR subfamilies. Each panel corresponds to the distributions of the density of a particular word (1, 2 or 3 unordered residues) in a specific region across all subfamilies. Regions are bounded by RRM and ZnK domains and defined as follows: (i) before the first domain; (ii) between the first and the second domain; and (iii) after the last domain (e.g., RS/SR before RRM1 or GG between RRM1 and ZnK1). For a given word/region combination, density is computed as the number of overlapping word occurrences in a sliding-window of 24 amino acid residues scanning the region in each protein. Maximal word densities (i.e., highest peaks) are then pooled by subfamily to allow joint plotting of resulting distributions on a single graph. Numbers in panel legends indicate the number of analyzed proteins in each subfamily. This number can vary between word/region combinations because of (i) mis-modeled proteins lacking one or more canonical domains of the family and (ii) non-prediction of divergent domains. Similarly, not all

combinations can be analyzed for every subfamily, e.g., single RRM SR proteins do not possess an interdomain region.

Figure S25. Architecture, conservation and compositional profile of SR subfamilies. Within each panel, an individual protein was selected as representative for compositional profiling while conservation was computed on the alignment of all subfamily sequences. The SRp40-55-75 (SRSF5-6-4) and animal ASF-like (SRSF1-ASF/SF2 and SRSF9/SRp30c) subfamilies were split in their individual component proteins thus resulting in 16 panels. In these cases, a single subfamily alignment was used for conservation profiling in respectively three (K, L, M) and two (H, I) panels. Only informative features (Table S5) are shown, with discriminating features used in the key (Table S6) in plain style and secondary features in dashed style. *(A)* SC35: human SRSF2/SC35 (ENSP00000293233); *(B)* SCL: Arabidopsis At-SCL33 (NP_001031195); *(C)* SRrp: mouse SRSF10/SRrp40 (ENSMUSP00000030437); *(D)* 9G8: human SRSF7/9G8 (ENSP00000325905); *(E)* SRp20: human SRSF3/SRp20 (ENSP00000362820); *(F)* RSZ: Arabidopsis At-RSZ21 (NP_564208); *(G)* RS2Z: rice Os-RS2Z36 (NP_001054489); *(H)* ASF/SF2: human SRSF1-ASF/SF2 (ENSP00000258962); *(I)* SRp30c: human SRSF9/SRp30c (ENSP00000229390); *(J)* ASF-like: rice Os-SR32 (NP_001050082); *(K)* SRp40: human SRSF5/SRp40 (ENSP00000342294); *(L)* SRp55: human SRSF6/SRp55 (ENSP00000244020); *(M)* SRp75: human SRSF4/SRp75 (ENSP00000362900); *(N)* RS: Arabidopsis At-RS41 (NP_200017); *(O)* SR45: Arabidopsis SR45 (NP_*173107); (P)* RNPS1: human RNPS1 (ENSP00000380275).

Figure S26. RAxML RRM tree of candidate SR proteins from selected proteomes (WAG+Γ_4 model). The tree was actually obtained under the very similar PROTMIXWAG model from the analysis of 64 unambiguously aligned amino acid positions across 270 sequences and arbitrarily rooted using RNPS1-like proteins. Positions corresponding to subfamily-specific insertions were discarded from the full alignment (82 amino acid positions) prior to analysis. RRM2 domains from RS proteins were also included to test the duplication hypothesis (see main text). Annotation was limited to KOG identifiers and reference RRM-containing proteins, supplemented with a subfamily code added during automatic alignment on subfamily consensus (e.g., asflkpl stands for plant ASF-like). Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S27. TreeFinder RRM tree of candidate SR proteins from selected proteomes (WAG+Γ_4 model). The tree was obtained from the analysis of 64 unambiguously aligned amino acid positions across 270 sequences and arbitrarily rooted using RNPS1-like proteins. Positions

corresponding to subfamily-specific insertions were discarded from the full alignment (82 amino acid positions) prior to analysis. RRM2 domains from RS proteins were also included to test the duplication hypothesis. Annotation was limited to KOG identifiers and reference RRM-containing proteins, supplemented with a subfamily code added during automatic alignment on subfamily consensus (e.g., asflkpl stands for plant ASF-like). Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S28. RAxML tree of candidate SR proteins from selected proteomes (LG+F+Γ_4 model). The tree was obtained from the analysis of 64 unambiguously aligned amino acid positions across 270 sequences and arbitrarily rooted using RNPS1-like proteins. Positions corresponding to subfamily-specific insertions were discarded from the full alignment (82 amino acid positions) prior to analysis. RRM2 domains from RS proteins were also included to test the duplication hypothesis. Annotation was limited to KOG identifiers and reference RRM-containing proteins, supplemented with a subfamily code added during automatic alignment on subfamily consensus (e.g., asflkpl stands for plant ASF-like). Bootstrap proportions >50% are shown. The scale bar at the bottom gives the number of substitutions per site.

Figure S29. PHYLOBAYES tree of candidate SR proteins from selected proteomes (CAT+Γ_4 model). The tree was obtained from the analysis of 64 unambiguously aligned amino acid positions across 270 sequences. It is the consensus of two independent chains ran for 55,000 cycles (burn-in = 5000 cycles; one tree sampled every 100 cycles) and is arbitrarily rooted using RNPS1-like proteins. Positions corresponding to subfamily-specific insertions were discarded from the full alignment (82 amino acid positions) prior to analysis. RRM2 domains from RS proteins were also included to test the duplication hypothesis. Annotation was limited to KOG identifiers and reference RRM-containing proteins, supplemented with a subfamily code added during automatic alignment on subfamily consensus (e.g., asflkpl stands for plant ASF-like). Posterior probabilities (PP) >50% are shown (whereas nodes with PP <5% are collapsed). The scale bar at the bottom gives the number of substitutions per site.

Supplemental References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:** 3389-3402

Armas P, Nasif S, Calcaterra NB (2008) Cellular nucleic acid binding protein binds G-rich single-stranded nucleic acids and may function as a nucleic acid chaperone. J Cell Biochem **103:** 1013-1036

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2009) GenBank. Nucleic Acids Res **37:** D26-31

Boucher L, Ouzounis CA, Enright AJ, Blencowe BJ (2001) A genome-wide survey of RS domain proteins. RNA **7:** 1693-1701

Cavaloc Y, Bourgeois CF, Kister L, Stevenin J (1999) The splicing factors 9G8 and SRp20 transactivate splicing through different and specific enhancers. RNA **5:** 468-483

Cavaloc Y, Popielarz M, Fuchs JP, Gattoni R, Stevenin J (1994) Characterization and cloning of the human splicing factor 9G8: a novel 35 kDa factor of the serine/arginine protein family. EMBO J **13:** 2639-2649

Cover TM, Thomas JA (2006) Elements of Information Theory. Wiley-Interscience, New York, Ed Second Edition

Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. Genome Res **14:** 1188-1190

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

Durbin R, Eddy S, Krogh A, Mitchinson G (1998) Biological sequence analysis: probabilistic models of proteins and nucleic acids. Cambridge University Press, Cambridge

Felsenstein J (1985) Confidence Limits on Phylogenies: An Approach Using the Bootstrap. Evolution **39:** 783-791

Felsenstein J (2005) PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author, Seattle, Department of Genome Sciences, University of Washington

Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer EL, Eddy SR, Bateman A (2010) The Pfam protein families database. Nucleic Acids Res **38:** D211-222

Green LM, Berg JM (1989) A retroviral Cys-Xaa2-Cys-Xaa4-His-Xaa4-Cys peptide binds metal ions: spectroscopic studies and a proposed three-dimensional structure. Proc Natl Acad Sci U S A **86:** 4047-4051

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

Lartillot N, Lepage T, Blanquart S (2009) PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics **25:** 2286-2288

Lartillot N, Philippe H (2004) A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Molecular biology and evolution **21:** 1095-1109

Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. Molecular biology and evolution **25:** 1307-1320

Letunic I, Doerks T, Bork P (2009) SMART 6: recent updates and new developments. Nucleic Acids Res **37:** D229-232

Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Tasneem A, Thanki N, Yamashita RA, Zhang D, Zhang N, Bryant SH (2009) CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Res **37:** D205-210

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

Philippe H, Brinkmann H, Lavrov DV, Littlewood DT, Manuel M, Worheide G, Baurain D (2011) Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS biology **9:** e1000602

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

Pruitt KD, Tatusova T, Maglott DR (2007) NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res **35:** D61-65

R Development Core Team (2010) A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria

Rice P, Longden I, Bleasby A (2000) EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet **16:** 276-277

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

Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Edgar R, Federhen S, Feolo M, Geer LY, Helmberg W, Kapustin Y, Landsman D, Lipman DJ, Madden TL, Maglott DR, Miller V, Mizrachi I, Ostell J, Pruitt KD, Schuler GD, Sequeira E, Sherry ST, Shumway M, Sirotkin K, Souvorov A, Starchenko G, Tatusova TA, Wagner L, Yaschenko E, Ye J (2009) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res **37:** D5-15

Schneider TD, Stephens RM (1990) Sequence logos: a new way to display consensus sequences. Nucleic Acids Res **18:** 6097-6100

Stamatakis A, Ludwig T, Meier H (2005) RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. Bioinformatics **21:** 456-463

Swofford DL (2002) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Inc., Sunderland, MA.

Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA (2003) The COG database: an updated version includes eukaryotes. BMC Bioinformatics **4:** 41

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22:** 4673-4680

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

Yang Z (1993) Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. Mol Biol Evol **10:** 1396-1401

organism	name	description	accession	#RRM	references
▷ **ARRS — acidic rich RS protein (RRM and RS domains)**					
Mus musculus	Psc1	–	AAS19274	1	Kavanagh et al. (2005)
Mus musculus	ARRS	–	BAC34721	1	Kavanagh et al. (2005)
Homo sapiens	KIAA1311	–	BAA92549	1	Kavanagh et al. (2005)
Homo sapiens	Se70-2	–	AAH41655	1	Kavanagh et al. (2005)
Xenopus laevis	RBM26	–	AAH43744	1	Kavanagh et al. (2005)
Drosophila melanogaster	ARRS	–	NP_609976	1	Kavanagh et al. (2005)
Anopheles gambiae	ARRS	–	XP_318628	1	Kavanagh et al. (2005)
Caenorhabditis elegans	ARRS	–	NP_498234	1	Kavanagh et al. (2005)
Dictyostelium discoideum	ARRS	–	AA051188	1	Kavanagh et al. (2005)
Homo sapiens	RBM 26	–	NP_071401	2	Kavanagh et al. (2005)
Homo sapiens	RBM 27	–	NP_061862	1	Kavanagh et al. (2005)
▷ **Bruno — CUG-BP1 and ETR-3 like factors (CELF)/BRUNO-like proteins**					
Homo sapiens	CUGBP 1	CUG triplet repeat, RNA-binding protein 1	NP_006551	3	Timchenko et al. (1996)
Homo sapiens	CUGBP 2	ETR-3	NP_001020247	3	Lu et al. (1999)
Homo sapiens	CELF 3	Bruno-like 3	NP_064565	3	Ladd et al. (2001)
Homo sapiens	CELF 4	Bruno-like 4	Q9BZC1	3	Ladd et al. (2001)
Homo sapiens	CELF 5	Bruno-like 5	NP_068757	3	Ladd et al. (2001)
Homo sapiens	CELF 6	Bruno-like 6	NP_443072	3	Ladd et al. (2004)
Mus musculus	CUGBP 1	–	NP_059064	3	Timchenko et al. (2006)
Mus musculus	BRUNOL-4	–	NP_573458	3	Meins et al. (2002)
Mus musculus	BRUNOL-6	–	NP_780444	3	McKee et al. (2005)
Gallus gallus	CUGBP 1	–	NP_001012539	3	Brimacombe and Ladd (2007)
Rattus norvegicus	CUGBP 1	–	NP_001020592	3	Karagiannides et al. (2006)
Danio rerio	CUGBP 1	–	NP_571688	3	Hashimoto et al. (2006)
Xenopus laevis	EDEN-BP	embryo deadenylation element binding protein	NP_001084196	3	Paillard et al. (1998)
▷ **CAPER — coactivator of activating protein-1 (AP-1) and estrogen receptors (Ers)**					
Homo sapiens	CAPER	–	NP_909122	3	Jung et al. (2002)
Mus musculus	CAPER	–	NP_573505	3	Jung et al. (2002)
▷ **CPEB — cytoplasmic polyadenylation element binding protein**					
Homo sapiens	CPEB 1	–	NP_085097	2	Welk et al. (2001)
Homo sapiens	CPEB 2	–	NP_872587	2	Hagele et al. (2008)
Homo sapiens	CPEB 3	–	NP_055727	2	Salehi-Ashtiani et al. (2006)
Homo sapiens	CPEB 4	–	NP_085130	2	Kurihara et al. (2003)

organism	name	description	accession	#RRM	references
Xenopus laevis	CPEB	–	NP_001084072	2	Paris et al. (1991)
Mus musculus	CPEB 1	–	NP_031781	2	Gebauer and Richter (1996)
Mus musculus	CPEB 2	–	NP_787951	2	Kurihara et al. (2003)
Mus musculus	CPEB 3	–	NP_938042	2	Theis et al. (2003)
Mus musculus	CPEB 4	–	NP_080528	2	Theis et al. (2003)
Rattus norvegicus	CPEB 1	–	NP_001099746	2	Wu et al. (1998)
▷ **Dazap — Deleted in azoospermia-associated protein**					
Homo sapiens	Dazap1	DAZ associated protein 1	NP_061832	2	Tsui et al. (2000)
Mus musculus	Dazap1	DAZ associated protein 1	NP_573451	2	Dai et al. (2001)
Mus musculus	Dazap1	DAZ associated protein 1	NP_001116077	2	Dai et al. (2001)
Rattus norvegicus	Dazap1	DAZ associated protein 1	NP_001020913	2	Pan et al. (2005)
Xenopus laevis	Dazap1	DAZ associated protein 1 and Prrp (proline-rich RNA binding protein)	Q98SJ2	2	Zhao et al. (2001)
▷ **ELAV — embryonic lethal, abnormal vision, Drosophila-like 1/Hu**					
Mus musculus	ELAVL 1/HuR	Hu antigen R	NP_034615	3	Atasoy et al. (1998)
Mus musculus	ELAVL 2	Hu antigen B	NP_034616	3	Abe et al. (1996b)
Mus musculus	ELAVL 3	Hu antigen C	NP_034617	3	Okano and Darnell (1997)
Mus musculus	ELAVL 4	Hu antigen D	NP_034618	3	Abe et al. (1994)
Homo sapiens	ELAVL 1/HuR	Hu antigen R	NP_001410	3	Ma et al. (1996)
Homo sapiens	ELAVL 2/HuB	Hu antigen B	NP_004423	3	Levine et al. (1993)
Homo sapiens	ELAVL 3	Hu antigen C	NP_001411	3	Van Tine et al. (1998)
Homo sapiens	ELAVL 4	Hu antigen D	NP_068771	3	Szabo et al. (1991)
Gallus gallus	ELAVL 4	Hu antigen D	NP_990161	3	Wakamatsu and Weston (1997)
Gallus gallus	ELAVL 1/HuR	Hu antigen R	NP_990164	3	Wakamatsu and Weston (1997)
Rattus norvegicus	ELAVL 1/HuR	Hu antigen R	NP_001102318	3	Tang et al. (2002)
Rattus norvegicus	ELAVL 2/HuB	Hu antigen B	NP_775431	3	Fornaro et al. (2007)
Rattus norvegicus	ELAVL 3	Hu antigen C	NP_758827	3	Abe et al. (1996a)
Rattus norvegicus	ELAVL 4	Hu antigen D	NP_001071119	3	Steller et al. (1996)
▷ **FCAFPA — other Arabidopsis RRM proteins FCA and FPA**					
Arabidopsis thaliana	FCA	–	CAB10407	2	Macknight et al. (1997)
Arabidopsis thaliana	FCA-like	–	AAB63831	2	Lorkovic and Barta (2002)
Arabidopsis thaliana	FPA	–	AAB64314	3	Schomburg et al. (2001)
Arabidopsis thaliana	AtY14	–	AAG52616	1	Park and Muench (2007)
Arabidopsis thaliana	AtREF	–	NP_200803	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	AtREF	–	AAG51767	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	AtREF	–	CAB85990	1	Lorkovic and Barta (2002)

organism	name	description	accession	#RRM	references
Arabidopsis thaliana	AtREF	–	NP_198588	1	Lorkovic and Barta (2002)
▷ **FOX — Feminizing locus on X**					
Caenorhabditis elegans	Fox-1	–	NP_508446	1	Hodgkin et al. (1994)
Homo sapiens	Fox-1	ataxin 2-binding protein 1	NP_665898	1	Shibata et al. (2000)
Homo sapiens	Fox-2	RBM 9	NP_001076045	1	Lieberman et al. (2001)
Homo sapiens	Fox-3	hexaribonucleotide binding protein	NP_001076044	1	Underwood et al. (2005)
Mus musculus	Fox-1	ataxin 2-binding protein	NP_067452	1	Kiehl et al. (2001)
Mus musculus	Fox-2	RBM 9	NP_444334	1	Lieberman et al. (2001)
Danio rerio	Fox-1	ataxin 2-binding protein-like	NP_999940	1	Jin et al. (2003)
▷ **G3BP — Ras-GTPase-activating protein SH3-domain-binding protein**					
Homo sapiens	G3BP 1	–	NP_938405	1	Parker et al. (1996)
Homo sapiens	G3BP 2	–	NP_987100	1	Prigent et al. (2000)
Mus musculus	G3BP 1	–	NP_038744	1	Kennedy et al. (1996)
▷ **GRSRBP — Glycine-Rich-RNA Binding (GR-RBP) and small RNA-binding protein (S-RBP)**					
Arabidopsis thaliana	GR-RBP1	–	AAD22311	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	GR-RBP2	–	CAB36849	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	GR-RBP3	–	BAB10366	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	GR-RBP4	–	BAB03001	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	GR-RBP5	–	AAG52402	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	GR-RBP6	–	AAF98412	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	GR-RBP7	–	AAD23639	1	Heintzen et al. (1997)
Arabidopsis thaliana	GR-RBP8	–	CAB43641	1	Carpenter et al. (1994)
Arabidopsis thaliana	AtRZ-1a	–	BAB02203	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	AtRZ-1b	–	AAB71977	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	AtRZ-1c	–	NP_196048	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	AtRZ-1-like	–	AAG51392	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	S-RBP1	–	CAB78428	1	Lorkovic and Barta (2002)
Zea mays	GRP	–	NP_001105707	1	Didierjean et al. (1992)
Nicotiana sylvestris	RZ-1	–	BAA06012	1	Hanano et al. (1996)
Homo sapiens	GRSF 1	G-rich RNA sequence binding factor 1	NP_002083	3	Qian and Wilusz (1994)
Homo sapiens	RBM3	–	NP_006734	1	Danno et al. (1997)
Homo sapiens	CIRP	–	NP_001271	1	Nishiyama et al. (1997)
Mus musculus	CIRP	–	NP_031731	1	Nishiyama et al. (1997)

organism	name	description	accession	#RRM	references
▷ **IGF2BP — insulin-like growth factor 2 RNA binding protein**					
Homo sapiens	IGF2BP 1	–	NP_006537	2	Nielsen et al. (1999)
Homo sapiens	IGF2BP 2	–	NP_006539	2	Nielsen et al. (1999)
Homo sapiens	IGF2BP 3	–	NP_006538	2	Nielsen et al. (1999)
▷ **LA — LA**					
Homo sapiens	La	–	NP_003133	1	Maraia and Intine (2001)
Homo sapiens	Larp 7	–	NP_056269	1	Krueger et al. (2008)
Mus musculus	Ssb	Sjogren syndrome antigen B	NP_001103615	1	Maraia and Intine (2001)
Mus musculus	Larp7	–	NP_613059	1	Krueger et al. (2008)
Arabidopsis thaliana	La1	–	NP_001031777	1	Fleurdepine et al. (2007)
Arabidopsis thaliana	La	–	NP_178106	2	Fleurdepine et al. (2007)
Arabidopsis thaliana	La	–	NP_974184	2	Fleurdepine et al. (2007)
Arabidopsis thaliana	La	–	NP_974185	2	Fleurdepine et al. (2007)
Arabidopsis thaliana	La	–	NP_188540	1	Fleurdepine et al. (2007)
Arabidopsis thaliana	La	–	NP_851141	1	Fleurdepine et al. (2007)
Oryza sativa	La	–	BAD19607	2	Fleurdepine et al. (2007)
Oryza sativa	La	–	CAE03115	2	Fleurdepine et al. (2007)
Schizosaccharomyces pombe	la1	–	NP_593315	1	Van Horn et al. (1997)
Saccharomyces cerevisiae	Lhp1	–	NP_010232	1	Yoo and Wolin (1994)
Drosophila melanogaster	La	–	NP_724257	1	Yoo and Wolin (1994)
▷ **LARK — LARK**					
Homo sapiens	RBM4	LARK	NP_002887	2	Markus and Morris (2006)
Mus musculus	RBM4	LARK	NP_033058	2	Kojima et al. (2007)
Drosophila melanogaster	LARK	–	NP_523957	2	McNeil et al. (2001)
Bombyx mori	LARK	–	NP_001037293	2	Wang et al. (2005)
▷ **Musashi — Musashi**					
Homo sapiens	musashi 1	neural-specific RNA binding protein	NP_002433	2	Good et al. (1998)
Homo sapiens	musashi 2	neural-specific RNA binding protein	NP_620412	2	Barbouti et al. (2003)
Mus musculus	Musashi 1	neural-specific RNA binding protein homolog 1	NP_032655	2	Sakakibara et al. (1996)
Mus musculus	Musashi 2	neural-specific RNA binding protein homolog 2	NP_473384	2	Sakakibara et al. (2001)
Xenopus laevis	nrp-1B	neural-specific RNA binding protein	NP_001084040	2	Richter et al. (1990)
Caenorhabditis elegans	Musashi-1	neural-specific RNA binding protein	NP_497799	2	Yoda et al. (2000)
Ciona intestinalis	Musashi	neural-specific RNA binding protein	NP_001027721	2	Kawashima et al. (2000)
Halocynthia roretzi	Musashi	neural-specific RNA binding protein	BAA82622	2	Kawashima et al. (2000)

organism	name	description	accession	#RRM	references
▷ **NTF — RRM proteins containing the NTF domain**					
Arabidopsis thaliana	NTF-RRM1	–	AAF20222	2	Lorkovic and Barta (2002)
Arabidopsis thaliana	NTF-RRM2	–	AAD20086	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	NTF-RRM3	–	AAF27060	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	NTF-RRM4	–	AAF81299	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	NTF-RRM5	–	BAB09056	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	NTF-RRM6	–	BAB10647	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	NTF-RRM7	–	BAB10698	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	NTF-RRM8	–	BAB02073	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	NTF-RRM9	–	AAF20221	2	Lorkovic and Barta (2002)
▷ **P54nrb — 54 kD nuclear RNA-binding protein/ non-POU domain containing, octamer-binding**					
Homo sapiens	p54nrb	–	NP_031389	2	Dong et al. (1993)
Mus musculus	NonO	–	NP_075633	2	Yang et al. (1993)
Rattus norvegicus	NonO/p54nrb	–	NP_001020442	2	Too et al. (1998)
Rattus norvegicus	NonO	–	NP_001012356	2	Brown et al. (2005)
Xenopus laevis	P54nrb	–	NP_001080735	2	Takahashi et al. (2005)
Drosophila melanogaster	NonA	no on or off transient A	NP_523367	2	Jones and Rubin (1990)
▷ **PABP — poly(A)-binding proteins**					
Homo sapiens	PABPC1	–	NP_002559	4	Hosoda et al. (2006)
Homo sapiens	PABP2	–	NP_004634	1	Wahle (1991)
Homo sapiens	PABPC3	–	NP_112241	3	Feral et al. (2001)
Homo sapiens	PABPC4	–	NP_003810	4	Yang et al. (1995)
Homo sapiens	PABPC5	–	NP_543022	4	Blanco et al. (2001)
Mus musculus	PAPB2	–	NP_001007463	1	Cosson et al. (2004)
Mus musculus	PABP1	–	NP_032800	4	Wang et al. (1992)
Mus musculus	PABP2	–	NP_035163	4	Kleene et al. (1998)
Mus musculus	PABPC5	–	NP_444344	4	Blanco et al. (2001)
Rattus norvegicus	PABP1	–	NP_599180	4	Lim et al. (2006)
Drosophila melanogaster	PABPC1	–	NP_725750	4	Behm-Ansmant et al. (2007)
Drosophila melanogaster	PABPN1	–	NP_476902	1	Behm-Ansmant et al. (2007)
Xenopus laevis	PABPC1	–	NP_001080204	4	Voeltz et al. (2001)
Xenopus laevis	ePAB	–	NP_001082094	4	Voeltz et al. (2001)
Bos taurus	PABPN1	–	NP_776994	1	Kuhn et al. (2003)
Arabidopsis thaliana	PABP1	–	NP_174676	3	Belostotsky (2003)
Arabidopsis thaliana	PABP2	–	NP_195137	2	Belostotsky (2003)
Arabidopsis thaliana	PABP3	–	NP_173690	4	Belostotsky (2003)

organism	name	description	accession	#RRM	references
Arabidopsis thaliana	PABP4	–	NP_179916	4	Belostotsky (2003)
Arabidopsis thaliana	PABP5	–	NP_177322	4	Belostotsky (2003)
Arabidopsis thaliana	PABP6	–	NP_188259	4	Belostotsky (2003)
Arabidopsis thaliana	PABP7	–	NP_181204	4	Belostotsky (2003)
Arabidopsis thaliana	PABP8	–	NP_564554	4	Belostotsky (2003)
Arabidopsis thaliana	PABP9	–	NP_175125	4	Belostotsky (2003)
Arabidopsis thaliana	PABN1	–	NP_568751	1	Forbes et al. (2006)
Chlamydomonas reinhardtii	RB57	–	XP_001689671	4	Yohn et al. (1998)
Saccharomyces cerevisiae	Pab1p	–	NP_011092	4	Otero et al. (1999)
Saccharomyces cerevisiae	Rbp29	–	NP_012266	1	Winstall et al. (2000)
Schizosaccharomyces pombe	Pab2	–	CAB16904	1	Perreault et al. (2007)
▷ **PDIP — DNA polymerase delta interacting protein 3**					
Homo sapiens	PDIP3	PDIP46	NP_115687	1	Smyk et al. (2006)
Mus musculus	PDIP3	–	NP_848742	1	Liu et al. (2003)
▷ **PSF — polypyrimidine tract binding protein associated, splicing factor proline/glutamine rich**					
Homo sapiens	PSF	sfpq	NP_005057	2	Gozani et al. (1994)
Mus musculus	PSF	sfpq	NP_076092	2	Shav-Tal et al. (2001)
Danio rerio	PSF	sfpq	NP_998443	2	Lowery et al. (2007)
▷ **PSP1 — paraspeckle protein**					
Homo sapiens	PSP1	paraspeckle protein 1	NP_001035879	2	Fox et al. (2002)
Mus musculus	PSP1	paraspeckle protein 1	NP_079958	2	Myojin et al. (2004)
▷ **PTBP — polypyrimidine tract-binding protein**					
Homo sapiens	PTBP1	–	NP_114368	4	Gil et al. (1991)
Homo sapiens	PTBP2	–	NP_067013	4	Markovtsov et al. (2000)
Mus musculus	PTBP1	–	NP_032982	4	Bothwell et al. (1991)
Mus musculus	PTBP2	brPTB	NP_062423	4	Polydorides et al. (2000)
Rattus norvegicus	PTBP1	PTB4	NP_071961	4	Gooding et al. (2003)
Rattus norvegicus	PTBP	smPTB	NP_877970	4	Gooding et al. (2003)
Rattus norvegicus	PTBP2	nPTB	NP_001005555	4	Gooding et al. (2003)
▷ **RALYL — RALY RNA binding protein and RALY-like**					
Homo sapiens	RALYL	–	NP_776247	1	Ji et al. (2003)
Homo sapiens	RALY	hnRNP-associated with lethal yellow	NP_031393	1	Khrebtukova et al. (1999)

organism	name	description	accession	#RRM	references
Mus musculus	RALY	hnRNP-associated with lethal yellow	NP_075619	1	Khrebtukova et al. (1999)
▷ **RBM — RNA Binding Motif proteins**					
Homo sapiens	RBM1	RNPC1	NP_906270	1	Shu et al. (2006)
Homo sapiens	RBM5	–	NP_005769	2	Shu et al. (2007)
Homo sapiens	RBM6	–	NP_005768	2	Wang et al. (2007b)
Homo sapiens	RBM8A	–	NP_005096	1	Salicioni et al. (2000)
Homo sapiens	RBM8B	–	AAG16782	1	Salicioni et al. (2000)
Homo sapiens	RBM10	–	NP_690595	2	Inoue et al. (1996)
Rattus norvegicus	RBM10	–	NP_690600	2	Inoue et al. (1996)
Homo sapiens	RBM12	–	NP_690051	4	Stover et al. (2001)
Pongo pygmaeus	RBM12	–	Q5RBM8	4	*GenBank*
Macaca mulatta	RBM12	–	Q8SQ27	4	*GenBank*
Mus musculus	RBM12	–	Q8R4X3	4	*GenBank*
Homo sapiens	RBM14	–	NP_006319	2	Iwasaki et al. (2001)
Homo sapiens	RBM15	–	NP_073605	3	Ma et al. (2001)
Mus musculus	RBM15	–	NP_001039272	3	Ma et al. (2007)
Homo sapiens	RBM22	–	NP_060517	1	Montaville et al. (2006)
Homo sapiens	RBM25	–	NP_067062	1	Fortes et al. (2007)
Homo sapiens	RBM28	–	NP_060547	4	Damianov et al. (2006)
Homo sapiens	RBM40	–	Q96LT9	2	Zhao et al. (2003)
Mus musculus	RBM40	–	Q3UZ01	2	Zhao et al. (2003)
Rattus norvegicus	RBM40	–	Q4G055	2	*GenBank*
Homo sapiens	RBM45	–	NP_694453	4	Tamada et al. (2002)
Mus musculus	RBM45	–	NP_700454	4	Tamada et al. (2002)
▷ **RBMS — RNA binding motif, single stranded interacting protein**					
Homo sapiens	RBMS1	scr2/MSSP1	NP_002888	2	Kanaoka and Nojima (1994) ; Negishi et al. (1994)
Homo sapiens	RBMS2	scr3	NP_002889	2	Kanaoka and Nojima (1994)
Homo sapiens	RBMS3	–	NP_001003792	2	Penkov et al. (2000) ; Fritz and Stefanovic (2007)
Mus musculus	RBMS1	–	NP_064692	2	Fujimoto et al. (2000)
Mus musculus	RBMS2	–	NP_062685	2	Fujimoto et al. (2000)
Gallus gallus	RBMS1	–	NP_990355	2	Kimura et al. (1998)
▷ **RBMY — RNA binding motif protein, Y-linked, family 1**					
Homo sapiens	RBMY A1	–	NP_005049	1	Chai et al. (1998)
Homo sapiens	RBMY B	–	NP_001006121	1	Chai et al. (1998)
Homo sapiens	RBMY D	–	NP_001006120	1	Chai et al. (1998)

organism	name	description	accession	#RRM	references
Homo sapiens	RBMY E	–	NP_001006118	1	Chai et al. (1998)
Homo sapiens	RBMY F	–	NP_689798	1	Chai et al. (1998)
Homo sapiens	RBMY J	–	NP_001006117	1	Chai et al. (1998)
▷ **RDM1 — RAD52 motif 1**					
Homo sapiens	RDM1	–	NP_663629	1	Hamimes et al. (2005)
Gallus gallus	RDM1	–	NP_989877	1	Hamimes et al. (2005)
Mus musculus	RDM1	–	NP_079930	1	Hamimes et al. (2005)
▷ **RNPS1 — RNA-binding protein S1**					
Homo sapiens	RNPS1	–	NP_006702	1	Badolato et al. (1995)
Mus musculus	RNPS1	–	NP_033096	1	Schmidt and Werner (1993)
▷ **SART3 — squamous cell carcinoma antigen recognized by T cells 3**					
Homo sapiens	SART 3	–	NP_055521	2	Harada et al. (2001)
Mus musculus	SART 3	–	NP_058622	2	Harada et al. (2000)
Danio rerio	SART 3	–	NP_001025289	2	Trede et al. (2007)
Rattus norvegicus	SART 3	–	NP_001100626	2	*GenBank*
▷ **SR — SR proteins**					
Homo sapiens	SRp20	–	P84103	1	Zahler et al. (1992)
Homo sapiens	SC35	–	NP_003007	1	Fu and Maniatis (1992)
Homo sapiens	SRp30c	–	AAA93069	2	Screaton et al. (1995)
Homo sapiens	9G8	–	NP_001026854	1	Cavaloc et al. (1994)
Homo sapiens	ASF/SF2	–	NP_008855	2	Krainer et al. (1990) ; Ge and Manley (1990)
Homo sapiens	SRp40	–	NP_008856	2	Screaton et al. (1995)
Homo sapiens	SRp46	–	NP_115285	1	Soret et al. (1998)
Homo sapiens	SRp55	–	Q13247	2	Screaton et al. (1995)
Homo sapiens	SRp54	–	NP_004759	1	Zhang and Wu (1996)
Homo sapiens	SRp75	–	Q08170	2	Zahler et al. (1993)
Arabidopsis thaliana	SRp30	–	NP_172386	2	Lopato et al. (1999b)
Arabidopsis thaliana	SR1/SRp34	–	O22315	2	Lazar et al. (1995)
Arabidopsis thaliana	SRp34a	–	NP_190512	2	Lorkovic and Barta (2002)
Arabidopsis thaliana	SRp34b	–	NP_567235	2	Lorkovic and Barta (2002)
Arabidopsis thaliana	RSp31	–	P92964	2	Lopato et al. (1996)
Arabidopsis thaliana	RSp31a	–	NP_182184	2	Kalyna and Barta (2004)
Arabidopsis thaliana	RSp40/SRp35	–	P92965	2	Lopato et al. (1996)

organism	name	description	accession	#RRM	references
Arabidopsis thaliana	RSp41	–	P92966	2	Lopato et al. (1996)
Arabidopsis thaliana	SRZ21/RSZ21	–	AAD12770	1	Golovkin and Reddy (1998)
Arabidopsis thaliana	SRZ22/RSZ22	–	NP_194886	1	Golovkin and Reddy (1998)
Arabidopsis thaliana	RSZ22a	–	NP_180035	1	Lopato et al. (1999a)
Arabidopsis thaliana	RSZ32	–	NP_190918	1	Lopato et al. (2002)
Arabidopsis thaliana	RSZ33	–	CAC03605	1	Lopato et al. (2002)
Arabidopsis thaliana	SC35	–	NP_201225	1	Lopato et al. (2002)
Arabidopsis thaliana	SR33/SCL33	–	NP_564685	1	Golovkin and Reddy (1999)
Arabidopsis thaliana	SCL30	–	NP_567021	1	Lopato et al. (2002)
Arabidopsis thaliana	SCL30a	–	NP_187966	1	Lopato et al. (2002)
Arabidopsis thaliana	SCL28	–	NP_197382	1	Lopato et al. (2002)
Arabidopsis thaliana	SR45	–	Q9SEE9	1	Golovkin and Reddy (1999)
Oryza sativa	SRp32	–	AK061903	2	Isshiki et al. (2006)
Oryza sativa	SRp33a	–	AK106176	2	Isshiki et al. (2006)
Oryza sativa	SRp33b	–	AK071503	2	Isshiki et al. (2006)
Oryza sativa	SR20	–	AK103534	2	Isshiki et al. (2006)
Oryza sativa	SC35a	–	AK103676	1	Isshiki et al. (2006)
Oryza sativa	SC35b	–	AK070292	1	Isshiki et al. (2006)
Oryza sativa	SC35c	–	AK073057	1	Isshiki et al. (2006)
Oryza sativa	SCL25	–	AK073451	1	Isshiki et al. (2006)
Oryza sativa	SCL30a	–	AK062577	1	Isshiki et al. (2006)
Oryza sativa	SCL30b	–	AK065531	1	Isshiki et al. (2006)
Oryza sativa	RSZp21a	–	AK063879	1	Isshiki et al. (2006)
Oryza sativa	RSZp21b	–	AK073210	1	Isshiki et al. (2006)
Oryza sativa	RSZp23	–	AK060088	1	Isshiki et al. (2006)
Oryza sativa	RSp29	–	AK102071	2	Isshiki et al. (2006)
Oryza sativa	RSp33	–	AK121372	2	Isshiki et al. (2006)
Oryza sativa	RSZ36	–	AK103897	1	Isshiki et al. (2006)
Oryza sativa	RSZ37a	–	AK060171	1	Isshiki et al. (2006)
Oryza sativa	RSZ37b	–	AK073348	1	Isshiki et al. (2006)
Oryza sativa	RSZ39	–	AK063518	1	Isshiki et al. (2006)
Zea mays	SRp30	–	AAU29331	2	Gao et al. (2004)
Zea mays	SRp30′	–	Q64HB9	2	Gao et al. (2004)
Zea mays	SRp31	–	Q64HB5	2	Gao et al. (2004)
Zea mays	SRp31"	–	Q64HB8	2	Gao et al. (2004)
Zea mays	SRp32	–	AAU29328	2	Gao et al. (2004)
Zea mays	SRp32′	–	Q64HC2	2	Gao et al. (2004)
Caenorhabditis elegans	CeSRp75/rsp-1	–	Q23121	2	Longman et al. (2000)

organism	name	description	accession	#RRM	references
Caenorhabditis elegans	CeSRp40/rsp-2	–	Q23120	2	Longman et al. (2000)
Caenorhabditis elegans	CeSF2/ASF/rsp-3	–	Q9NEW6	2	Longman et al. (2000)
Caenorhabditis elegans	CeSC35/rsp-4	–	Q09511	1	Longman et al. (2000)
Caenorhabditis elegans	CeSC35-2/rsp-5	–	Q10021	1	Longman et al. (2000)
Caenorhabditis elegans	CeSRp20/rsp-6	–	Q18409	1	Longman et al. (2000)
Caenorhabditis elegans	Cep54/rsp-7	–	O01159	2	Longman et al. (2000)
Drosophila melanogaster	SC35	–	NP_652612	1	Allemand et al. (2001)
Drosophila melanogaster	9G8	–	AAF43414	1	Allemand et al. (2001)
Drosophila melanogaster	SRp55	–	CAA41556	2	Roth et al. (1991)
Schizosaccharomyces pombe	Srp1	–	AAC49909	1	Gross et al. (1998)
Chironomus tentans	Hrp45	–	Z54163	2	Alzhanova-Ericsson et al. (1996)
▷ **SRrp — SR related protein**					
Homo sapiens	SRp40	–	AAL57514	1	Cowper et al. (2001)
Homo sapiens	SRrp35	–	Q8WXF0	1	Cowper et al. (2001)
Homo sapiens	SRrp86	–	NP_001070667	2	Barnard and Patton (2000)
Homo sapiens	SRp38	–	NP_473357	1	Shin et al. (2004)
▷ **TARDBP — TAR DNA binding protein**					
Homo sapiens	TARDBP	–	NP_031401	2	Ou et al. (1995)
Mus musculus	TARDBP	–	NP_663531	2	Buratti and Baralle (2008)
▷ **TIATIAR — T-cell internal antigen-1 (TIA-1) and TIA-1-related protein (TIAR)**					
Homo sapiens	TIAR	–	NP_003243	3	Kawakami et al. (1992)
Homo sapiens	TIA-1	–	NP_071320	3	Tian et al. (1991)
Mus musculus	TIA-1	–	NP_035715	3	Lowin et al. (1996)
Gallus gallus	TIA-1	–	NP_989687	3	Le Guiner et al. (2003)
Bombyx mori	TIA-1	–	NP_001036895	3	Kotani et al. (2003)
Rattus norvegicus	TIA-1	–	NP_001012096	3	*GenBank*
Mus musculus	TiaL1	–	NP_033409	3	Beck et al. (1996)
Rattus norvegicus	TiaL1	–	NP_001013211	3	*GenBank*
▷ **TIF3 — Translation initiation factor 3**					
Arabidopsis thaliana	TIF3	–	AAG51445	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	TIF3	–	BAB10805	1	Lorkovic and Barta (2002)

organism	name	description	accession	#RRM	references
▷ **UMH — Proteins with the UHM domain (RRM variant)**					
Homo sapiens	KIS	kinase interacting stathmin	NP_787062	1	Bieche et al. (2003)
Mus musculus	KIS	kinase interacting stathmin	NP_034763	1	Maucuer et al. (1995)
Rattus norvegicus	KIS	kinase interacting stathmin	NP_058989	1	Maucuer et al. (1997)
Homo sapiens	PUF60	poly-U binding splicing factor 60KDa	NP_510965	3	Page-McCaw et al. (1999)
Drosophila melanogaster	PUF60	poly-U binding splicing factor 60KDa	NP_525123	2	Van Buskirk and Schupbach (2002)
Mus musculus	PUF60	poly-U binding splicing factor 60KDa	AAH10601	2	Kielkopf et al. (2004)
Danio rerio	PUF60	poly-U binding splicing factor 60KDa	CAD61099	2	Kielkopf et al. (2004)
Anopheles gambiae	PUF60	poly-U binding splicing factor 60KDa	EAA09944	2	Kielkopf et al. (2004)
Caenorhabditis elegans	PUF60	poly-U binding splicing factor 60KDa	AAF60676	2	Kielkopf et al. (2004)
Homo sapiens	SPF45	RBM17	NP_116294	1	Sampath et al. (2003)
Mus musculus	SPF45	RBM17	NP_690037	1	Kielkopf et al. (2004)
Danio rerio	SPF45	–	AAH45473	1	Kielkopf et al. (2004)
Anopheles gambiae	SPF45	–	EAA15153	1	Kielkopf et al. (2004)
Drosophila melanogaster	SPF45	–	EAA46008	1	Chaouki and Salz (2006)
Caenorhabditis elegans	SPF45	–	CAA97799	1	Kielkopf et al. (2004)
Homo sapiens	Tat-SF1	–	AAB18823	1	Zhou and Sharp (1996)
Mus musculus	Tat-SF1	–	AAH37711	1	Kielkopf et al. (2004)
Danio rerio	Tat-SF1	–	AAH55565	1	Kielkopf et al. (2004)
Drosophila melanogaster	Tat-SF1	–	AAF51719	1	Kielkopf et al. (2004)
Anopheles gambiae	Tat-SF1	–	EAA10753	1	Kielkopf et al. (2004)
Caenorhabditis elegans	Tat-SF1	–	AAK29956	1	Kielkopf et al. (2004)
Homo sapiens	HCC1	Hepatocellular carcinoma protein 1	Q14498	2	Imai et al. (1993)
Mus musculus	HCC1	–	Q8VH51	2	Kielkopf et al. (2004)
Danio rerio	HCC1	–	AAH44487	2	Kielkopf et al. (2004)
Drosophila melanogaster	HCC1	–	AAF52478	2	Kielkopf et al. (2004)
Anopheles gambiae	HCC1	–	EAA12873	2	Kielkopf et al. (2004)
Caenorhabditis elegans	HCC1	–	AAM97977	2	Kielkopf et al. (2004)
Arabidopsis thaliana	HCC1	–	AAM20703	2	Kielkopf et al. (2004)
▷ **ZCRB1 — zinc finger CCHC-type and RNA binding motif**					
Homo sapiens	ZCRB1	–	NP_149105	1	Wang et al. (2007a)
Mus musculus	ZCRB1	–	NP_080301	1	Wang et al. (2007a)
▷ **cpRNP — chloroplast RRM containing proteins**					
Arabidopsis thaliana	cpRNP28	–	AAD14476	2	Lorkovic and Barta (2002)
Arabidopsis thaliana	cpRNP28b	–	AAF26472	2	Lorkovic and Barta (2002)
Arabidopsis thaliana	cpRNP29	–	CAB67653	2	Ohta et al. (1995)

organism	name	description	accession	#RRM	references
Arabidopsis thaliana	cpRNP29b	–	AAC98043	2	Lorkovic and Barta (2002)
Arabidopsis thaliana	cpRNP31	–	CAA22986	2	Ohta et al. (1995)
Arabidopsis thaliana	cpRNP31b	–	BAB09396	2	Lorkovic and Barta (2002)
Arabidopsis thaliana	cpRNP33	–	CAB43448	2	Ohta et al. (1995)
Arabidopsis thaliana	cpRNP33b	–	AAC36180	2	Lorkovic and Barta (2002)
▷ **cyclophilin — cyclophilins**					
Arabidopsis thaliana	AtCyp59	–	NP_175776	1	Gullerova et al. (2006)
Rattus norvegicus	CypE	–	NP_001041333	1	*GenBank*
Caenorhabditis elegans	cyp-13	–	NP_503034	1	Zorio and Blumenthal (1999)
Mus musculus	Cyp-33	–	Q9QZH3	1	Anderson et al. (2002)
Drosophila melanogaster	Cyp-33	–	Q9V3G3	1	Anderson et al. (2002)
Homo sapiens	Cyp-33	–	Q9UNP9	1	Mi et al. (1996)
Homo sapiens	PPIL4	–	Q8WUA2	1	Zeng et al. (2001)
Mus musculus	PPIL4	–	Q9CXG3	1	Zeng et al. (2001)
Schizosaccharomyces pombe	PPIL4	–	Q9UUE4	1	Pemberton and Kay (2005)
Haemonchus contortus	CYP	–	AAW82658	1	Valle et al. (2005)
▷ **eIF4B — eukaryotic translation initiation factor 4B**					
Homo sapiens	eIF-4B	–	NP_001408	1	Milburn et al. (1990)
Mus musculus	eIF-4B	–	NP_663600	1	Shahbazian et al. (2006)
▷ **hnRNP — Heterogeneous nuclear ribonucleoprotein**					
Arabidopsis thaliana	AtRNP A/B1	–	CAB10209	2	Lorkovic et al. (2000b)
Arabidopsis thaliana	AtRNP A/B2	–	AAB80680	2	Lorkovic et al. (2000b)
Arabidopsis thaliana	AtRNP A/B3	–	BAB08572	2	Lorkovic et al. (2000b)
Arabidopsis thaliana	AtRNP A/B4	–	CAB43861	2	Lorkovic et al. (2000b)
Arabidopsis thaliana	AtRNP A/B5	–	BAB09088	2	Lorkovic et al. (2000b)
Arabidopsis thaliana	AtRNP A/B6	–	AAF21191	2	Lorkovic et al. (2000b)
Arabidopsis thaliana	AtRNP H/F1	–	BAB10406	2	Lorkovic et al. (2000b)
Arabidopsis thaliana	AtRNP H/F2	–	BAB02497	2	Lorkovic et al. (2000b)
Arabidopsis thaliana	AtRNP I/PTB	–	BAB08421	2	Lorkovic et al. (2000b)
Arabidopsis thaliana	AtRNP I/PTB	–	AAF26159	2	Lorkovic and Barta (2002)
Arabidopsis thaliana	AtRNP I/PTB	–	NP_175010	4	Lorkovic and Barta (2002)
Homo sapiens	hnRNP A2/B1	–	NP_112533	2	Burd et al. (1989)
Homo sapiens	hnRNP A1/A1	–	NP_002127	2	Buvoli et al. (1988)
Homo sapiens	hnRNP A/B	–	NP_112556	2	Khan et al. (1991)
Homo sapiens	hnRNP A0	–	NP_006796	2	Myer and Steitz (1995)

organism	name	description	accession	#RRM	references
Homo sapiens	hnRNP C1/C2	–	AAH07052	1	Swanson et al. (1987)
Homo sapiens	hnRNP C	–	NP_001070910	1	Swanson et al. (1987)
Homo sapiens	hnRNP D	–	NP_112737	2	Park et al. (2007)
Homo sapiens	hnRNP F	–	NP_001091676	3	Matunis et al. (1994)
Homo sapiens	hnRNP G	–	P38159	1	Soulard et al. (1993)
Homo sapiens	hnRNP H3	–	NP_036339	2	Mahe et al. (1997)
Homo sapiens	hnRNP L	–	NP_001524	4	Pinol-Roma et al. (1989)
Homo sapiens	hnRNP L-like	–	NP_612403	3	Shur et al. (2004)
Homo sapiens	hnRNP R	–	NP_005817	3	Hassfeld et al. (1998)
Rattus norvegicus	hnRNP A2/B1	–	NP_001098083	2	Kamma et al. (1999)
Rattus norvegicus	hnRNP M	–	NP_001103381	3	Kafasla et al. (2000)
Chironomus tentans	Hrp59 protein	–	CAH05070	3	Kiesler et al. (2005)
Mus musculus	HNRP A/B	–	NP_034578	2	Inoue et al. (2001)
Mus musculus	HNRP L	–	NP_796275	3	Costain and Mishra (2003)
Mus musculus	HNRP M	–	NP_084080	3	Mahe et al. (2000)
Drosophila melanogaster	Hrp36	–	CAA44502	2	Matunis et al. (1992)
▷ **nucleolin — nucleolins**					
Arabidopsis thaliana	nucleolin	AtNUC-L1	NP_175322	2	Pontvianne et al. (2007)
Arabidopsis thaliana	nucleolin	AtNUC-L2	NP_188491	2	Pontvianne et al. (2007)
Homo sapiens	nucleolin/C23	–	NP_005372	4	Srivastava et al. (1989)
Rattus norvegicus	nucleolin/C23	–	NP_036881	4	Bourbon et al. (1988b)
Schizosaccharomyces pombe	GAR2	–	NP_593531	2	Gulli et al. (1995)
Saccharomyces cerevisiae	NSR1	–	NP_011675	2	Edwards et al. (2000)
Xenopus laevis	NCL	–	P20397	4	Heine et al. (1993)
Mus musculus	NCL	–	NP_035010	4	Bourbon et al. (1988a)
▷ **oligoU — oligouridylated-specific RRM proteins (UBP1, RBP45, RBP47, UBA1 et UBA2)**					
Arabidopsis thaliana	RBP45b	–	NP_172630	3	Lorkovic et al. (2000a)
Arabidopsis thaliana	RBP47a	–	AAG13046	3	Lorkovic et al. (2000a)
Arabidopsis thaliana	RBP47b	–	BAB02953	3	Lorkovic et al. (2000a)
Arabidopsis thaliana	RBP47c	–	AAD46038	3	Lorkovic et al. (2000a)
Arabidopsis thaliana	RBP47c′	–	AAD46037	3	Lorkovic et al. (2000a)
Nicotiana plumbaginifolia	RBP45	–	CAC01237	3	Lorkovic et al. (2000a)
Arabidopsis thaliana	UBP1a	–	AAD25780	3	Lambermon et al. (2000)
Arabidopsis thaliana	UBP1b	–	AAF79492	3	Lambermon et al. (2000)
Arabidopsis thaliana	UBP1c	–	BAB02974	3	Lambermon et al. (2000)
Nicotiana plumbaginifolia	UBP1	–	CAB75429	3	Lambermon et al. (2000)

organism	name	description	accession	#RRM	references
Arabidopsis thaliana	UBA2a	–	NP_850711	2	Lorkovic and Barta (2002)
Arabidopsis thaliana	UBA2b	–	NP_001078035	2	Lorkovic and Barta (2002)
Arabidopsis thaliana	UBA2c	–	BAA97064	2	Lorkovic and Barta (2002)
Arabidopsis thaliana	UBA1a	–	AAD25815	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	UBA1b	–	AAD25814	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	UBA1c	–	AAC16468	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	S-RBP2	–	BAB09544	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	S-RBP3	–	AAC61284	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	S-RBP4	–	CAB88326	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	S-RBP5	–	CAB78306	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	S-RBP7	–	AAG51218	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	S-RBP8	–	AAD20390	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	S-RBP9	–	AAC23648	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	S-RBP10	–	AAF21210	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	S-RBP11	–	BAB09686	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	S-RBP12	–	BAB09337	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	S-RBP13	–	BAB08354	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	S-RBP14	–	AAB88654	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	S-RBP15	–	BAB09540	1	Lorkovic and Barta (2002)
▷ **snRNP — small nuclear ribonucleoprotein (snRNP)**					
Solanum tuberosum	U1A	–	CAA90282	2	Simpson et al. (1995)
Homo sapiens	U1A	–	NP_004587	2	Sillekens et al. (1987)
Mus musculus	U1A	–	NP_056597	2	Bennett et al. (1993)
Xenopus laevis	U1A	–	CAA41021	2	Scherly et al. (1991)
Solanum tuberosum	U2B″	–	AAA33847	2	Simpson et al. (1991)
Arabidopsis thaliana	U1A	–	NP_182280	2	Simpson et al. (1995)
Arabidopsis thaliana	U170K	–	CAB62436	1	Golovkin and Reddy (1996)
Arabidopsis thaliana	U2B	–	NP_180585	2	Simpson et al. (1995)
Arabidopsis thaliana	U2B	–	AAF82223	2	Simpson et al. (1991)
Arabidopsis thaliana	U2AF65a	–	CAB80335	2	Domon et al. (1998)
Arabidopsis thaliana	U2AF65b	–	AAG51641	3	Domon et al. (1998)
Arabidopsis thaliana	U2SF3b53a	–	AAD12222	2	Lorkovic and Barta (2002)
Arabidopsis thaliana	U2SF3b53b	–	AAD15475	2	Lorkovic and Barta (2002)
Arabidopsis thaliana	U11-35kD	–	AAB64330	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	U2AF35	–	BAB10638	1	Domon et al. (1998)
Nicotiana plumbaginifolia	U2AF65	–	CAA77136	3	Domon et al. (1998)
Zea mays	U170K	–	AAT64945	1	Gupta et al. (2006)

organism	name	description	accession	#RRM	references
Danio rerio	U170K	–	Q6DRE8	1	Amsterdam et al. (2004)
Mus musculus	U11/U1235K	–	Q9D384	1	*GenBank*
Homo sapiens	U2AF35	–	NP_006749	1	Zhang et al. (1992)
Mus musculus	U2AF35	–	Q9D883	1	Pacheco et al. (2004)
Danio rerio	U2AF35	–	AAM34646	1	Pacheco et al. (2004)
Drosophila melanogaster	U2AF35	–	AAB17271	1	Rudner et al. (1996)
Caenorhabditis elegans	U2AF35	–	CAB55137	1	Zorio and Blumenthal (1999)
Gallus gallus	U2AF35	–	NP_989986	1	Pacheco et al. (2004)
Schizosaccharomyces pombe	U2AF23	–	AAC49805	1	Wentz-Hunter and Potashkin (1996)
Homo sapiens	U2AF65	–	NP_009210	3	Zamore et al. (1992)
Mus musculus	U2AF65	–	NP_598432	3	Sailer et al. (1992)
Danio rerio	U2AF65	–	AAH65869	3	Kielkopf et al. (2004)
Caenorhabditis elegans	U2AF65	–	AAC26982	3	Zorio et al. (1997)
Homo sapiens	U11/U12 35K	–	NP_073208	1	Will et al. (2004)
Mus musculus	U2AF1-rs 1	–	NP_035793	1	Kitagawa et al. (1995)
Homo sapiens	U2A1L3	–	Q8WU68	1	Shepard et al. (2002)
Homo sapiens	U2AF1L4	–	NP_001035515	1	Shepard et al. (2002)
▷ **30kRRM — 30K-RRM proteins**					
Arabidopsis thaliana	30K-RRM1	–	CAB77597	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	30K-RRM2	–	AAF27001	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	30K-RRM3	–	AAD30604	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	30K-RRM4	–	AAG51948	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	30K-RRM5	–	AAF87259	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	30K-RRM6	–	AAF71806	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	30K-RRM7	–	AAC33496	1	Lorkovic and Barta (2002)
Arabidopsis thaliana	30K-RRM8	–	AAK32921	1	Lorkovic and Barta (2002)

References

R. Abe, Y. Uyeno, K. Yamamoto, and H. Sakamoto. Tissue-specific expression of the gene encoding a mouse RNA binding protein homologous to human HuD antigen. *DNA Res*, 1:175–80, 1994.

R. Abe, E. Sakashita, K. Yamamoto, and H. Sakamoto. Two different RNA binding activities for the AU-rich element and the poly(A) sequence of the mouse neuronal protein mHuC. *Nucleic Acids Res*, 24:4895–901, 1996a.

R. Abe, K. Yamamoto, and H. Sakamoto. Target specificity of neuronal RNA-binding protein, Mel-N1: direct binding to the 3′ untranslated region of its own mRNA. *Nucleic Acids Res*, 24:2011–6, 1996b.

E. Allemand, R. Gattoni, H. M. Bourbon, J. Stevenin, J. F. Caceres, J. Soret, and J. Tazi. Distinctive features of Drosophila alternative splicing factor RS domain: implication for specific phosphorylation, shuttling, and splicing activation. *Mol Cell Biol*, 21:1345–59, 2001.

A. T. Alzhanova-Ericsson, X. Sun, N. Visa, E. Kiseleva, T. Wurtz, and B. Daneholt. A protein of the SR family of splicing factors binds extensively to exonic Balbiani ring pre-mRNA and accompanies the RNA from the gene to the nuclear pore. *Genes Dev*, 10:2881–93, 1996.

A. Amsterdam, R. M. Nissen, Z. Sun, E. C. Swindell, S. Farrington, and N. Hopkins. Identification of 315 genes essential for early zebrafish development. *Proc Natl Acad Sci U S A*, 101:12792–7, 2004.

M. Anderson, K. Fair, S. Amero, S. Nelson, P. J. Harte, and M. O. Diaz. A new family of cyclophilins with an RNA recognition motif that interact with members of the trx/MLL protein family in Drosophila and human cells. *Dev Genes Evol*, 212:107–13, 2002.

U. Atasoy, J. Watson, D. Patel, and J. D. Keene. ELAV protein HuA (HuR) can redistribute between nucleus and cytoplasm and is upregulated during serum stimulation and T cell activation. *J Cell Sci*, 111 (Pt 21): 3145–56, 1998.

J. Badolato, E. Gardiner, N. Morrison, and J. Eisman. Identification and characterisation of a novel human RNA-binding protein. *Gene*, 166:323–7, 1995.

A. Barbouti, M. Hoglund, B. Johansson, C. Lassen, P. G. Nilsson, A. Hagemeijer, F. Mitelman, and T. Fioretos. A novel gene, MSI2, encoding a putative RNA-binding protein is recurrently rearranged at disease progression of chronic myeloid leukemia and forms a fusion gene with HOXA9 as a result of the cryptic t(7;17)(p15;q23). *Cancer Res*, 63:1202–6, 2003.

D. C. Barnard and J. G. Patton. Identification and characterization of a novel serine-arginine-rich splicing regulatory protein. *Mol Cell Biol*, 20:3049–57, 2000.

A. R. Beck, Q. G. Medley, S. O'Brien, P. Anderson, and M. Streuli. Structure, tissue distribution and genomic organization of the murine RRM-type RNA binding proteins TIA-1 and TIAR. *Nucleic Acids Res*, 24:3829–35, 1996.

I. Behm-Ansmant, D. Gatfield, J. Rehwinkel, V. Hilgers, and E. Izaurralde. A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. *Embo J*, 26:1591–601, 2007.

D. A. Belostotsky. Unexpected complexity of poly(A)-binding protein gene families in flowering plants: three conserved lineages that are at least 200 million years old and possible auto- and cross-regulation. *Genetics*, 163: 311–9, 2003.

M. M. Bennett, M. A. Baron, and J. Craft. Nucleotide sequence analysis of the A protein of the U1 small nuclear ribonucleoprotein particle: the murine protein contains a 5′ amino-terminal tag. *Nucleic Acids Res*, 21:4404, 1993.

I. Bieche, V. Manceau, P. A. Curmi, I. Laurendeau, S. Lachkar, K. Leroy, D. Vidaud, A. Sobel, and A. Maucuer. Quantitative RT-PCR reveals a ubiquitous but preferentially neural expression of the KIS gene in rat and human. *Brain Res Mol Brain Res*, 114:55–64, 2003.

P. Blanco, C. A. Sargent, C. A. Boucher, G. Howell, M. Ross, and N. A. Affara. A novel poly(A)-binding protein gene (PABPC5) maps to an X-specific subinterval in the Xq21.3/Yp11.2 homology block of the human sex chromosomes. *Genomics*, 74:1–11, 2001.

A. L. Bothwell, D. W. Ballard, W. M. Philbrick, G. Lindwall, S. E. Maher, M. M. Bridgett, S. F. Jamison, and M. A. Garcia-Blanco. Murine polypyrimidine tract binding protein. Purification, cloning, and mapping of the RNA binding domain. *J Biol Chem*, 266:24657–63, 1991.

H. M. Bourbon, B. Lapeyre, and F. Amalric. Structure of the mouse nucleolin gene. The complete sequence reveals that each RNA binding domain is encoded by two independent exons. *J Mol Biol*, 200:627–38, 1988a.

H. M. Bourbon, M. Prudhomme, and F. Amalric. Sequence and structure of the nucleolin promoter in rodents: characterization of a strikingly conserved CpG island. *Gene*, 68:73–84, 1988b.

K. R. Brimacombe and A. N. Ladd. Cloning and embryonic expression patterns of the chicken CELF family. *Dev Dyn*, 236:2216–24, 2007.

S. A. Brown, J. Ripperger, S. Kadener, F. Fleury-Olela, F. Vilbois, M. Rosbash, and U. Schibler. PERIOD1-associated proteins modulate the negative limb of the mammalian circadian oscillator. *Science*, 308:693–6, 2005.

E. Buratti and F. E. Baralle. Multiple roles of TDP-43 in gene expression, splicing regulation, and human disease. *Front Biosci*, 13:867–78, 2008.

C. G. Burd, M. S. Swanson, M. Gorlach, and G. Dreyfuss. Primary structures of the heterogeneous nuclear ribonucleoprotein A2, B1, and C2 proteins: a diversity of RNA binding proteins is generated by small peptide inserts. *Proc Natl Acad Sci U S A*, 86:9788–92, 1989.

M. Buvoli, G. Biamonti, P. Tsoulfas, M. T. Bassi, A. Ghetti, S. Riva, and C. Morandi. cDNA cloning of human hnRNP protein A1 reveals the existence of multiple mRNA isoforms. *Nucleic Acids Res*, 16:3751–70, 1988.

C. D. Carpenter, J. A. Kreps, and A. E. Simon. Genes encoding glycine-rich Arabidopsis thaliana proteins with RNA-binding motifs are influenced by cold treatment and an endogenous circadian rhythm. *Plant Physiol*, 104:1015–25, 1994.

Y. Cavaloc, M. Popielarz, J. P. Fuchs, R. Gattoni, and J. Stevenin. Characterization and cloning of the human splicing factor 9G8: a novel 35 kDa factor of the serine/arginine protein family. *Embo J*, 13:2639–49, 1994.

N. N. Chai, H. Zhou, J. Hernandez, H. Najmabadi, S. Bhasin, and P. H. Yen. Structure and organization of the RBMY genes on the human Y chromosome: transposition and amplification of an ancestral autosomal hnRNPG gene. *Genomics*, 49:283–9, 1998.

A. S. Chaouki and H. K. Salz. Drosophila SPF45: a bifunctional protein with roles in both splicing and DNA repair. *PLoS Genet*, 2:e178, 2006.

B. Cosson, F. Braun, L. Paillard, P. Blackshear, and H. Beverley Osborne. Identification of a novel Xenopus laevis poly (A) binding protein. *Biol Cell*, 96:519–27, 2004.

W. J. Costain and R. K. Mishra. PLG regulates hnRNP-L expression in the rat striatum and pre-frontal cortex: identification by ddPCR. *Peptides*, 24:137–46, 2003.

A. E. Cowper, J. F. Caceres, A. Mayeda, and G. R. Screaton. Serine-arginine (SR) protein-like factors that antagonize authentic SR proteins and regulate alternative splicing. *J Biol Chem*, 276:48908–14, 2001.

T. Dai, Y. Vera, E. C. Salido, and P. H. Yen. Characterization of the mouse Dazap1 gene encoding an RNA-binding protein that interacts with infertility factors DAZ and DAZL. *BMC Genomics*, 2:6, 2001.

A. Damianov, M. Kann, W. S. Lane, and A. Bindereif. Human RBM28 protein is a specific nucleolar component of the spliceosomal snRNPs. *Biol Chem*, 387: 1455–60, 2006.

S. Danno, H. Nishiyama, H. Higashitsuji, H. Yokoi, J. H. Xue, K. Itoh, T. Matsuda, and J. Fujita. Increased transcript level of RBM3, a member of the glycine-rich RNA-binding protein family, in human cells in response to cold stress. *Biochem Biophys Res Commun*, 236:804–7, 1997.

L. Didierjean, P. Frendo, and G. Burkard. Stress responses in maize: sequence analysis of cDNAs encoding glycine-rich proteins. *Plant Mol Biol*, 18:847–9, 1992.

C. Domon, Z. J. Lorkovic, J. Valcarcel, and W. Filipowicz. Multiple forms of the U2 small nuclear ribonucleoprotein auxiliary factor U2AF subunits expressed in higher plants. *J Biol Chem*, 273:34603–10, 1998.

B. Dong, D. S. Horowitz, R. Kobayashi, and A. R. Krainer. Purification and cDNA cloning of HeLa cell p54nrb, a nuclear protein with two RNA recognition motifs and extensive homology to human splicing factor PSF and Drosophila NONA/BJ6. *Nucleic Acids Res*, 21:4085–92, 1993.

T. K. Edwards, A. Saleem, J. A. Shaman, T. Dennis, C. Gerigk, E. Oliveros, M. R. Gartenberg, and E. H. Rubin. Role for nucleolin/Nsr1 in the cellular localization of topoisomerase I. *J Biol Chem*, 275:36181–8, 2000.

C. Feral, G. Guellaen, and A. Pawlak. Human testis expresses a specific poly(A)-binding protein. *Nucleic Acids Res*, 29:1872–83, 2001.

S. Fleurdepine, J. M. Deragon, M. Devic, J. Guilleminot, and C. Bousquet-Antonelli. A bona fide La protein is required for embryogenesis in Arabidopsis thaliana. *Nucleic Acids Res*, 35:3306–21, 2007.

K. P. Forbes, B. Addepalli, and A. G. Hunt. An Arabidopsis Fip1 homolog interacts with RNA and provides conceptual links with a number of other polyadenylation factor subunits. *J Biol Chem*, 281:176–86, 2006.

M. Fornaro, S. Raimondo, J. M. Lee, and M. G. Giacobini-Robecchi. Neuron-specific Hu proteins subcellular localization in primary sensory neurons. *Ann Anat*, 189:223–8, 2007.

P. Fortes, D. Longman, S. McCracken, J. Y. Ip, R. Poot, I. W. Mattaj, J. F. Caceres, and B. J. Blencowe. Identification and characterization of RED120: a conserved PWI domain protein with links to splicing and 3'-end formation. *FEBS Lett*, 581:3087–97, 2007.

A. H. Fox, Y. W. Lam, A. K. Leung, C. E. Lyon, J. Andersen, M. Mann, and A. I. Lamond. Paraspeckles: a novel nuclear domain. *Curr Biol*, 12:13–25, 2002.

D. Fritz and B. Stefanovic. RNA-binding protein RBMS3 is expressed in activated hepatic stellate cells and liver

fibrosis and increases expression of transcription factor Prx1. *J Mol Biol*, 371:585–95, 2007.

X. D. Fu and T. Maniatis. Isolation of a complementary DNA that encodes the mammalian splicing factor SC35. *Science*, 256:535–8, 1992.

M. Fujimoto, K. I. Matsumoto, S. M. Iguchi-Ariga, and H. Ariga. Structures and comparison of genomic and complementary DNAs of mouse MSSP, a c-Myc binding protein. *Int J Oncol*, 16:245–51, 2000.

H. Gao, W. J. Gordon-Kamm, and L. A. Lyznik. ASF/SF2-like maize pre-mRNA splicing factors affect splice site utilization and their transcripts are alternatively spliced. *Gene*, 339:25–37, 2004.

H. Ge and J. L. Manley. A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. *Cell*, 62:25–34, 1990.

F. Gebauer and J. D. Richter. Mouse cytoplasmic polyadenylylation element binding protein: an evolutionarily conserved protein that interacts with the cytoplasmic polyadenylylation elements of c-mos mRNA. *Proc Natl Acad Sci U S A*, 93:14602–7, 1996.

A. Gil, P. A. Sharp, S. F. Jamison, and M. A. Garcia-Blanco. Characterization of cDNAs encoding the polypyrimidine tract-binding protein. *Genes Dev*, 5: 1224–36, 1991.

M. Golovkin and A. S. Reddy. An SC35-like protein and a novel serine/arginine-rich protein interact with Arabidopsis U1-70K protein. *J Biol Chem*, 274:36428–38, 1999.

M. Golovkin and A. S. Reddy. Structure and expression of a plant U1 snRNP 70K gene: alternative splicing of U1 snRNP 70K pre-mRNAs produces two different transcripts. *Plant Cell*, 8:1421–35, 1996.

M. Golovkin and A. S. Reddy. The plant U1 small nuclear ribonucleoprotein particle 70K protein interacts with two novel serine/arginine-rich proteins. *Plant Cell*, 10:1637–48, 1998.

P. Good, A. Yoda, S. Sakakibara, A. Yamamoto, T. Imai, H. Sawa, T. Ikeuchi, S. Tsuji, H. Satoh, and H. Okano. The human Musashi homolog 1 (MSI1) gene encoding the homologue of Musashi/Nrp-1, a neural RNA-binding protein putatively expressed in CNS stem cells and neural progenitor cells. *Genomics*, 52:382–4, 1998.

C. Gooding, P. Kemp, and C. W. Smith. A novel polypyrimidine tract-binding protein paralog expressed in smooth muscle cells. *J Biol Chem*, 278:15201–7, 2003.

O. Gozani, J. G. Patton, and R. Reed. A novel set of spliceosome-associated proteins and the essential splicing factor PSF bind stably to pre-mRNA prior to catalytic step II of the splicing reaction. *Embo J*, 13:3356–67, 1994.

T. Gross, K. Richert, C. Mierke, M. Lutzelberger, and N. F. Kaufer. Identification and characterization of srp1, a gene of fission yeast encoding a RNA binding domain and a RS domain typical of SR splicing factors. *Nucleic Acids Res*, 26:505–11, 1998.

M. Gullerova, A. Barta, and Z. J. Lorkovic. AtCyp59 is a multidomain cyclophilin from Arabidopsis thaliana that interacts with SR proteins and the C-terminal domain of the RNA polymerase II. *Rna*, 12:631–43, 2006.

M. P. Gulli, J. P. Girard, D. Zabetakis, B. Lapeyre, T. Melese, and M. Caizergues-Ferrer. gar2 is a nucleolar protein from Schizosaccharomyces pombe required for 18S rRNA and 40S ribosomal subunit accumulation. *Nucleic Acids Res*, 23:1912–8, 1995.

S. Gupta, A. Ciungu, N. Jameson, and S. K. Lal. Alternative splicing expression of U1 snRNP 70K gene is evolutionary conserved between different plant species. *DNA Seq*, 17:254–61, 2006.

S. Hagele, U. Kuhn, M. Boning, and D. M. Katschinski. Cytoplasmic polyadenylation element binding protein (CPEB)1 and 2 bind to the HIF-1alpha mRNA 3'UTR and modulate HIF-1 alpha protein expression. *Biochem J*, 2008.

S. Hamimes, H. Arakawa, A. Z. Stasiak, A. M. Kierzek, S. Hirano, Y. G. Yang, M. Takata, A. Stasiak, J. M. Buerstedde, and E. Van Dyck. RDM1, a novel RNA recognition motif (RRM)-containing protein involved in the cell response to cisplatin in vertebrates. *J Biol Chem*, 280: 9225–35, 2005.

S. Hanano, M. Sugita, and M. Sugiura. Isolation of a novel RNA-binding protein and its association with a large ribonucleoprotein particle present in the nucleoplasm of tobacco cells. *Plant Mol Biol*, 31:57–68, 1996.

K. Harada, A. Yamada, T. Mine, N. Kawagoe, H. Takasu, and K. Itoh. Mouse homologue of the human SART3 gene encoding tumor-rejection antigen. *Jpn J Cancer Res*, 91:239–47, 2000.

K. Harada, A. Yamada, D. Yang, K. Itoh, and S. Shichijo. Binding of a SART3 tumor-rejection antigen to a pre-mRNA splicing factor RNPS1: a possible regulation of splicing by a complex formation. *Int J Cancer*, 93:623–8, 2001.

Y. Hashimoto, H. Suzuki, Y. Kageyama, K. Yasuda, and K. Inoue. Bruno-like protein is localized to zebrafish germ plasm during the early cleavage stages. *Gene Expr Patterns*, 6:201–5, 2006.

W. Hassfeld, E. K. Chan, D. A. Mathison, D. Portman, G. Dreyfuss, G. Steiner, and E. M. Tan. Molecular definition of heterogeneous nuclear ribonucleoprotein R (hnRNP R) using autoimmune antibody: immunological relationship with hnRNP P. *Nucleic Acids Res*, 26:439–45, 1998.

M. A. Heine, M. L. Rankin, and P. J. DiMario. The Gly/Arg-rich (GAR) domain of Xenopus nucleolin facilitates in vitro nucleic acid binding and in vivo nucleolar localization. *Mol Biol Cell*, 4:1189–204, 1993.

C. Heintzen, M. Nater, K. Apel, and D. Staiger. AtGRP7, a nuclear RNA-binding protein as a component of a circadian-regulated negative feedback loop in Arabidopsis thaliana. *Proc Natl Acad Sci U S A*, 94:8515–20, 1997.

J. Hodgkin, J. D. Zellan, and D. G. Albertson. Identification of a candidate primary sex determination locus, fox-1, on the X chromosome of Caenorhabditis elegans. *Development*, 120:3681–9, 1994.

N. Hosoda, F. Lejeune, and L. E. Maquat. Evidence that poly(A) binding protein C1 binds nuclear pre-mRNA poly(A) tails. *Mol Cell Biol*, 26:3085–97, 2006.

H. Imai, E. K. Chan, K. Kiyosawa, X. D. Fu, and E. M. Tan. Novel nuclear autoantigen with splicing factor motifs identified with antibody from hepatocellular carcinoma. *J Clin Invest*, 92:2419–26, 1993.

A. Inoue, K. P. Takahashi, M. Kimura, T. Watanabe, and S. Morisawa. Molecular cloning of a RNA binding protein, S1-1. *Nucleic Acids Res*, 24:2990–7, 1996.

A. Inoue, A. Omori, S. Ichinose, K. P. Takahashi, Y. Kinoshita, and S. Mita. S1 proteins C2 and D2 are novel hnRNPs similar to the transcriptional repressor, CArG box motif-binding factor A. *Eur J Biochem*, 268:3654–63, 2001.

M. Isshiki, A. Tsumoto, and K. Shimamoto. The serine/arginine-rich protein family in rice plays important roles in constitutive and alternative splicing of pre-mRNA. *Plant Cell*, 18:146–58, 2006.

T. Iwasaki, W. W. Chin, and L. Ko. Identification and characterization of RRM-containing coactivator activator (CoAA) as TRBP-interacting protein, and its splice variant as a coactivator modulator (CoAM). *J Biol Chem*, 276: 33375–83, 2001.

C. N. Ji, J. Z. Chen, Y. Xie, S. Wang, J. Qian, E. Zhao, W. Jin, X. Z. Wu, W. X. Xu, K. Ying, and Y. M. Mao. A novel cDNA encodes a putative hRALY-like protein, hRALYL. *Mol Biol Rep*, 30:61–7, 2003.

Y. Jin, H. Suzuki, S. Maegawa, H. Endo, S. Sugano, K. Hashimoto, K. Yasuda, and K. Inoue. A vertebrate RNA-binding protein Fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG. *Embo J*, 22:905–12, 2003.

K. R. Jones and G. M. Rubin. Molecular analysis of no-on-transient A, a gene required for normal vision in Drosophila. *Neuron*, 4:711–23, 1990.

D. J. Jung, S. Y. Na, D. S. Na, and J. W. Lee. Molecular cloning and characterization of CAPER, a novel coactivator of activating protein-1 and estrogen receptors. *J Biol Chem*, 277:1229–34, 2002.

P. Kafasla, M. Patrinou-Georgoula, and A. Guialis. The 72/74-kDa polypeptides of the 70-110 S large heterogeneous nuclear ribonucleoprotein complex (LH-nRNP) represent a discrete subset of the hnRNP M protein family. *Biochem J*, 350 Pt 2:495–503, 2000.

M. Kalyna and A. Barta. A plethora of plant serine/arginine-rich proteins: redundancy or evolution of novel gene functions? *Biochem Soc Trans*, 32:561–4, 2004.

H. Kamma, H. Horiguchi, L. Wan, M. Matsui, M. Fujiwara, M. Fujimoto, T. Yazawa, and G. Dreyfuss. Molecular characterization of the hnRNP A2/B1 proteins: tissue-specific expression and novel isoforms. *Exp Cell Res*, 246:399–411, 1999.

Y. Kanaoka and H. Nojima. SCR: novel human suppressors of cdc2/cdc13 mutants of Schizosaccharomyces pombe harbour motifs for RNA binding proteins. *Nucleic Acids Res*, 22:2687–93, 1994.

I. Karagiannides, T. Thomou, T. Tchkonia, T. Pirtskhalava, K. E. Kypreos, A. Cartwright, G. Dalagiorgou, T. L. Lash, S. R. Farmer, N. A. Timchenko, and J. L. Kirkland. Increased CUG triplet repeat-binding protein-1 predisposes to impaired adipogenesis with aging. *J Biol Chem*, 281:23025–33, 2006.

S. J. Kavanagh, T. C. Schulz, P. Davey, C. Claudianos, C. Russell, and P. D. Rathjen. A family of RS domain proteins with novel subcellular localization and trafficking. *Nucleic Acids Res*, 33:1309–22, 2005.

A. Kawakami, Q. Tian, X. Duan, M. Streuli, S. F. Schlossman, and P. Anderson. Identification and functional characterization of a TIA-1-related nucleolysin. *Proc Natl Acad Sci U S A*, 89:8681–5, 1992.

T. Kawashima, A. R. Murakami, M. Ogasawara, K. Tanaka, R. Isoda, Y. Sasakura, T. Nishikata, H. Okano, and K. W. Makabe. Expression patterns of musashi homologs of the ascidians, Halocynthia roretzi and Ciona intestinalis. *Dev Genes Evol*, 210:162–5, 2000.

D. Kennedy, S. A. Wood, T. Ramsdale, P. P. Tam, K. A. Steiner, and J. S. Mattick. Identification of a mouse orthologue of the human ras-GAP-SH3-domain binding protein and structural confirmation that these proteins contain an RNA recognition motif. *Biomed Pept Proteins Nucleic Acids*, 2:93–9, 1996.

F. A. Khan, A. K. Jaiswal, and W. Szer. Cloning and sequence analysis of a human type A/B hnRNP protein. *FEBS Lett*, 290:159–61, 1991.

I. Khrebtukova, A. Kuklin, R. P. Woychik, and E. J. Michaud. Alternative processing of the human and mouse raly genes(1). *Biochim Biophys Acta*, 1447:107–12, 1999.

T. R. Kiehl, H. Shibata, T. Vo, D. P. Huynh, and S. M. Pulst. Identification and expression of a mouse ortholog of A2BP1. *Mamm Genome*, 12:595–601, 2001.

C. L. Kielkopf, S. Lucke, and M. R. Green. U2AF homology motifs: protein recognition in the RRM world. *Genes Dev*, 18:1513–26, 2004.

E. Kiesler, M. E. Hase, D. Brodin, and N. Visa. Hrp59, an hnRNP M protein in Chironomus and Drosophila, binds to exonic splicing enhancers and is required for expression of a subset of mRNAs. *J Cell Biol*, 168:1013–25, 2005.

K. Kimura, H. Saga, K. Hayashi, H. Obata, Y. Chimori, H. Ariga, and K. Sobue. c-Myc gene single-strand binding protein-1, MSSP-1, suppresses transcription of alpha-smooth muscle actin gene in chicken visceral smooth muscle cells. *Nucleic Acids Res*, 26:2420–5, 1998.

K. Kitagawa, X. Wang, I. Hatada, T. Yamaoka, H. Nojima, J. Inazawa, T. Abe, K. Mitsuya, M. Oshimura, A. Murata, and et al. Isolation and mapping of human homologues of an imprinted mouse gene U2af1-rs1. *Genomics*, 30:257–63, 1995.

K. C. Kleene, E. Mulligan, D. Steiger, K. Donohue, and M. A. Mastrangelo. The mouse gene encoding the testis-specific isoform of Poly(A) binding protein (Pabp2) is an expressed retroposon: intimations that gene expression in spermatogenic cells facilitates the creation of new genes. *J Mol Evol*, 47:275–81, 1998.

S. Kojima, K. Matsumoto, M. Hirose, M. Shimada, M. Nagano, Y. Shigeyoshi, S. Hoshino, K. Ui-Tei, K. Saigo, C. B. Green, Y. Sakaki, and H. Tei. LARK activates posttranscriptional expression of an essential mammalian clock protein, PERIOD1. *Proc Natl Acad Sci U S A*, 104:1859–64, 2007.

E. Kotani, T. Ohba, T. Niwa, K. B. Storey, J. S. Storey, S. Hara, H. Saito, Y. Sugimura, and T. Furusawa. De novo gene expression and antisense inhibition in cultured cells of BmTRN-1, cloned from the midgut of the silkworm, Bombyx mori, which is homologous with mammalian TIA-1/R. *Gene*, 320:67–79, 2003.

A. R. Krainer, G. C. Conway, and D. Kozak. Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes Dev*, 4:1158–71, 1990.

B. J. Krueger, C. Jeronimo, B. B. Roy, A. Bouchard, C. Barrandon, S. A. Byers, C. E. Searcey, J. J. Cooper, O. Bensaude, E. A. Cohen, B. Coulombe, and D. H. Price. LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Res*, 36:2219–29, 2008.

U. Kuhn, A. Nemeth, S. Meyer, and E. Wahle. The RNA binding domains of the nuclear poly(A)-binding protein. *J Biol Chem*, 278:16916–25, 2003.

Y. Kurihara, M. Tokuriki, R. Myojin, T. Hori, A. Kuroiwa, Y. Matsuda, T. Sakurai, M. Kimura, N. B. Hecht, and S. Uesugi. CPEB2, a novel putative translational regulator in mouse haploid germ cells. *Biol Reprod*, 69:261–8, 2003.

A. N. Ladd, N. Charlet, and T. A. Cooper. The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing. *Mol Cell Biol*, 21:1285–96, 2001.

A. N. Ladd, N. H. Nguyen, K. Malhotra, and T. A. Cooper. CELF6, a member of the CELF family of RNA-binding proteins, regulates muscle-specific splicing enhancer-dependent alternative splicing. *J Biol Chem*, 279:17756–64, 2004.

M. H. Lambermon, G. G. Simpson, D. A. Wieczorek Kirk, M. Hemmings-Mieszczak, U. Klahre, and W. Filipowicz. UBP1, a novel hnRNP-like protein that functions at multiple steps of higher plant nuclear pre-mRNA maturation. *Embo J*, 19:1638–49, 2000.

G. Lazar, T. Schaal, T. Maniatis, and H. M. Goodman. Identification of a plant serine-arginine-rich protein similar to the mammalian splicing factor SF2/ASF. *Proc Natl Acad Sci U S A*, 92:7672–6, 1995.

C. Le Guiner, M. C. Gesnel, and R. Breathnach. TIA-1 or TIAR is required for DT40 cell viability. *J Biol Chem*, 278:10465–76, 2003.

T. D. Levine, F. Gao, P. H. King, L. G. Andrews, and J. D. Keene. Hel-N1: an autoimmune RNA-binding protein with specificity for 3′ uridylate-rich untranslated regions of growth factor mRNAs. *Mol Cell Biol*, 13:3494–504, 1993.

A. P. Lieberman, D. L. Friedlich, G. Harmison, B. W. Howell, C. L. Jordan, S. M. Breedlove, and K. H. Fischbeck. Androgens regulate the mammalian homologues of invertebrate sex determination genes tra-2 and fox-1. *Biochem Biophys Res Commun*, 282:499–506, 2001.

N. S. Lim, G. Kozlov, T. C. Chang, O. Groover, N. Siddiqui, L. Volpon, G. De Crescenzo, A. B. Shyu, and K. Gehring. Comparative peptide binding studies of the PABC domains from the ubiquitin-protein isopeptide ligase HYD and poly(A)-binding protein. Implications for HYD function. *J Biol Chem*, 281:14376–82, 2006.

L. Liu, E. M. Rodriguez-Belmonte, N. Mazloum, B. Xie, and M. Y. Lee. Identification of a novel protein, PDIP38, that interacts with the p50 subunit of DNA polymerase delta and proliferating cell nuclear antigen. *J Biol Chem*, 278:10041–7, 2003.

D. Longman, I. L. Johnstone, and J. F. Caceres. Functional characterization of SR and SR-related genes in Caenorhabditis elegans. *Embo J*, 19:1625–37, 2000.

S. Lopato, E. Waigmann, and A. Barta. Characterization of a novel arginine/serine-rich splicing factor in Arabidopsis. *Plant Cell*, 8:2255–64, 1996.

S. Lopato, R. Gattoni, G. Fabini, J. Stevenin, and A. Barta. A novel family of plant splicing factors with a Zn knuckle motif: examination of RNA binding and splicing activities. *Plant Mol Biol*, 39:761–73, 1999a.

S. Lopato, M. Kalyna, S. Dorner, R. Kobayashi, A. R. Krainer, and A. Barta. atSRp30, one of two SF2/ASF-like proteins from Arabidopsis thaliana, regulates splicing of specific plant genes. *Genes Dev*, 13:987–1001, 1999b.

S. Lopato, C. Forstner, M. Kalyna, J. Hilscher, U. Langhammer, K. Indrapichate, Z. J. Lorkovic, and A. Barta. Network of interactions of a novel plant-specific Arg/Ser-rich protein, atRSZ33, with atSC35-like splicing factors. *J Biol Chem*, 277:39989–98, 2002.

Z. J. Lorkovic and A. Barta. Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant Arabidopsis thaliana. *Nucleic Acids Res*, 30:623–35, 2002.

Z. J. Lorkovic, D. A. Wieczorek Kirk, U. Klahre, M. Hemmings-Mieszczak, and W. Filipowicz. RBP45 and RBP47, two oligouridylate-specific hnRNP-like proteins interacting with poly(A)+ RNA in nuclei of plant cells. *Rna*, 6:1610–24, 2000a.

Z. J. Lorkovic, D. A. Wieczorek Kirk, M. H. Lambermon, and W. Filipowicz. Pre-mRNA splicing in higher plants. *Trends Plant Sci*, 5:160–7, 2000b.

L. A. Lowery, J. Rubin, and H. Sive. Whitesnake/sfpq is required for cell survival and neuronal development in the zebrafish. *Dev Dyn*, 236:1347–57, 2007.

B. Lowin, L. French, J. C. Martinou, and J. Tschopp. Expression of the CTL-associated protein TIA-1 during murine embryogenesis. *J Immunol*, 157:1448–54, 1996.

X. Lu, N. A. Timchenko, and L. T. Timchenko. Cardiac elav-type RNA-binding protein (ETR-3) binds to RNA CUG repeats expanded in myotonic dystrophy. *Hum Mol Genet*, 8:53–60, 1999.

W. J. Ma, S. Cheng, C. Campbell, A. Wright, and H. Furneaux. Cloning and characterization of HuR, a ubiquitously expressed Elav-like protein. *J Biol Chem*, 271:8144–51, 1996.

X. Ma, M. J. Renda, L. Wang, E. C. Cheng, C. Niu, S. W. Morris, A. S. Chi, and D. S. Krause. Rbm15 modulates Notch-induced transcriptional activation and affects myeloid differentiation. *Mol Cell Biol*, 27:3056–64, 2007.

Z. Ma, S. W. Morris, V. Valentine, M. Li, J. A. Herbrick, X. Cui, D. Bouman, Y. Li, P. K. Mehta, D. Nizetic, Y. Kaneko, G. C. Chan, L. C. Chan, J. Squire, S. W. Scherer, and J. K. Hitzler. Fusion of two novel genes, RBM15 and MKL1, in the t(1;22)(p13;q13) of acute megakaryoblastic leukemia. *Nat Genet*, 28:220–1, 2001.

R. Macknight, I. Bancroft, T. Page, C. Lister, R. Schmidt, K. Love, L. Westphal, G. Murphy, S. Sherson, C. Cobbett, and C. Dean. FCA, a gene controlling flowering time in Arabidopsis, encodes a protein containing RNA-binding domains. *Cell*, 89:737–45, 1997.

D. Mahe, P. Mahl, R. Gattoni, N. Fischer, M. G. Mattei, J. Stevenin, and J. P. Fuchs. Cloning of human 2H9 heterogeneous nuclear ribonucleoproteins. Relation with splicing and early heat shock-induced splicing arrest. *J Biol Chem*, 272:1827–36, 1997.

D. Mahe, N. Fischer, D. Decimo, and J. P. Fuchs. Spatiotemporal regulation of hnRNP M and 2H9 gene expression during mouse embryonic development. *Biochim Biophys Acta*, 1492:414–24, 2000.

R. J. Maraia and R. V. Intine. Recognition of nascent RNA by the human La antigen: conserved and divergent features of structure and function. *Mol Cell Biol*, 21: 367–79, 2001.

V. Markovtsov, J. M. Nikolic, J. A. Goldman, C. W. Turck, M. Y. Chou, and D. L. Black. Cooperative assembly of an hnRNP complex induced by a tissue-specific homolog of polypyrimidine tract binding protein. *Mol Cell Biol*, 20: 7463–79, 2000.

M. A. Markus and B. J. Morris. Lark is the splicing factor RBM4 and exhibits unique subnuclear localization properties. *DNA Cell Biol*, 25:457–64, 2006.

E. L. Matunis, M. J. Matunis, and G. Dreyfuss. Characterization of the major hnRNP proteins from Drosophila melanogaster. *J Cell Biol*, 116:257–69, 1992.

M. J. Matunis, J. Xing, and G. Dreyfuss. The hnRNP F protein: unique primary structure, nucleic acid-binding properties, and subcellular localization. *Nucleic Acids Res*, 22:1059–67, 1994.

A. Maucuer, J. H. Camonis, and A. Sobel. Stathmin interaction with a putative kinase and coiled-coil-forming protein domains. *Proc Natl Acad Sci U S A*, 92:3100–4, 1995.

A. Maucuer, S. Ozon, V. Manceau, O. Gavet, S. Lawler, P. Curmi, and A. Sobel. KIS is a protein kinase with an RNA recognition motif. *J Biol Chem*, 272:23151–6, 1997.

A. E. McKee, E. Minet, C. Stern, S. Riahi, C. D. Stiles, and P. A. Silver. A genome-wide in situ hybridization map of RNA-binding proteins reveals anatomically restricted expression in the developing mouse brain. *BMC Dev Biol*, 5:14, 2005.

G. P. McNeil, A. J. Schroeder, M. A. Roberts, and F. R. Jackson. Genetic analysis of functional domains within the Drosophila LARK RNA-binding protein. *Genetics*, 159:229–40, 2001.

M. Meins, S. Schlickum, C. Wilhelm, J. Missbach, S. Yadav, B. Glaser, M. Grzmil, P. Burfeind, and F. Laccone. Identification and characterization of murine Brunol4, a new member of the elav/bruno family. *Cytogenet Genome Res*, 97:254–60, 2002.

H. Mi, O. Kops, E. Zimmermann, A. Jaschke, and M. Tropschug. A nuclear RNA-binding cyclophilin in human T cells. *FEBS Lett*, 398:201–5, 1996.

S. C. Milburn, J. W. Hershey, M. V. Davies, K. Kelleher, and R. J. Kaufman. Cloning and expression of eukaryotic initiation factor 4B cDNA: sequence determination identifies a common RNA recognition motif. *Embo J*, 9: 2783–90, 1990.

P. Montaville, Y. Dai, C. Y. Cheung, K. Giller, S. Becker, M. Michalak, S. E. Webb, A. L. Miller, and J. Krebs. Nuclear translocation of the calcium-binding protein ALG-2 induced by the RNA-binding protein RBM22. *Biochim Biophys Acta*, 1763:1335–43, 2006.

V. E. Myer and J. A. Steitz. Isolation and characterization of a novel, low abundance hnRNP protein: A0. *Rna*, 1: 171–82, 1995.

R. Myojin, S. Kuwahara, T. Yasaki, T. Matsunaga, T. Sakurai, M. Kimura, S. Uesugi, and Y. Kurihara. Expression and functional significance of mouse paraspeckle protein 1 on spermatogenesis. *Biol Reprod*, 71:926–32, 2004.

Y. Negishi, Y. Nishita, Y. Saegusa, I. Kakizaki, I. Galli, F. Kihara, K. Tamai, N. Miyajima, S. M. Iguchi-Ariga, and H. Ariga. Identification and cDNA cloning of single-stranded DNA binding proteins that interact with the region upstream of the human c-myc gene. *Oncogene*, 9: 1133–43, 1994.

J. Nielsen, J. Christiansen, J. Lykke-Andersen, A. H. Johnsen, U. M. Wewer, and F. C. Nielsen. A family of insulin-like growth factor II mRNA-binding proteins represses translation in late development. *Mol Cell Biol*, 19:1262–70, 1999.

H. Nishiyama, H. Higashitsuji, H. Yokoi, K. Itoh, S. Danno, T. Matsuda, and J. Fujita. Cloning and characterization of human CIRP (cold-inducible RNA-binding protein) cDNA and chromosomal assignment of the gene. *Gene*, 204:115–20, 1997.

M. Ohta, M. Sugita, and M. Sugiura. Three types of nuclear genes encoding chloroplast RNA-binding proteins (cp29, cp31 and cp33) are present in Arabidopsis thaliana: presence of cp31 in chloroplasts and its homologue in nuclei/cytoplasms. *Plant Mol Biol*, 27:529–39, 1995.

H. J. Okano and R. B. Darnell. A hierarchy of Hu RNA binding proteins in developing and adult neurons. *J Neurosci*, 17:3024–37, 1997.

L. J. Otero, M. P. Ashe, and A. B. Sachs. The yeast poly(A)-binding protein Pab1p stimulates in vitro poly(A)-dependent and cap-dependent translation by distinct mechanisms. *Embo J*, 18:3153–63, 1999.

S. H. Ou, F. Wu, D. Harrich, L. F. Garcia-Martinez, and R. B. Gaynor. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *J Virol*, 69:3584–96, 1995.

T. R. Pacheco, A. Q. Gomes, N. L. Barbosa-Morais, V. Benes, W. Ansorge, M. Wollerton, C. W. Smith, J. Valcarcel, and M. Carmo-Fonseca. Diversity of vertebrate splicing factor U2AF35: identification of alternatively spliced U2AF1 mRNAS. *J Biol Chem*, 279:27039–49, 2004.

P. S. Page-McCaw, K. Amonlirdviman, and P. A. Sharp. PUF60: a novel U2AF65-related splicing activity. *Rna*, 5:1548–60, 1999.

L. Paillard, F. Omilli, V. Legagneux, T. Bassez, D. Maniey, and H. B. Osborne. EDEN and EDEN-BP, a cis element and an associated factor that mediate sequence-specific mRNA deadenylation in Xenopus embryos. *Embo J*, 17:278–87, 1998.

H. A. Pan, Y. S. Lin, K. H. Lee, J. R. Huang, Y. H. Lin, and P. L. Kuo. Expression patterns of the DAZ-associated protein DAZAP1 in rat and human ovaries. *Fertil Steril*, 84 Suppl 2:1089–94, 2005.

J. Paris, K. Swenson, H. Piwnica-Worms, and J. D. Richter. Maturation-specific polyadenylation: in vitro activation by p34cdc2 and phosphorylation of a 58-kD CPE-binding protein. *Genes Dev*, 5:1697–708, 1991.

H. G. Park, J. Y. Yoon, and M. Choi. Heterogeneous nuclear ribonucleoprotein D/AUF1 interacts with heterogeneous nuclear ribonucleoprotein L. *J Biosci*, 32:1263–72, 2007.

N. I. Park and D. G. Muench. Biochemical and cellular characterization of the plant ortholog of PYM, a protein that interacts with the exon junction complex core proteins Mago and Y14. *Planta*, 225:625–39, 2007.

F. Parker, F. Maurier, I. Delumeau, M. Duchesne, D. Faucher, L. Debussche, A. Dugue, F. Schweighoffer, and B. Tocque. A Ras-GTPase-activating protein SH3-domain-binding protein. *Mol Cell Biol*, 16:2561–9, 1996.

T. J. Pemberton and J. E. Kay. The cyclophilin repertoire of the fission yeast Schizosaccharomyces pombe. *Yeast*, 22:927–45, 2005.

D. Penkov, R. Ni, C. Else, S. Pinol-Roma, F. Ramirez, and S. Tanaka. Cloning of a human gene closely related to the genes coding for the c-myc single-strand binding proteins. *Gene*, 243:27–36, 2000.

A. Perreault, C. Lemieux, and F. Bachand. Regulation of the nuclear poly(A)-binding protein by arginine methylation in fission yeast. *J Biol Chem*, 282:7552–62, 2007.

S. Pinol-Roma, M. S. Swanson, J. G. Gall, and G. Dreyfuss. A novel heterogeneous nuclear RNP protein with a unique distribution on nascent transcripts. *J Cell Biol*, 109:2575–87, 1989.

A. D. Polydorides, H. J. Okano, Y. Y. Yang, G. Stefani, and R. B. Darnell. A brain-enriched polypyrimidine tract-binding protein antagonizes the ability of Nova to regulate neuron-specific alternative splicing. *Proc Natl Acad Sci U S A*, 97:6350–5, 2000.

F. Pontvianne, I. Matia, J. Douet, S. Tourmente, F. J. Medina, M. Echeverria, and J. Saez-Vasquez. Characterization of AtNUC-L1 reveals a central role of nucleolin in nucleolus organization and silencing of AtNUC-L2 gene in Arabidopsis. *Mol Biol Cell*, 18:369–79, 2007.

M. Prigent, I. Barlat, H. Langen, and C. Dargemont. IkappaBalpha and IkappaBalpha /NF-kappa B complexes are retained in the cytoplasm through interaction with a novel partner, RasGAP SH3-binding protein 2. *J Biol Chem*, 275:36441–9, 2000.

Z. Qian and J. Wilusz. GRSF-1: a poly(A)+ mRNA binding protein which interacts with a conserved G-rich element. *Nucleic Acids Res*, 22:2334–43, 1994.

K. Richter, P. J. Good, and I. B. Dawid. A developmentally regulated, nervous system-specific gene in Xenopus encodes a putative RNA-binding protein. *New Biol*, 2: 556–65, 1990.

M. B. Roth, A. M. Zahler, and J. A. Stolk. A conserved family of nuclear phosphoproteins localized to sites of polymerase II transcription. *J Cell Biol*, 115:587–96, 1991.

D. Z. Rudner, R. Kanaar, K. S. Breger, and D. C. Rio. Mutations in the small subunit of the Drosophila U2AF splicing factor cause lethality and developmental defects. *Proc Natl Acad Sci U S A*, 93:10333–7, 1996.

A. Sailer, N. J. MacDonald, and C. Weissmann. Cloning and sequencing of the murine homologue of the human splicing factor U2AF65. *Nucleic Acids Res*, 20:2374, 1992.

S. Sakakibara, T. Imai, K. Hamaguchi, M. Okabe, J. Aruga, K. Nakajima, D. Yasutomi, T. Nagata, Y. Kurihara, S. Uesugi, T. Miyata, M. Ogawa, K. Mikoshiba, and H. Okano. Mouse-Musashi-1, a neural RNA-binding protein highly enriched in the mammalian CNS stem cell. *Dev Biol*, 176:230–42, 1996.

S. Sakakibara, Y. Nakamura, H. Satoh, and H. Okano. Rna-binding protein Musashi2: developmentally regulated expression in neural precursor cells and subpopulations of neurons in mammalian CNS. *J Neurosci*, 21: 8091–107, 2001.

K. Salehi-Ashtiani, A. Luptak, A. Litovchick, and J. W. Szostak. A genomewide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene. *Science*, 313:1788–92, 2006.

A. M. Salicioni, M. Xi, L. A. Vanderveer, B. Balsara, J. R. Testa, J. Dunbrack, R. L., and A. K. Godwin. Identification and structural analysis of human RBM8A and RBM8B: two highly conserved RNA-binding motif proteins that interact with OVCA1, a candidate tumor suppressor. *Genomics*, 69:54–62, 2000.

J. Sampath, P. R. Long, R. L. Shepard, X. Xia, V. Devanarayan, G. E. Sandusky, r. Perry, W. L., A. H. Dantzig, M. Williamson, M. Rolfe, and R. E. Moore. Human SPF45, a splicing factor, has limited expression in normal tissues, is overexpressed in many tumors, and can confer a multidrug-resistant phenotype to cells. *Am J Pathol*, 163:1781–90, 2003.

D. Scherly, C. Kambach, W. Boelens, W. J. van Venrooij, and I. W. Mattaj. Conserved amino acid residues within and outside of the N-terminal ribonucleoprotein motif of U1A small nuclear ribonucleoprotein involved in U1 RNA binding. *J Mol Biol*, 219:577–84, 1991.

G. Schmidt and D. Werner. Sequence of a complete murine cDNA reflecting an S phase-prevalent transcript encoding a protein with two types of nucleic acid binding motifs. *Biochim Biophys Acta*, 1216:317–20, 1993.

F. M. Schomburg, D. A. Patton, D. W. Meinke, and R. M. Amasino. FPA, a gene involved in floral induction in Arabidopsis, encodes a protein containing RNA-recognition motifs. *Plant Cell*, 13:1427–36, 2001.

G. R. Screaton, J. F. Caceres, A. Mayeda, M. V. Bell, M. Plebanski, D. G. Jackson, J. I. Bell, and A. R. Krainer. Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. *Embo J*, 14:4336–49, 1995.

D. Shahbazian, P. P. Roux, V. Mieulet, M. S. Cohen, B. Raught, J. Taunton, J. W. Hershey, J. Blenis, M. Pende, and N. Sonenberg. The mTOR/PI3K and MAPK pathways converge on eIF4B to control its phosphorylation and activity. *Embo J*, 25:2781–91, 2006.

Y. Shav-Tal, M. Cohen, S. Lapter, B. Dye, J. G. Patton, J. Vandekerckhove, and D. Zipori. Nuclear relocalization of the pre-mRNA splicing factor PSF during apoptosis involves hyperphosphorylation, masking of antigenic epitopes, and changes in protein interactions. *Mol Biol Cell*, 12:2328–40, 2001.

J. Shepard, M. Reick, S. Olson, and B. R. Graveley. Characterization of U2AF(6), a splicing factor related to U2AF(35). *Mol Cell Biol*, 22:221–30, 2002.

H. Shibata, D. P. Huynh, and S. M. Pulst. A novel protein with RNA-binding motifs interacts with ataxin-2. *Hum Mol Genet*, 9:1303–13, 2000.

C. Shin, Y. Feng, and J. L. Manley. Dephosphorylated SRp38 acts as a splicing repressor in response to heat shock. *Nature*, 427:553–8, 2004.

L. Shu, W. Yan, and X. Chen. RNPC1, an RNA-binding protein and a target of the p53 family, is required for maintaining the stability of the basal and stress-induced p21 transcript. *Genes Dev*, 20:2961–72, 2006.

Y. Shu, N. D. Rintala-Maki, V. E. Wall, K. Wang, C. A. Goard, C. E. Langdon, and L. C. Sutherland. The apoptosis modulator and tumour suppressor protein RBM5 is a phosphoprotein. *Cell Biochem Funct*, 25:643–53, 2007.

I. Shur, D. Ben-Avraham, and D. Benayahu. Alternatively spliced isoforms of a novel stromal RNA regulating factor. *Gene*, 334:113–21, 2004.

P. T. Sillekens, W. J. Habets, R. P. Beijer, and W. J. van Venrooij. cDNA cloning of the human U1 snRNA-associated A protein: extensive homology between U1 and U2 snRNP-specific proteins. *Embo J*, 6:3841–8, 1987.

G. G. Simpson, P. Vaux, G. Clark, R. Waugh, J. D. Beggs, and J. W. Brown. Evolutionary conservation of the spliceosomal protein, U2B″. *Nucleic Acids Res*, 19: 5213–7, 1991.

G. G. Simpson, G. P. Clark, H. M. Rothnie, W. Boelens, W. van Venrooij, and J. W. Brown. Molecular characterization of the spliceosomal proteins U1A and U2B" from higher plants. *Embo J*, 14:4540–50, 1995.

A. Smyk, M. Szuminska, K. A. Uniewicz, L. M. Graves, and P. Kozlowski. Human enhancer of rudimentary is a molecular partner of PDIP46/SKAR, a protein interacting with DNA polymerase delta and S6K1 and regulating cell growth. *Febs J*, 273:4728–41, 2006.

J. Soret, R. Gattoni, C. Guyon, A. Sureau, M. Popielarz, E. Le Rouzic, S. Dumon, F. Apiou, B. Dutrillaux, H. Voss, W. Ansorge, J. Stevenin, and B. Perbal. Characterization of SRp46, a novel human SR splicing factor encoded by a PR264/SC35 retropseudogene. *Mol Cell Biol*, 18:4924–34, 1998.

M. Soulard, V. Della Valle, M. C. Siomi, S. Pinol-Roma, P. Codogno, C. Bauvy, M. Bellini, J. C. Lacroix, G. Monod, G. Dreyfuss, and et al. hnRNP G: sequence and characterization of a glycosylated RNA-binding protein. *Nucleic Acids Res*, 21:4210–7, 1993.

M. Srivastava, P. J. Fleming, H. B. Pollard, and A. L. Burns. Cloning and sequencing of the human nucleolin cDNA. *FEBS Lett*, 250:99–105, 1989.

U. Steller, S. Kohls, B. Muller, R. Soller, R. Muller, J. Schlender, and D. H. Blohm. The RNA binding protein HuD: rat cDNA and analysis of the alternative spliced mRNA in neuronal differentiating cell lines P19 and PC12. *Brain Res Mol Brain Res*, 35:285–96, 1996.

C. Stover, G. Gradl, I. Jentsch, M. R. Speicher, R. Wieser, and W. Schwaeble. cDNA cloning, chromosome assignment, and genomic structure of a human gene encoding a novel member of the RBM family. *Cytogenet Cell Genet*, 92:225–30, 2001.

M. S. Swanson, T. Y. Nakagawa, K. LeVan, and G. Dreyfuss. Primary structure of human nuclear ribonucleoprotein particle C proteins: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA, and pre-rRNA-binding proteins. *Mol Cell Biol*, 7:1731–9, 1987.

A. Szabo, J. Dalmau, G. Manley, M. Rosenfeld, E. Wong, J. Henson, J. B. Posner, and H. M. Furneaux. HuD, a paraneoplastic encephalomyelitis antigen, contains RNA-binding domains and is homologous to Elav and Sex-lethal. *Cell*, 67:325–33, 1991.

N. Takahashi, N. Tochimoto, S. Y. Ohmori, H. Mamada, M. Itoh, M. Inamori, J. Shinga, S. Osada, and M. Taira. Systematic screening for genes specifically expressed in the anterior neuroectoderm during early Xenopus development. *Int J Dev Biol*, 49:939–51, 2005.

H. Tamada, E. Sakashita, K. Shimazaki, E. Ueno, T. Hamamoto, Y. Kagawa, and H. Endo. cDNA cloning and characterization of Drb1, a new member of RRM-type neural RNA-binding protein. *Biochem Biophys Res Commun*, 297:96–104, 2002.

K. Tang, E. C. Breen, and P. D. Wagner. Hu protein R-mediated posttranscriptional regulation of VEGF expression in rat gastrocnemius muscle. *Am J Physiol Heart Circ Physiol*, 283:H1497–504, 2002.

M. Theis, K. Si, and E. R. Kandel. Two previously undescribed members of the mouse CPEB family of genes and their inducible expression in the principal cell layers of the hippocampus. *Proc Natl Acad Sci U S A*, 100: 9602–7, 2003.

Q. Tian, M. Streuli, H. Saito, S. F. Schlossman, and P. Anderson. A polyadenylate binding protein localized to the granules of cytolytic lymphocytes induces DNA fragmentation in target cells. *Cell*, 67:629–39, 1991.

L. T. Timchenko, J. W. Miller, N. A. Timchenko, D. R. DeVore, K. V. Datar, L. Lin, R. Roberts, C. T. Caskey, and M. S. Swanson. Identification of a (CUG)n triplet repeat RNA-binding protein and its expression in myotonic dystrophy. *Nucleic Acids Res*, 24:4407–14, 1996.

L. T. Timchenko, E. Salisbury, G. L. Wang, H. Nguyen, J. H. Albrecht, J. W. Hershey, and N. A. Timchenko. Age-specific CUGBP1-eIF2 complex increases translation of CCAAT/enhancer-binding protein beta in old liver. *J Biol Chem*, 281:32806–19, 2006.

C. K. Too, R. Knee, A. L. Pinette, A. W. Li, and P. R. Murphy. Prolactin induces expression of FGF-2 and a novel FGF-responsive NonO/p54nrb-related mRNA in rat lymphoma cells. *Mol Cell Endocrinol*, 137:187–95, 1998.

N. S. Trede, J. Medenbach, A. Damianov, L. H. Hung, G. J. Weber, B. H. Paw, Y. Zhou, C. Hersey, A. Zapata, M. Keefe, B. A. Barut, A. B. Stuart, T. Katz, C. T. Amemiya, L. I. Zon, and A. Bindereif. Network of coregulated spliceosome components revealed by zebrafish mutant in recycling factor p110. *Proc Natl Acad Sci U S A*, 104:6608–13, 2007.

S. Tsui, T. Dai, S. Roettger, W. Schempp, E. C. Salido, and P. H. Yen. Identification of two novel proteins that interact with germ-cell-specific RNA-binding proteins DAZ and DAZL1. *Genomics*, 65:266–73, 2000.

J. G. Underwood, P. L. Boutz, J. D. Dougherty, P. Stoilov, and D. L. Black. Homologues of the Caenorhabditis el-

egans Fox-1 protein are neuronal splicing regulators in mammals. *Mol Cell Biol*, 25:10005–16, 2005.

C. Valle, A. R. Troiani, P. Lazzaretti, J. Bouvier, D. Cioli, and M. Q. Klinkert. Molecular and biochemical characterization of a protein cyclophilin from the nematode Haemonchus contortus(P). *Parasitol Res*, 96:199–205, 2005.

C. Van Buskirk and T. Schupbach. Half pint regulates alternative splice site selection in Drosophila. *Dev Cell*, 2:343–53, 2002.

D. J. Van Horn, C. J. Yoo, D. Xue, H. Shi, and S. L. Wolin. The La protein in Schizosaccharomyces pombe: a conserved yet dispensable phosphoprotein that functions in tRNA maturation. *Rna*, 3:1434–43, 1997.

B. A. Van Tine, J. F. Knops, A. Butler, P. Deloukas, G. M. Shaw, and P. H. King. Localization of HuC (ELAVL3) to chromosome 19p13.2 by fluorescence in situ hybridization utilizing a novel tyramide labeling technique. *Genomics*, 53:296–9, 1998.

G. K. Voeltz, J. Ongkasuwan, N. Standart, and J. A. Steitz. A novel embryonic poly(A) binding protein, ePAB, regulates mRNA deadenylation in Xenopus egg extracts. *Genes Dev*, 15:774–88, 2001.

E. Wahle. A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell*, 66:759–68, 1991.

Y. Wakamatsu and J. A. Weston. Sequential expression and role of Hu RNA-binding proteins during neurogenesis. *Development*, 124:3449–60, 1997.

H. Wang, M. X. Gao, L. Li, B. Wang, N. Hori, and K. Sato. Isolation, expression, and characterization of the human ZCRB1 gene mapped to 12q12. *Genomics*, 89:59–69, 2007a.

K. Wang, G. Ubriaco, and L. C. Sutherland. RBM6-RBM5 transcription-induced chimeras are differentially expressed in tumours. *BMC Genomics*, 8:348, 2007b.

M. Y. Wang, M. Cutler, I. Karimpour, and K. C. Kleene. Nucleotide sequence of a mouse testis poly(A) binding protein cDNA. *Nucleic Acids Res*, 20:3519, 1992.

Z. L. Wang, J. Li, Q. Y. Xia, P. Zhao, J. Duan, X. F. Zha, and Z. H. Xiang. Identification and expression pattern of Bmlark, a homolog of the Drosophila gene lark in Bombyx mori. *DNA Seq*, 16:224–9, 2005.

J. F. Welk, A. Charlesworth, G. D. Smith, and A. M. MacNicol. Identification and characterization of the gene encoding human cytoplasmic polyadenylation element binding protein. *Gene*, 263:113–20, 2001.

K. Wentz-Hunter and J. Potashkin. The small subunit of the splicing factor U2AF is conserved in fission yeast. *Nucleic Acids Res*, 24:1849–54, 1996.

C. L. Will, C. Schneider, M. Hossbach, H. Urlaub, R. Rauhut, S. Elbashir, T. Tuschl, and R. Luhrmann. The human 18S U11/U12 snRNP contains a set of novel proteins not found in the U2-dependent spliceosome. *Rna*, 10:929–41, 2004.

E. Winstall, M. Sadowski, U. Kuhn, E. Wahle, and A. B. Sachs. The Saccharomyces cerevisiae RNA-binding protein Rbp29 functions in cytoplasmic mRNA metabolism. *J Biol Chem*, 275:21817–26, 2000.

L. Wu, D. Wells, J. Tay, D. Mendis, M. A. Abbott, A. Barnitt, E. Quinlan, A. Heynen, J. R. Fallon, and J. D. Richter. CPEB-mediated cytoplasmic polyadenylation and the regulation of experience-dependent translation of alpha-CaMKII mRNA at synapses. *Neuron*, 21:1129–39, 1998.

H. Yang, C. S. Duckett, and T. Lindsten. iPABP, an inducible poly(A)-binding protein detected in activated human T cells. *Mol Cell Biol*, 15:6770–6, 1995.

Y. S. Yang, J. H. Hanke, L. Carayannopoulos, C. M. Craft, J. D. Capra, and P. W. Tucker. NonO, a non-POU-domain-containing, octamer-binding protein, is the mammalian homolog of Drosophila nonAdiss. *Mol Cell Biol*, 13:5593–603, 1993.

A. Yoda, H. Sawa, and H. Okano. MSI-1, a neural RNA-binding protein, is involved in male mating behaviour in Caenorhabditis elegans. *Genes Cells*, 5:885–895, 2000.

C. B. Yohn, A. Cohen, C. Rosch, M. R. Kuchka, and S. P. Mayfield. Translation of the chloroplast psbA mRNA requires the nuclear-encoded poly(A)-binding protein, RB47. *J Cell Biol*, 142:435–42, 1998.

C. J. Yoo and S. L. Wolin. La proteins from Drosophila melanogaster and Saccharomyces cerevisiae: a yeast homolog of the La autoantigen is dispensable for growth. *Mol Cell Biol*, 14:5412–24, 1994.

A. M. Zahler, W. S. Lane, J. A. Stolk, and M. B. Roth. SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev*, 6:837–47, 1992.

A. M. Zahler, K. M. Neugebauer, J. A. Stolk, and M. B. Roth. Human SR proteins and isolation of a cDNA encoding SRp75. *Mol Cell Biol*, 13:4023–8, 1993.

P. D. Zamore, J. G. Patton, and M. R. Green. Cloning and domain structure of the mammalian splicing factor U2AF. *Nature*, 355:609–14, 1992.

L. Zeng, Z. Zhou, J. Xu, W. Zhao, W. Wang, Y. Huang, C. Cheng, M. Xu, Y. Xie, and Y. Mao. Molecular cloning, structure and expression of a novel nuclear RNA-binding cyclophilin-like gene (PPIL4) from human fetal brain. *Cytogenet Cell Genet*, 95:43–7, 2001.

M. Zhang, P. D. Zamore, M. Carmo-Fonseca, A. I. Lamond, and M. R. Green. Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit. *Proc Natl Acad Sci U S A*, 89: 8769–73, 1992.

W. J. Zhang and J. Y. Wu. Functional properties of p54, a novel SR protein active in constitutive and alternative splicing. *Mol Cell Biol*, 16:5400–8, 1996.

E. Zhao, J. Li, Y. Xie, W. Jin, Z. Zhang, J. Chen, L. Zeng, G. Yin, J. Qian, H. Wu, K. Ying, R. C. Zhao, and Y. Mao. Cloning and identification of a novel human RNPC3 gene that encodes a protein with two RRM domains and is expressed in the cell nucleus. *Biochem Genet*, 41: 315–23, 2003.

W. M. Zhao, C. Jiang, T. T. Kroll, and P. W. Huber. A proline-rich protein binds to the localization element of Xenopus Vg1 mRNA and to ligands involved in actin polymerization. *Embo J*, 20:2315–25, 2001.

Q. Zhou and P. A. Sharp. Tat-SF1: cofactor for stimulation of transcriptional elongation by HIV-1 Tat. *Science*, 274:605–10, 1996.

D. A. Zorio and T. Blumenthal. U2AF35 is encoded by an essential gene clustered in an operon with RRM/cyclophilin in Caenorhabditis elegans. *Rna*, 5: 487–94, 1999.

D. A. Zorio, K. Lea, and T. Blumenthal. Cloning of Caenorhabditis U2AF65: an alternatively spliced RNA containing a novel exon. *Mol Cell Biol*, 17:946–53, 1997.

i

node	RRM# (KOG#) KOG name	reference sequences	MP	RW	TW	RL	mRW
1	RRM1 (0110) RBP	–	y	y	y	y[1]	y
2	RRM1 (0117) hnRNPr	hnRNP (excl. flowering plants)	y[2]	y	y	y	y[1]
3	RRM1 (0123) PABP	PABP	y[3]	y	y[1]	y	y
4	RRM2 (0123) PABP RRM2 (0131) splicing factor 2b [RRM2 (0115) RBP p54 nrb]	PABP – (fungi) p54 nrb/PSP1/PSF	y[4]	y[5]	y[4]	y[4]	y[5]
5	RRM3 (0123) PABP	PABP (minor subgroup)	y[3]	y	y	y	y
6	RRM4 (0123) PABP	PABP	y[3]	y	y	y	y
7	RRM1 (0144) BRUNO	BRUNO FCAFPA (land plants)	y[2]	y	y	y	y
8	RRM2 (0144) BRUNO	BRUNO FCAFPA (land plants)	y[2]	y	y	y	y
9	RRM3 (0144) BRUNO	BRUNO	y[3]	y	y	y	y
10	RRM1 (0145) ELAV/HU	ELAV	y[3]	y	y	y	y
11	RRM2 (0145) ELAV/HU	ELAV/RBMS	y[3]	y	y	y	y
12	RRM3 (0145) ELAV/HU	ELAV	y[3]	y	y	y	y
13	[RRM1 (0151) predicted splicing regulator]	–	y[3]	y	y[1]	y[1]	y[1]
14	[RRM1 (0153) RBP]	RBM	y[3]	y	y[1]	y[1]	y[1]
15	–	–	n[6]	y	y	n[6]	y
16	**RRM1 (0105) ASF/SF2** **RRM1 (0106) SRp55/B52/SRp75**	**dual RRM SR proteins**	y	y	y	y	y
17	**[RRM1 (0107) SRp20/9G8]**	**single RRM Zn-K SR proteins**	y	y	y	y	y
18	–	see footnote [7]	n	y	n	y	y
19	[RRM1 (0111) cyclophilin-type PPCTI[8]]	cyclophilin	y[9]	y	y[9]	y[9]	y
20	[RRM1 (0130) RBP RBM8[10]]	RBM	y[9]	y	y[9]	y	y
21	–	–	n	n	n	n	y
22	**RRM1 (4207) predicted SR protein**	**single RRM SR proteins**	y	y	y	y	y
23	RRM1 (0111) cyclophilin-type PPCTI[8]	**SRrp**	y	y	y	y	y
24	[RRM1 (0121) CBP20]	–	y	y	y	y	y[9]
25	–	–	y	y	y	y	n[11]
26	**RRM1 (-)**	**SR45 (green plants)**	y	y	y	y	y
27	RRM1 (4209) splicing factor RNPS1	RNPS1	y	y	y	y	y
28	–	–	y	y	y	y	y
29	**RRM2 (0105) ASF/SF2** **RRM2 (0106) SRp55/B52/SRp75**	**dual RRM SR proteins (excl. plants)**	y	y	y	y	y
30	RRM1/2/3 (4212) hnRNPm [RRM1 (0533) RRM-containing protein]	hnRNP PDIP/FCAFPA	y[12]	y[13]	y[13]	y[12]	y[12]
31	RRM1 (0117) hnRNPr	– (flowering plants)	y	y	y	y	y
32	RRM3 (0110) RBP	–	y[14]	y	y	y	y
33	[RRM1 (0115) RBP p54 nrb]	p54 nrb-PSP1-PSF	n	y	n	n	y
34	RRM1 (2277) CID1	–	y	y	y[15]	y[15]	y
35	RRM1 (0116) Rasputin RRM1 (0149) SEB4 RRM1/2 (4205) HRP1	G3BP RBM/30KRRM hnRNP/Musashi/Dazap/TARDBP/OligoU	y[16]	y	n	y[16]	y
36	RRM1 (0113) U1 snRNP	snRNP	y[15]	y	y	y	y
37	RRM2 (0128) SART3	SART3	y[15]	y	y	y[17]	y[17]
38	[RRM1 (0415) PPCTI[8]]	cyclophilin	y[15]	y[15]	y	y	y
39	–	–	n[6]	n[18]	y[19]	y[20]	y[20]
40	RRM1 (4206) U1A/U2B snRNP	snRNP	y	y	y	y	y
41	RRM2 (4206) U1A/U2B snRNP	snRNP	y	y	y	y	y
42	RRM1 (4660) Mei2	–	y	y	y	y	y
43	RRM2 (4660) Mei2	–	y	y	y	y	y
44	[RRM1 (0114) RBP]	–	y	y	y	y	y
45	–	–	n[21]	n[22]	n[23]	n[24]	n[24]
46	RRM1 (0131) splicing factor 2b	see footnote [25]	y	y	y	y	y
47	RRM1 (4454) RBP	–	y	y	y	y	y
48	[RRM1 (2193) IGF2BP]	IGF2BP	y	y	y	y	y

node	RRM# (KOG#) KOG name	reference sequences	MP	RW	TW	RL	mRW
49	–	–	y	y	n[26]	y	n[26]
50	RRM1 (0109) LARK	LARK	y	y	y	y	y
51	RRM2 (0109) LARK	LARK	y	y	y	y	y
52	–	–	n	y	y	n	y
53	RRM2 (0112) RBP	RBM	y	y	y	y	y[27]
54	RRM3 (0112) RBP	–	y	y	y	y	y[27]
55	–	–	n	y	y	y	y
56	**RRM2 (0106) SRp55/B52/SRp75**	**plant-specific RS group of SR proteins**	y	y	y	y	y
57	RRM1 (0108) RNA15	ZCRB1	y	y	y	y	y
58	RRM1 (0108) RNA15	GRSRBP/hnRNP/RBMY	y	y	y	y	y
59	RRM4/5 (0110) RBP	–	y	y	y	y	y
60	RRM4/5/6 (0110) RBP	–	y	y	y	y	y
61	RRM4 (0112) RBP	–	y	y	y	y	y
62	RRM1 (0117) hnRNPr	– (land plants)	y	y	y	y	y
63	RRM3 (0117) hnRNPr	hnRNP	n[28]	y	y	n[28]	n[28]
64	RRM2 (0120) U2AF	snRNP	y	y	y	y	y
65	RRM1 (0122) TIF3 (eIF3)	–	y	y	y	y	y
66	RRM3 (0123) PABP	PABP (major subgroup)	y[29]	y	y	y[2930]	y[30]
67	RRM2 (0124) PUF60 RRM2 (0147) CAPER/UMH [RRM1 (0124) PUF60]	UMH UMH UMH	y[31]	y[31]	y[32]	y[32]	y[32]
68	RRM1 (0125) ataxin-2 binding protein RRM1 (0127) fibrillarin RRM1 (4661) SAF-B	FOX – (land plants) –	n	y	y	y	y
69	RRM1 (0126) RBP	–	y	y	y	y	y
70	RRM1 (0127) fibrillarin	RBM	y	y	y	y	y
71	RRM2 (0127) fibrillarin	RBM	y	y	y	y	y
72	RRM3 (0127) fibrillarin	RBM	y	y	y	y	y
73	RRM1 (0128) SART3	SART3 (excl. green plants)	y	y	y	y	y
74	RRM1 (0128) SART3	– (green plants)	y	y	y	y	y
75	RRM1 (0147) CAPER/UMH	CAPER/UMH	y	y	y	y	y
76	RRM1 (0148) TIA1/TIAR	TIATIAR/OligoU	y	n[6]	y	y	y
77	RRM2 (0148) TIA1/TIAR RRM1 (0226) RBP	TIATIAR/OligoU –	y	y	y	y	y
78	RRM3 (0148) TIA1/TIAR	OligoU	y	y	y	y	y
79	RRM3 (0148) TIA1/TIAR	TIATIAR/OligoU	n[6]	y	y	y	y
80	RRM? (1365) fusillin RRM2 (4211) hnRNPf	– hnRNP	y	y	y	y	y
81	RRM1 (1855) predicted RBP	La	y	y	y	y	y
82	RRM1 (4208) NIFK	–	y	y	y	y	y
83	RRM1 (4209) splicing factor RNPS1	PABP (26%)	y	y	y	y	y
84	RRM1 (4213) RBP La	La	y	y	y	y	y
85	RRM2 (4676) splicing factor SR rich	SRrp	y	y	y	y	y
86	RRM1 (–)	hnRNP/RALYL	y	y	y	y	y
87	RRM1 (–)	eIF4b	y	y	y	y	y
88	RRM1 (0108)[33]	prokaryotic RRMs	n[6]	y	y	y	y

[1]not associated to node (15); [2]also associated in MP tree but independently of node (15); [3]forming a mini node in MP tree (15); [4]RRM2 of p54 nrb not associated to the group; [5]RRM2 of p54 nrb associated to the group; [6]split; [7]9G8 and SC35/SCL groups associated in TF and MP trees; [8]PPCTI: peptidyl-propyl cis-trans isomerase; [9]not associated to SR proteins; [10]always associated to 9G8 SR proteins; [11]RRM1 of CBP20 not associated to the group; [12]RRM1 of KOG0533 not associated to the group; [13]RRM1 of KOG0533 associated to the group; [14]fungal RBP missing from the group; [15]not associated to the group; [16]RRM1 of Rasputin not associated to the group; [17]more distantly related to the group; [18]nodes (36, 37) associated; [19]nodes (36, 37, 38) associated; [20]nodes (36, 37) + (38) associated; [21]4 groups: nodes (42, 44), node (40), node (41), and node (43); [22]2 groups: nodes (41, 42, 43, 44) and node (40); [23]2 groups: nodes (40, 41, 43, 44) and node (42); [24]4 groups: nodes (40, 43), node (41), node (42), and node (44); [25]snRNPs represent 8% of group proteins; [26]RRM1 of IGF2BP not associated to the group; [27]nodes (53) and (54) merged; [28]split into plant and animal groups; [29]associated to node (5); [30]associated to node (6); [31]RRM1 of PUF60 not associated to the group; [32]RRM1 of PUF60 associated to the group; [33]mostly KOG0108, but also a lower % of KOG0116 and KOG0145

Table S4

dataset	seq#/AA#	meth.	heuristic	model	tree	node associations
original	1266x72	MP	PAUP*	–	Fig. S5	[16] vs. [17 + (22-23) + (26-27)]
original	1266x72	ML	RAxML	WAG + G4	Fig. S6	[16 + 17 + (22-23)] vs. [(26-27)]
original	1266x72	ML	TreeFinder	WAG + G4	Fig. S7	[16] vs. [17 + (22-23)] vs. [(26-27)]
original	1266x72	ML	RAxML	LG + F + G4	Fig. S8	[16 + 17 + (26-27)] vs. [(22-23)]
enlarged	1831x72	ML	RAxML	WAG + G4	Fig. S9	[16 + 17 + (26-27)] vs. [(22-23)]

Table S5

SR subfamily	key features	secondary features
SC35	K P S KS	R RS
SCL	P R GR RS	E G S
SRrp	K P R S KS	RS
9G8	R RS	S
SRp20	G	P R S RS
RSZ	G GG GR	P R S RS
RS2Z		G P R S GR RS
ASF/SF2	G	P R S RS
SRp30c	G	P R S RS
ASF-like (plants)	G KS RS SS	R S
SRp40	G K S KS	R RS
SRp55	G K S KS	R RS
SRp75	G K S KS	R RS
RS		E P R S RS
SR45	K P R S PP RS SS SSS	
RNPS1	P R S PP RS SS ST SSS	

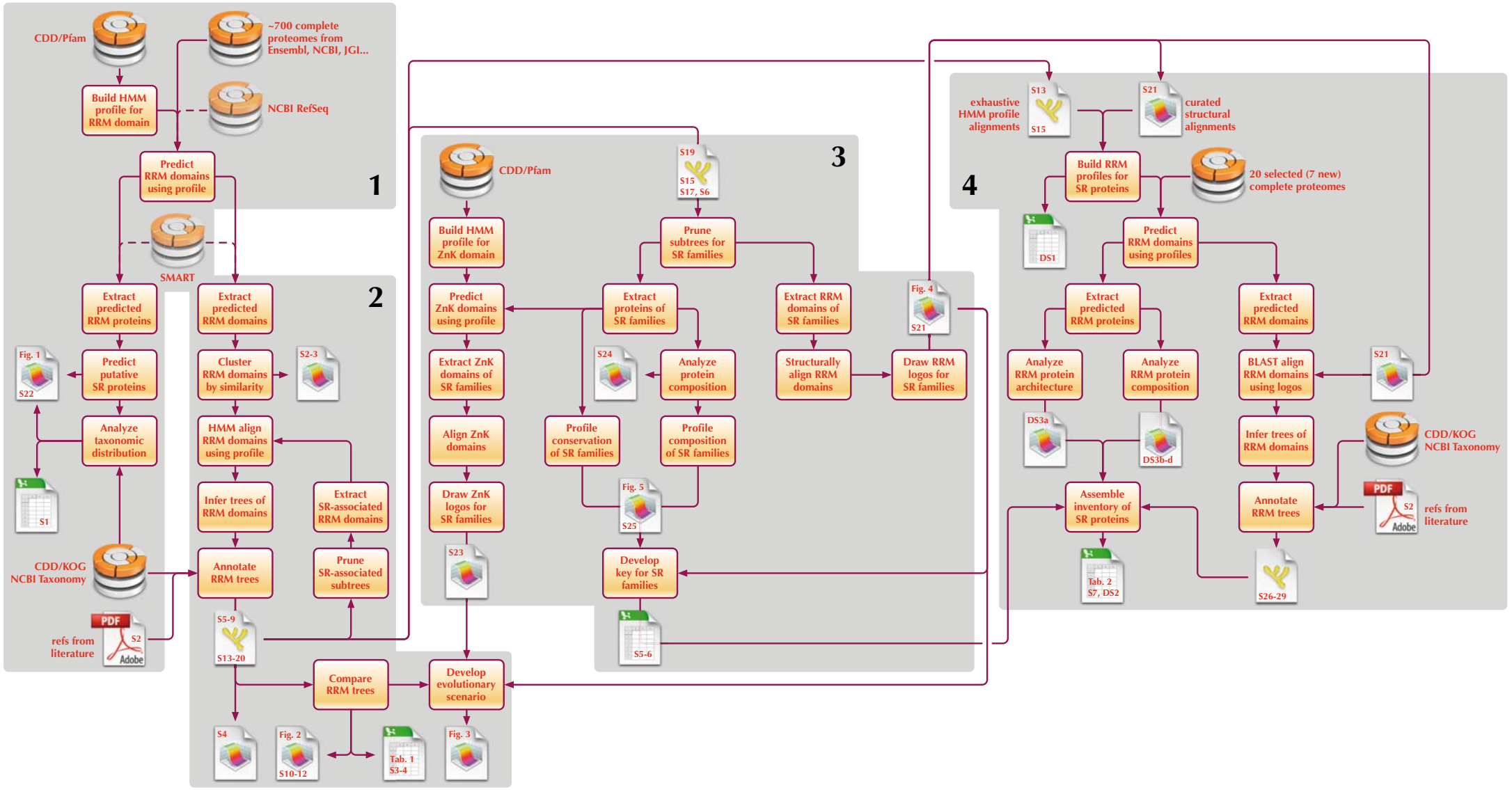
Table S6

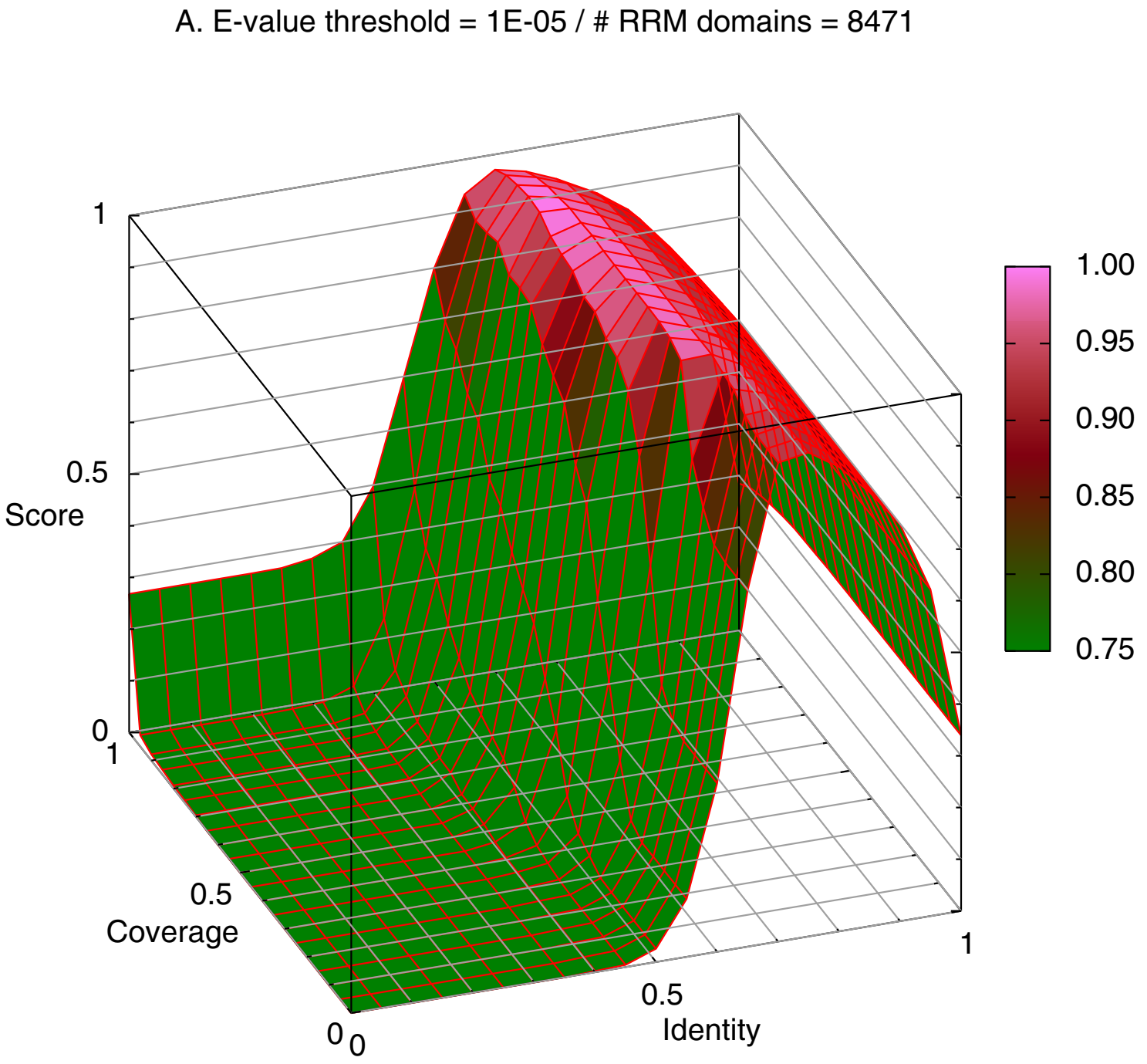
step	criteria	result
1	one RRM domain	step 2
	two RRM domains	step 9
2	2 Zn knuckles	▷ **RS2Z**
	1 Zn knuckle	step 3
	no Zn knuckle	step 4
3	enriched in G, GG, GR, GGG between RRM and ZnK	▷ **RSZ**
	enriched in R, RS between RRM and ZnK	▷ **9G8**
4	enriched in P, R, RS before RRM	step 5
	otherwise	step 7
5	enriched in GR before RRM	▷ **SCL**
	enriched in S, GS, SS before RRM and P, PP, PPP after RRM	step 6
6	enriched in R, ST, SSS before RRM	▷ **RNPS1**
	enriched in K, SSS after RRM	▷ **SR45**
7	enriched in K, KS after RRM and P, S before RRM	step 8
	enriched G before RRM	▷ **SRp20**
8	enriched in R before RRM	▷ **SRrp**
	otherwise	▷ **SC35**
9	second RRM lacks SWKDLKD motif	▷ **RS**
	enriched in G between RRMs	step 10
10	enriched in KS after second RRM	step 11
	enriched in G before first RRM	▷ **ASF/SF2, SRp30c**
11	enriched in RS, SS before RRM	▷ **ASF-like (plant)**
	enriched in S, K, KS after last domain	▷ **SRp40-55-75**

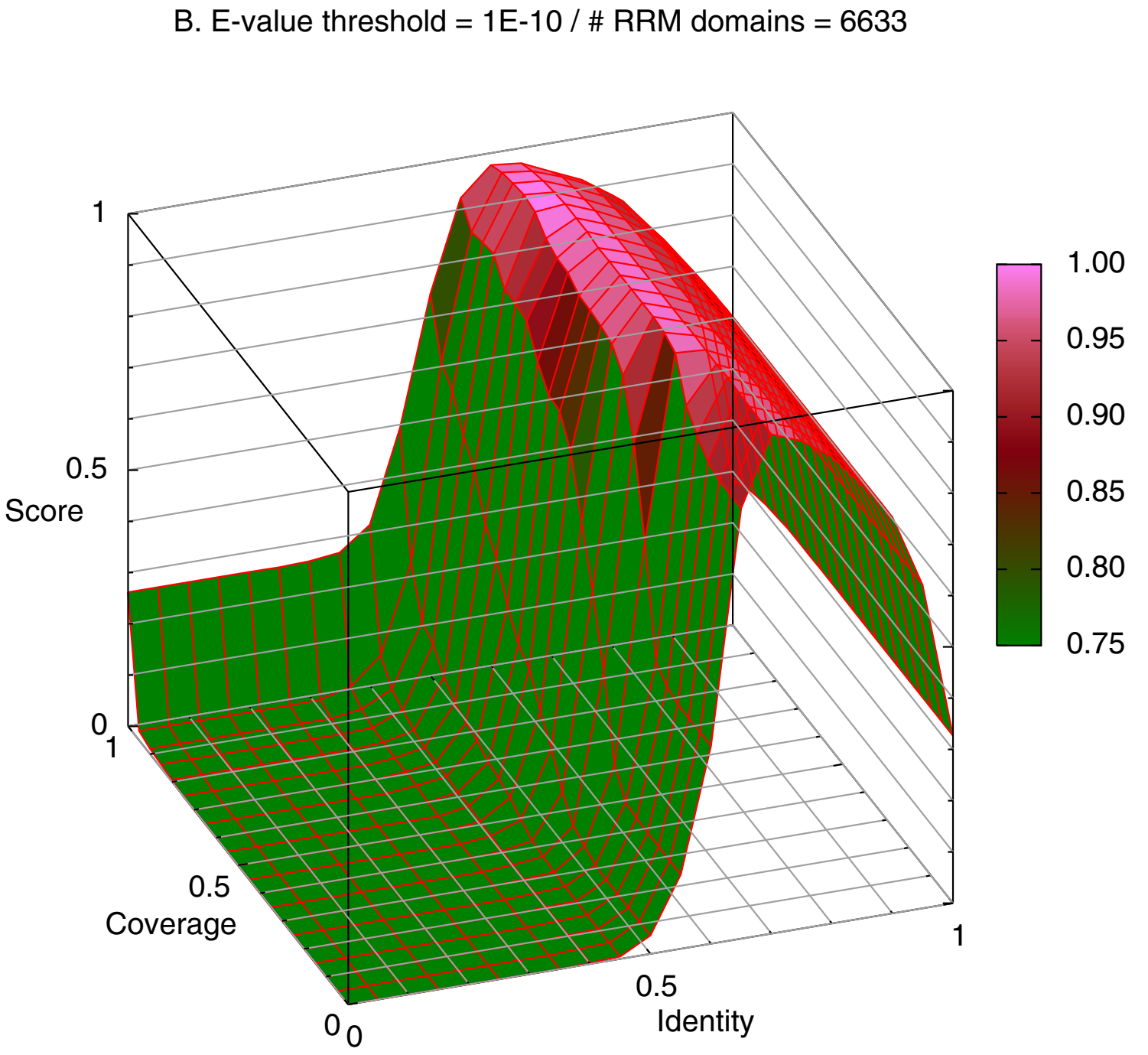
Table S7

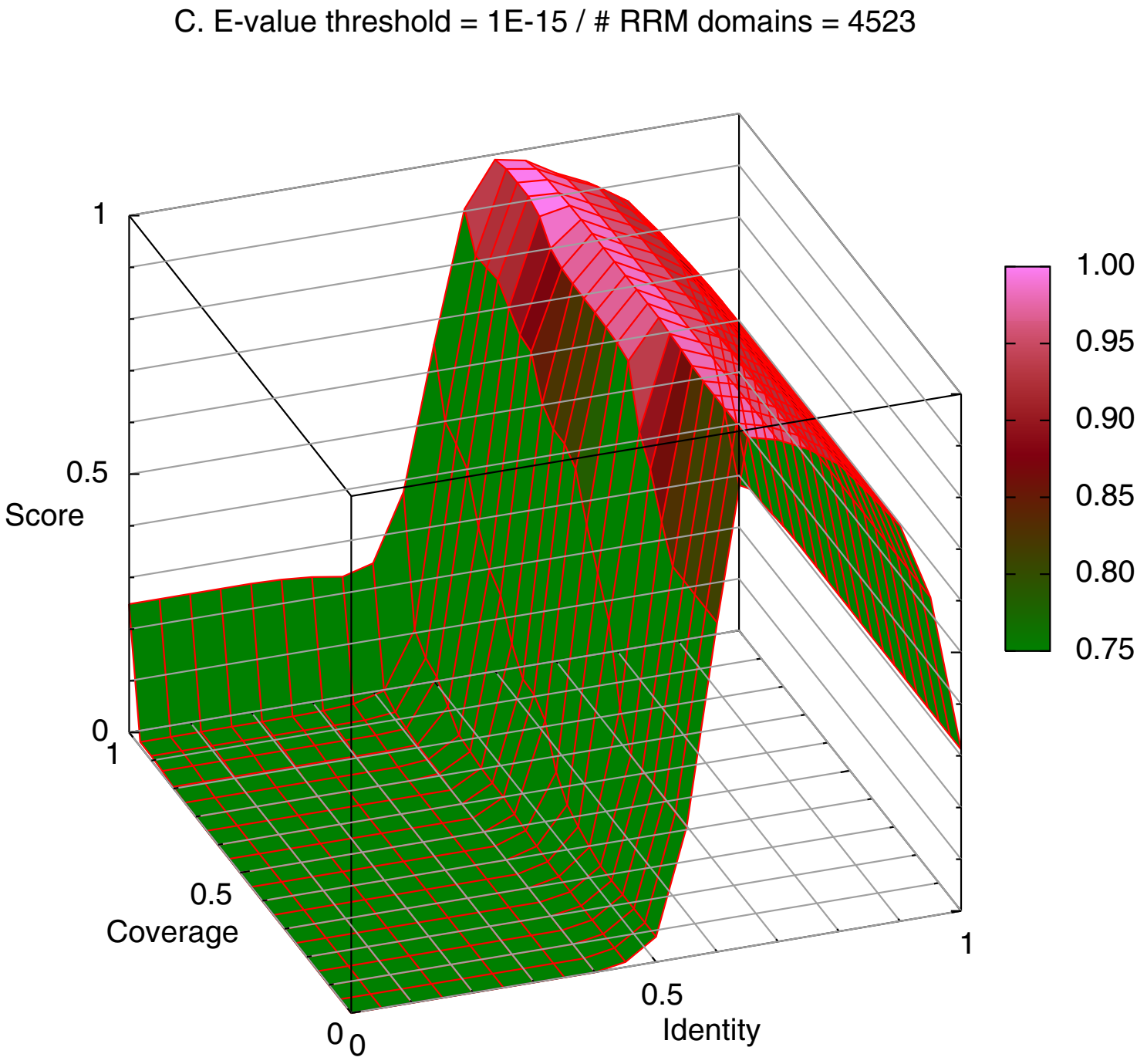
organism	lineage	cand	rej	iso	conf	key	trunc
Homo sapiens	Eutheria	23	0	11	12	11	0
Mus musculus	Eutheria	26	1	13	12	12	0
Drosophila melanogaster	Insecta	15	0	6	9	5	0
Caenorhabditis elegans	Nematoda	14	0	7	7	4	0
Aspergillus fumigatus	Ascomycota	0	0	0	0	0	0
Arabidopsis thaliana	eudicotyledons	29	0	10	19	18	0
Populus trichocarpa	eudicotyledons	30	0	0	30	21	6
Oryza sativa	Liliopsida	23	1	0	22	17	0
Sorghum bicolor	Liliopsida	21	0	0	21	15	4
Selaginella moellendorfii	Lycopodiophyta	15	0	5	10	4	4
Physcomitrella patens	Bryophyta	12	0	0	12	11	0
Chlamydomonas reinhardtii	Chlorophyta	9	0	0	9	4	3
Volvox carteri	Chlorophyta	8	1	0	7	1	1
***Chlorella* sp. NC64A**	Chlorophyta	7	0	0	7	3	1
***Micromonas pusilla* CCMP1545**	Chlorophyta	3	0	0	3	0	3
***Micromonas pusilla* RCC299**	Chlorophyta	3	0	0	3	1	1
***Ostreococcus* sp. RCC809**	Chlorophyta	3	0	0	3	0	3
Ostreococcus lucimarinus	Chlorophyta	3	0	0	3	0	3
Ostreococcus tauri	Chlorophyta	1	0	0	1	0	1
Cyanidioschyzon merolae	Rhodophyta	2	0	0	2	1	0
Total		**247**	**3**	**52**	**192**	**128**	**30**

Fig. S1


CDD/Pfam
~700 complete proteomes from Ensembl, NCBI, JGI...
Build HMM profile for RRM domain
NCBI RefSeq
Predict RRM domains using profile
1
SMART
2
Extract predicted RRM proteins
Extract predicted RRM domains
Fig. 1
S22
Predict putative SR proteins
Cluster RRM domains by similarity
S2-3
Analyze taxonomic distribution
HMM align RRM domains using profile
S1
Infer trees of RRM domains
Extract SR-associated RRM domains
CDD/KOG NCBI Taxonomy
Annotate RRM trees
Prune SR-associated subtrees
refs from literature
S2
Adobe
S5-9
S13-20
Compare RRM trees
Develop evolutionary scenario
S4
Fig. 2
S10-12
Tab. 1
S3-4
Fig. 3
CDD/Pfam
S19
S15
S17, S6
3
Build HMM profile for ZnK domain
Prune subtrees for SR families
Predict ZnK domains using profile
Extract proteins of SR families
Extract RRM domains of SR families
Fig. 4
S21
Extract ZnK domains of SR families
S24
Analyze protein composition
Structurally align RRM domains
Draw RRM logos for SR families
Align ZnK domains
Profile conservation of SR families
Profile composition of SR families
Draw ZnK logos for SR families
Fig. 5
S25
S23
Develop key for SR families
S5-6
exhaustive HMM profile alignments
S13
S15
S21
curated structural alignments
4
Build RRM profiles for SR proteins
20 selected (7 new) complete proteomes
DS1
Predict RRM domains using profiles
Extract predicted RRM proteins
Extract predicted RRM domains
Analyze RRM protein architecture
Analyze RRM protein composition
BLAST align RRM domains using logos
S21
DS3a
DS3b-d
Infer trees of RRM domains
CDD/KOG NCBI Taxonomy
Assemble inventory of SR proteins
Annotate RRM trees
PDF
S2
Adobe
refs from literature
Tab. 2
S7, DS2
S26-29


A. E-value threshold = 1E-05 / # RRM domains = 8471
1
0.5
Score
0
1
0.5
Coverage
0
0
0.5
Identity
1
1.00
0.95
0.90
0.85
0.80
0.75


B. E-value threshold = 1E-10 / # RRM domains = 6633
1
0.5
Score
0
1
0.5
Coverage
0
0
0.5
Identity
1
1.00
0.95
0.90
0.85
0.80
0.75


C. E-value threshold = 1E-15 / # RRM domains = 4523
1
0.5
Score
0
1
0.5
Coverage
0
0
0.5
Identity
1
1.00
0.95
0.90
0.85
0.80
0.75

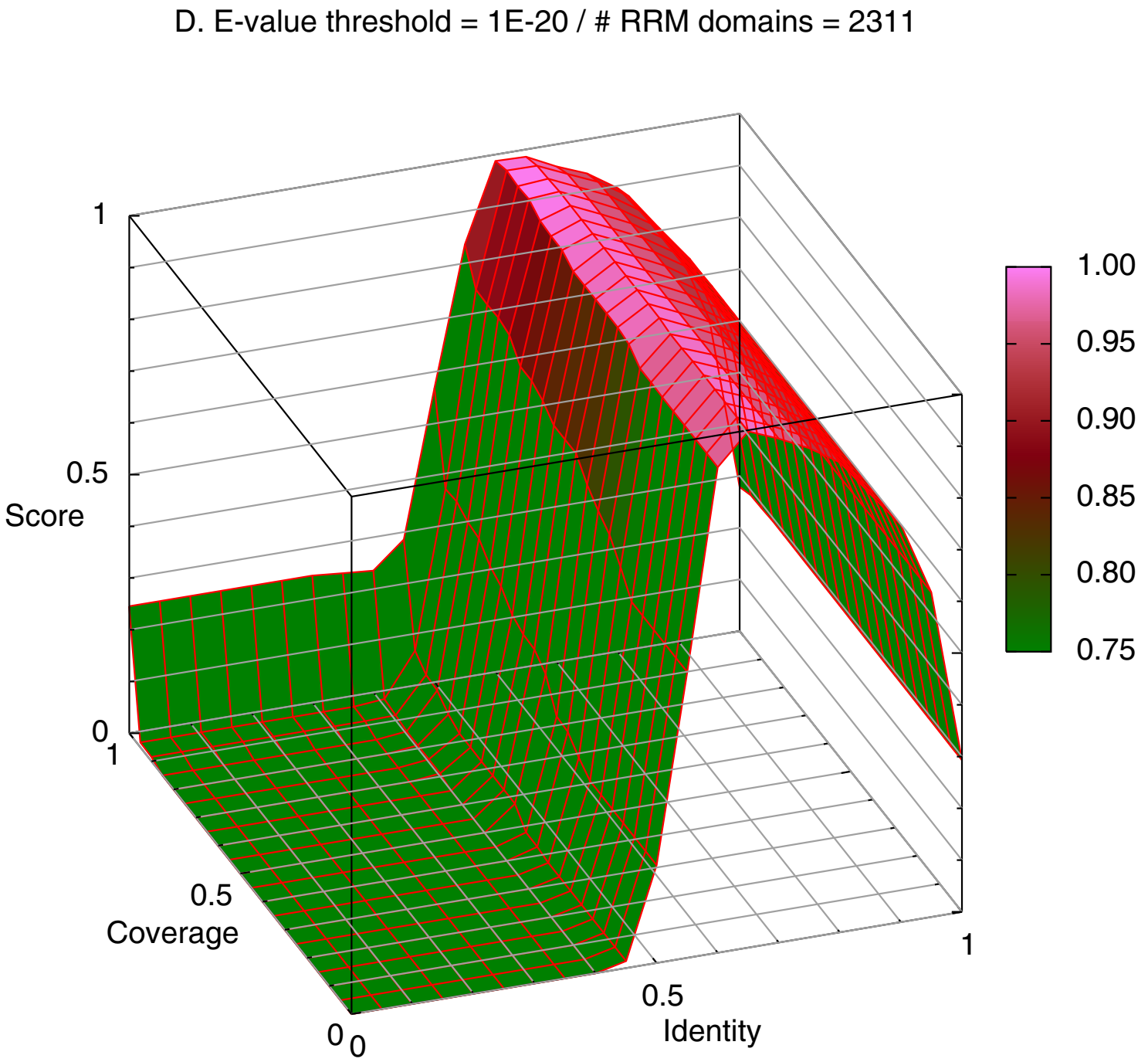

D. E-value threshold = 1E-20 / # RRM domains = 2311
1
0.5
Score
0
1
0.5
Coverage
0
0
0.5
Identity
1
1.00
0.95
0.90
0.85
0.80
0.75

Fig. S2

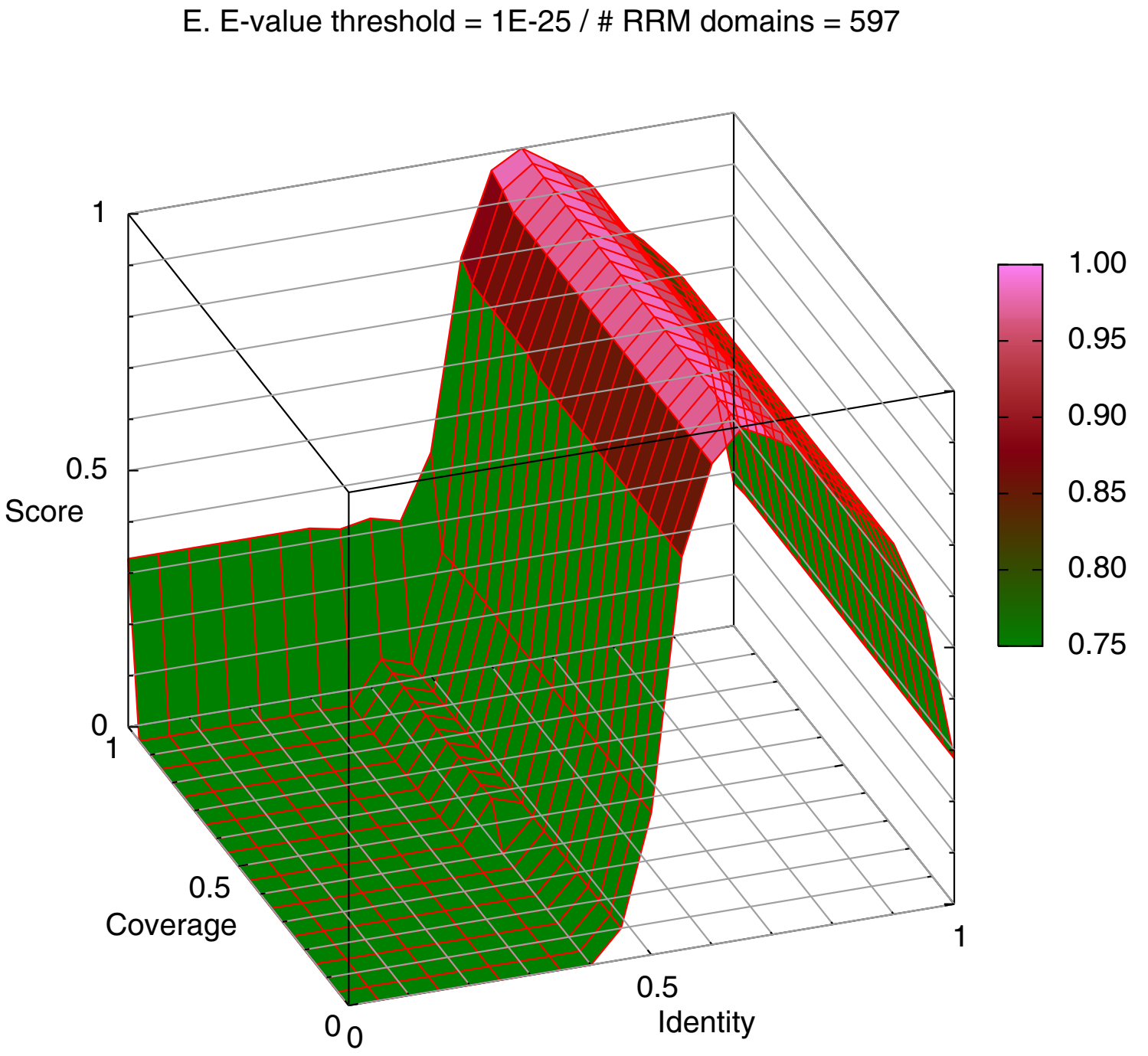

E. E-value threshold = 1E-25 / # RRM domains = 597
Score
Coverage
Identity
0
0.5
1
1.00
0.95
0.90
0.85
0.80
0.75

Fig. S3

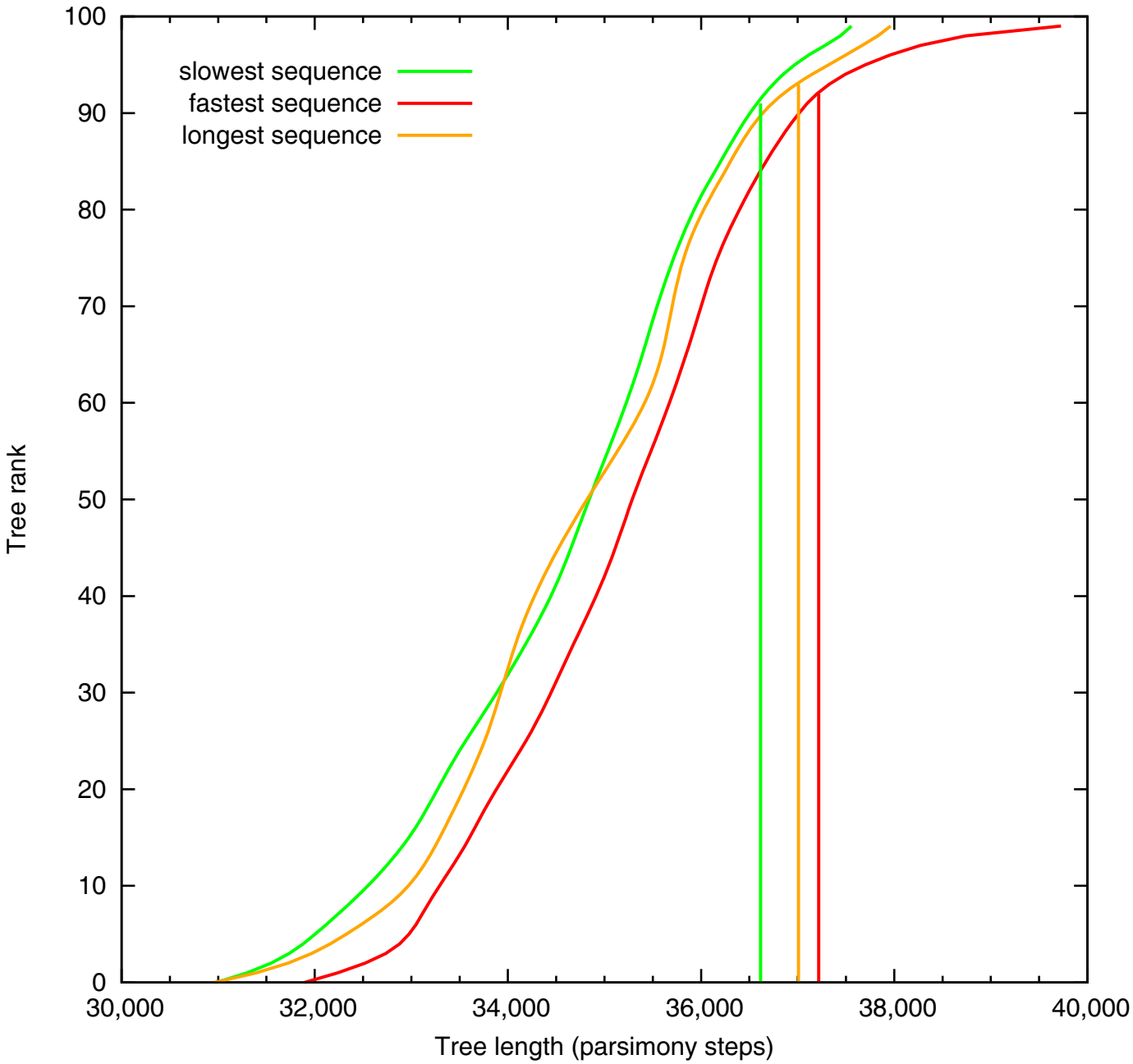

slowest sequence
fastest sequence
longest sequence
Tree rank
Tree length (parsimony steps)
0
10
20
30
40
50
60
70
80
90
100
30,000
32,000
34,000
36,000
38,000
40,000

Fig. S4

identifier	count	process	description
[KOG4205]	1554	[RNA processing and modification]	RNA-binding protein musashi/mRNA cleavage and polyadenylation factor I complex, subunit HRP1
[KOG0123]	1303	[RNA processing and modification, Translation, ribosomal structure and biogenesis]	Polyadenylate-binding protein (RRM superfamily)
[KOG0145]	972	[RNA processing and modification]	RNA-binding protein ELAV/HU (RRM superfamily)
[KOG0148]	534	[RNA processing and modification, Translation, ribosomal structure and biogenesis]	Apoptosis-promoting RNA-binding protein TIA-1/TIAR (RRM superfamily)
[KOG0117]	515	[RNA processing and modification]	Heterogeneous nuclear ribonucleoprotein R (RRM superfamily)
[KOG0144]	435	[RNA processing and modification]	RNA-binding protein CUGBP1/BRUNO (RRM superfamily)
[KOG0108]	359	[RNA processing and modification]	mRNA cleavage and polyadenylation factor I complex, subunit RNA15
[KOG4212]	328	[RNA processing and modification]	RNA-binding protein hnRNP-M
[KOG0146]	297	[RNA processing and modification]	RNA-binding protein ETR-3 (RRM superfamily)
[KOG0105]	292	[RNA processing and modification]	Alternative splicing factor ASF/SF2 (RRM superfamily)
[KOG0127]	246	[RNA processing and modification]	Nucleolar protein fibrillarin NOP77 (RRM superfamily)
[KOG0110]	233	[General function prediction only]	RNA-binding protein (RRM superfamily)
[KOG0115]	209	[RNA processing and modification]	RNA-binding protein p54nrb (RRM superfamily)
[KOG0131]	205	[RNA processing and modification]	Splicing factor 3b, subunit 4
[KOG0147]	203	[Transcription]	Transcriptional coactivator CAPER (RRM superfamily)
[KOG0109]	199	[RNA processing and modification, General function prediction only]	RNA-binding protein LARK, contains RRM and retroviral-type Zn-finger domains
[KOG0149]	178	[General function prediction only]	Predicted RNA-binding protein SEB4 (RRM superfamily)
[KOG0125]	167	[General function prediction only]	Ataxin 2-binding protein (RRM superfamily)
[KOG0106]	164	[RNA processing and modification]	Alternative splicing factor SRp55/B52/SRp75 (RRM superfamily)
[KOG0116]	149	[Signal transduction mechanisms]	RasGAP SH3 binding protein rasputin, contains NTF2 and RRM domains
[KOG4209]	145	[RNA processing and modification]	Splicing factor RNPS1, SR protein superfamily
[KOG0124]	144	[RNA processing and modification]	Polypyrimidine tract-binding protein PUF60 (RRM superfamily)
[KOG0111]	140	[Posttranslational modification, protein turnover, chaperones]	Cyclophilin-type peptidyl-prolyl cis-trans isomerase
[KOG4211]	135	[RNA processing and modification]	Splicing factor hnRNP-F and related RNA-binding proteins
[KOG0113]	123	[RNA processing and modification]	U1 small nuclear ribonucleoprotein (RRM superfamily)
[KOG0107]	113	[RNA processing and modification]	Alternative splicing factor SRp20/9G8 (RRM superfamily)
[KOG4206]	107	[RNA processing and modification]	Spliceosomal protein snRNP-U1A/U2B
[KOG0128]	106	[RNA processing and modification]	RNA-binding protein SART3 (RRM superfamily)
[KOG0122]	102	[Translation, ribosomal structure and biogenesis]	Translation initiation factor 3, subunit g (eIF-3g)
[KOG0112]	97	[General function prediction only]	Large RNA-binding protein (RRM superfamily)
[KOG0533]	94	[RNA processing and modification]	RRM motif-containing protein
[KOG4661]	93	[Transcription]	Hsp27-ERE-TATA-binding protein/Scaffold attachment factor (SAF-B)
[KOG0126]	88	[General function prediction only]	Predicted RNA-binding protein (RRM superfamily)
[KOG0120]	86	[RNA processing and modification]	Splicing factor U2AF, large subunit (RRM superfamily)
[KOG4208]	77	[General function prediction only]	Nucleolar RNA-binding protein NIFK
[KOG0415]	77	[Posttranslational modification, protein turnover, chaperones]	Predicted peptidyl prolyl cis-trans isomerase
[KOG0226]	76	[General function prediction only]	RNA-binding proteins
[KOG0121]	74	[RNA processing and modification]	Nuclear cap-binding protein complex, subunit CBP20 (RRM superfamily)
[KOG1995]	66	[General function prediction only]	Conserved Zn-finger protein
[KOG4207]	58	[RNA processing and modification]	Predicted splicing factor, SR protein superfamily
[KOG0114]	56	[General function prediction only]	Predicted RNA-binding protein (RRM superfamily)
[KOG0130]	56	[General function prediction only]	RNA-binding protein RBM8/Tsunagi (RRM superfamily)
[KOG0151]	53	[General function prediction only]	Predicted splicing regulator, contains RRM, SWAP and RPR domains
[KOG0921]	44	[Transcription]	Dosage compensation complex, subunit MLE
[KOG0153]	42	[General function prediction only]	Predicted RNA-binding protein (RRM superfamily)
[KOG4676]	39	[RNA processing and modification]	Splicing factor, arginine/serine-rich
[KOG4660]	34	[Cell cycle control, cell division, chromosome partitioning]	Protein Mei2, essential for commitment to meiosis, and related proteins
[KOG4454]	34	[General function prediction only]	RNA binding protein (RRM superfamily)
[KOG1855]	34	[General function prediction only]	Predicted RNA-binding protein
[KOG0132]	22	[RNA processing and modification, Transcription]	RNA polymerase II C-terminal domain-binding protein RA4, contains RPR and RRM domains
[KOG3598]	21	[Transcription]	Thyroid hormone receptor-associated protein complex, subunit TRAP230
[KOG4213]	19	[RNA processing and modification]	RNA-binding protein La
[KOG1190]	19	[RNA processing and modification]	Polypyrimidine tract-binding protein
[KOG4368]	19	[General function prediction only]	Predicted RNA binding protein, contains SWAP, RPR and G-patch domains
[KOG4210]	13	[Transcription]	Nuclear localization sequence binding protein
[KOG0548]	9	[Posttranslational modification, protein turnover, chaperones]	Molecular co-chaperone STI1
[KOG2253]	8	[RNA processing and modification]	U1 snRNP complex, subunit SNU71 and related PWI-motif proteins
[KOG0154]	8	[General function prediction only]	RNA-binding protein RBM5 and related proteins, contain G-patch and RRM domains
[KOG2193]	8	[RNA processing and modification, General function prediction only]	IGF-II mRNA-binding protein IMP, contains RRM and KH domains
[KOG0670]	7	[RNA processing and modification]	U4/U6-associated splicing factor PRP4
[KOG1015]	6	[Transcription]	Transcription regulator XNP/ATRX, DEAD-box superfamily
[KOG0119]	6	[RNA processing and modification]	Splicing factor 1/branch point binding protein (RRM superfamily)
[KOG1080]	6	[Chromatin structure and dynamics, Transcription]	Histone H3 (Lys4) methyltransferase complex, subunit SET1 and related methyltransferases
[KOG2277]	6	[Cell cycle control, cell division, chromosome partitioning]	S-M checkpoint control protein CID1 and related nucleotidyltransferases
[KOG1240]	5	[Signal transduction mechanisms]	Protein kinase containing WD40 repeats
[KOG3702]	5	[RNA processing and modification]	Nuclear polyadenylated RNA binding protein
[KOG1548]	4	[Transcription]	Transcription elongation factor TAT-SF1
[KOG2044]	4	[Replication, recombination and repair, RNA processing and modification]	5'-3' exonuclease HKE1/RAT1
[KOG0576]	4	[Signal transduction mechanisms]	Mitogen-activated protein kinase kinase kinase kinase (MAP4K), germinal center kinase family
[KOG1365]	4	[RNA processing and modification, General function prediction only]	RNA-binding protein Fusilli, contains RRM domain
[KOG0598]	4	[General function prediction only, Signal transduction mechanisms]	Ribosomal protein S6 kinase and related proteins
[KOG4574]	4	[General function prediction only]	RNA-binding protein (contains RRM and Pumilio-like repeats)
[KOG1984]	3	[Intracellular trafficking, secretion, and vesicular transport]	Vesicle coat complex COPII, subunit SFB3
[KOG2248]	3	[Replication, recombination and repair]	3'-5' exonuclease
[KOG4849]	3	[RNA processing and modification]	mRNA cleavage factor I subunit/CPSF subunit
[KOG1144]	3	[Translation, ribosomal structure and biogenesis]	Translation initiation factor 5B (eIF-5B)
[KOG0055]	3	[Secondary metabolites biosynthesis, transport and catabolism]	Multidrug/pheromone exporter, ABC superfamily
[KOG3257]	2	[Translation, ribosomal structure and biogenesis]	Mitochondrial/chloroplast ribosomal protein L11
[KOG1833]	2	[Nuclear structure, Intracellular trafficking, secretion, and vesicular transport]	Nuclear pore complex, gp210 component
[KOG2141]	2	[Signal transduction mechanisms]	Protein involved in high osmolarity signaling pathway
[KOG3001]	2	[Chromatin structure and dynamics, Transcription]	Dosage compensation regulatory complex/histone acetyltransferase complex, subunit MSL-3/MRG15/EAF3, and related CHROMO domain-containing proteins
[KOG4307]	2	[General function prediction only]	RNA binding protein RBM12/SWAN
[KOG3070]	2	[Translation, ribosomal structure and biogenesis]	Predicted RNA-binding protein containing PIN domain and invovled in translation or RNA processing
[KOG0331]	2	[RNA processing and modification]	ATP-dependent RNA helicase
[KOG3766]	2	[Transcription]	Polycomb group protein SCM/L(3)MBT (tumor-supressor in Drosophila and humans)
[KOG1457]	2	[General function prediction only]	RNA binding protein (contains RRM repeats)
[KOG1274]	2	[General function prediction only]	WD40 repeat protein
[KOG1832]	2	[Cell cycle control, cell division, chromosome partitioning]	HIV-1 Vpr-binding protein
[KOG1178]	2	[Secondary metabolites biosynthesis, transport and catabolism]	Non-ribosomal peptide synthetase/alpha-aminoadipate reductase and related enzymes
[KOG2087]	1	[Signal transduction mechanisms]	Glycoprotein hormone receptor
[KOG2476]	1	[Function unknown]	Uncharacterized conserved protein
[KOG1860]	1	[Intracellular trafficking, secretion, and vesicular transport, Cell cycle control, cell division, chromosome partitioning]	Nuclear protein export factor
[KOG2186]	1	[Cell cycle control, cell division, chromosome partitioning]	Cell growth-regulating nucleolar protein
[KOG0899]	1	[Translation, ribosomal structure and biogenesis]	Mitochondrial/chloroplast ribosomal protein S19
[KOG0941]	1	[Posttranslational modification, protein turnover, chaperones]	E3 ubiquitin protein ligase
[KOG2029]	1	[Function unknown]	Uncharacterized conserved protein
[KOG4177]	1	[Cell wall/membrane/envelope biogenesis]	Ankyrin
[KOG0462]	1	[Translation, ribosomal structure and biogenesis]	Elongation factor-type GTP-binding protein
[KOG1924]	1	[Signal transduction mechanisms, Cytoskeleton]	RhoA GTPase effector DIA/Diaphanous
[KOG0433]	1	[Translation, ribosomal structure and biogenesis]	Isoleucyl-tRNA synthetase
[KOG0585]	1	[Signal transduction mechanisms]	Ca2+/calmodulin-dependent protein kinase kinase beta and related serine/threonine protein kinases
[KOG0687]	1	[Posttranslational modification, protein turnover, chaperones]	26S proteasome regulatory complex, subunit RPN7/PSMD6
[KOG0591]	1	[Cell cycle control, cell division, chromosome partitioning]	NIMA (never in mitosis)-related G2-specific serine/threonine protein kinase
[KOG1311]	1	[General function prediction only]	DHHC-type Zn-finger proteins
[KOG0475]	1	[Inorganic ion transport and metabolism]	Cl- channel CLC-3 and related proteins (CLC superfamily)
[KOG0197]	1	[Signal transduction mechanisms]	Tyrosine kinases
[KOG4246]	1	[General function prediction only]	Predicted DNA-binding protein, contains SAP domain
[KOG0474]	1	[Inorganic ion transport and metabolism]	Cl- channel CLC-7 and related proteins (CLC superfamily)
[KOG2553]	1	[Translation, ribosomal structure and biogenesis]	Pseudouridylate synthase
[KOG1267]	1	[Transcription, General function prediction only]	Mitochondrial transcription termination factor, mTERF
[KOG0161]	1	[Cytoskeleton]	Myosin class II heavy chain
[KOG1337]	1	[General function prediction only]	N-methyltransferase
[KOG0307]	1	[Intracellular trafficking, secretion, and vesicular transport]	Vesicle coat complex COPII, subunit SEC31
[KOG1611]	1	[General function prediction only]	Predicted short chain-type dehydrogenase
[KOG2490]	1	[Function unknown]	Predicted membrane protein
[KOG4198]	1	[General function prediction only]	RNA-binding Ran Zn-finger protein and related proteins
[KOG3208]	1	[Intracellular trafficking, secretion, and vesicular transport]	SNARE protein GS28

Fig. S5

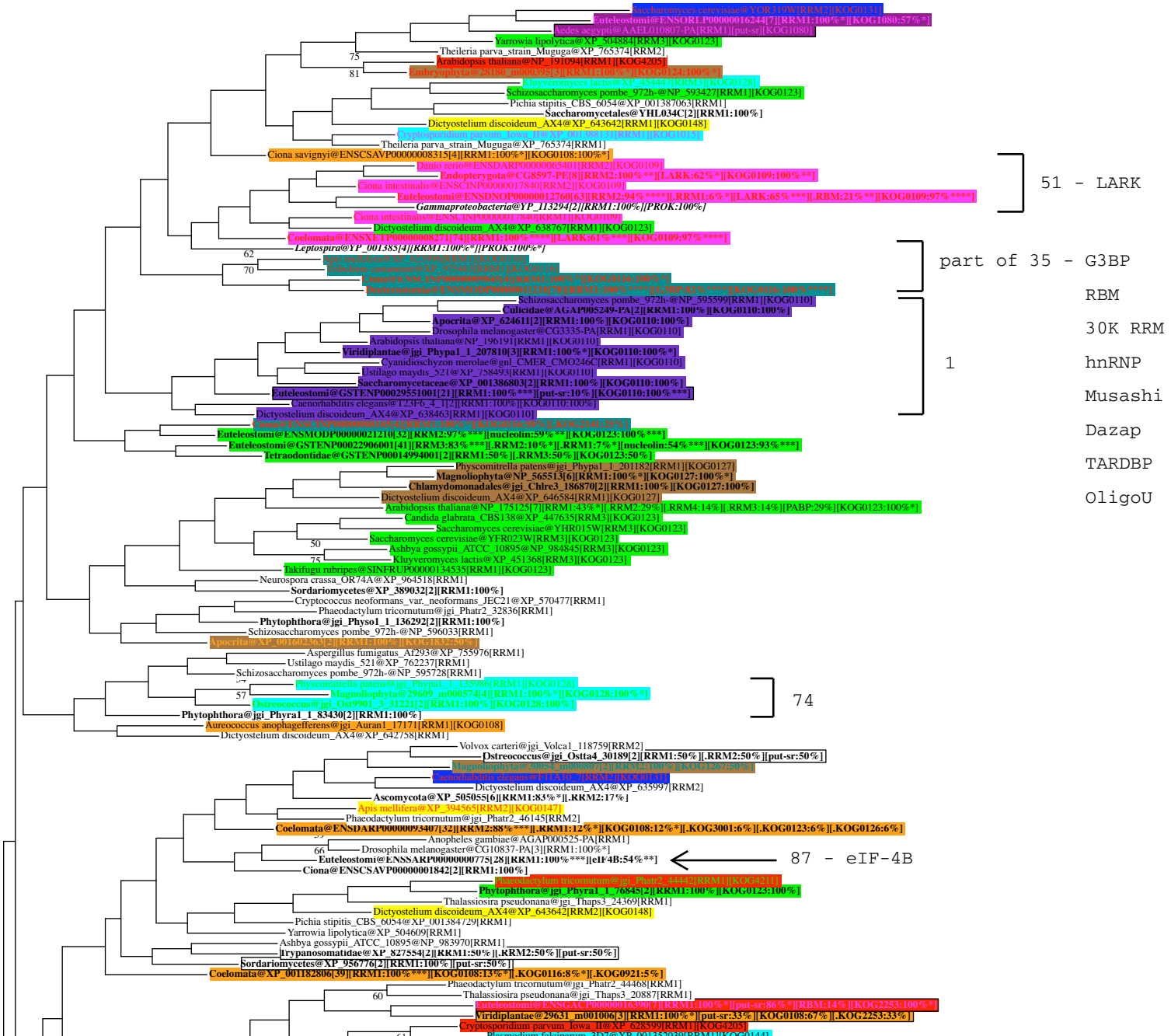
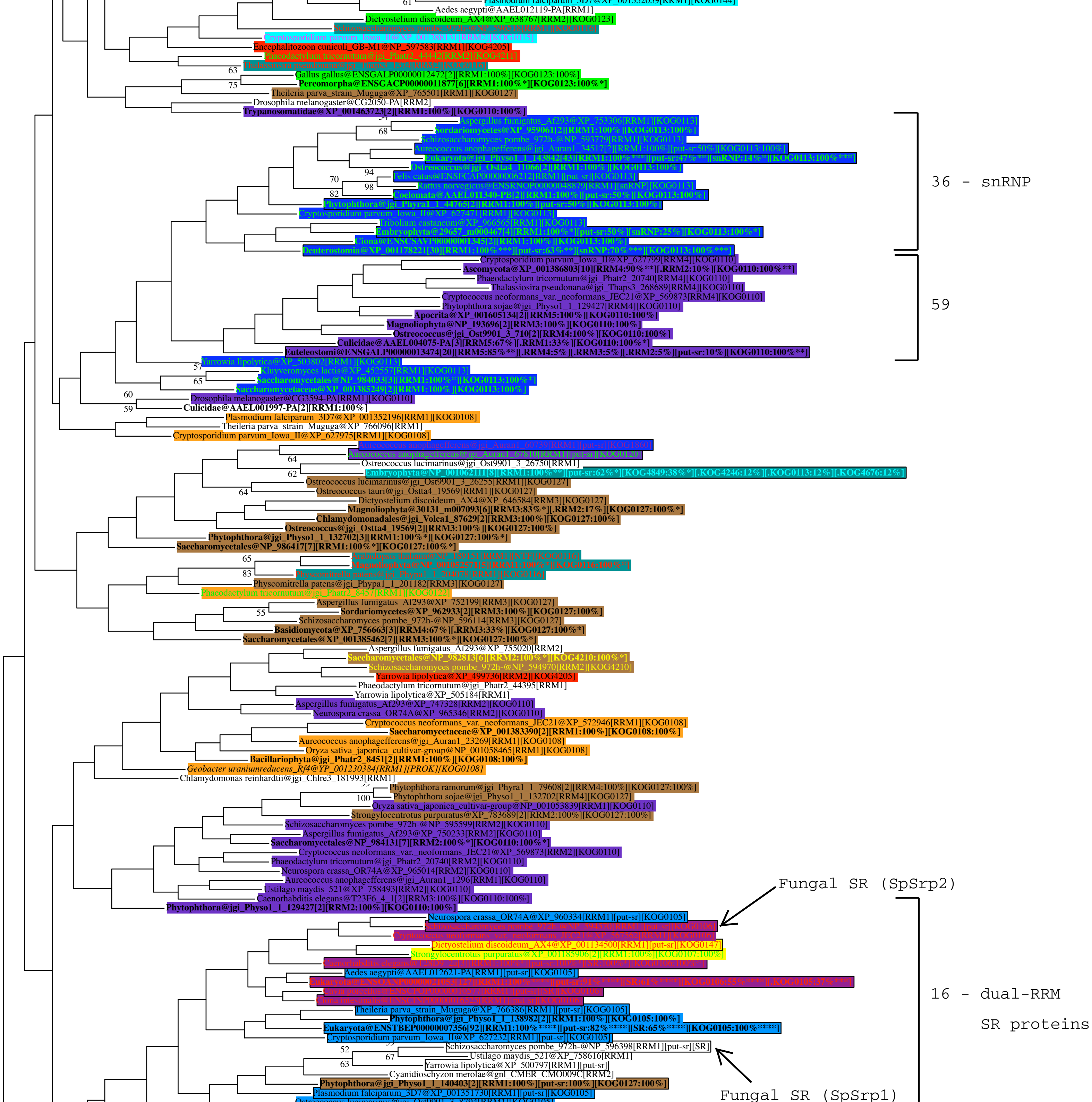

51 - LARK
part of 35 - G3BP
RBM
30K RRM
1
hnRNP
Musashi
Dazap
TARDBP
OligoU
74
87 - eIF-4B

Fig. S5



36 - snRNP
59
Fungal SR (SpSrp2)
16 - dual-RRM
SR proteins
Fungal SR (SpSrp1)

Fig. S5

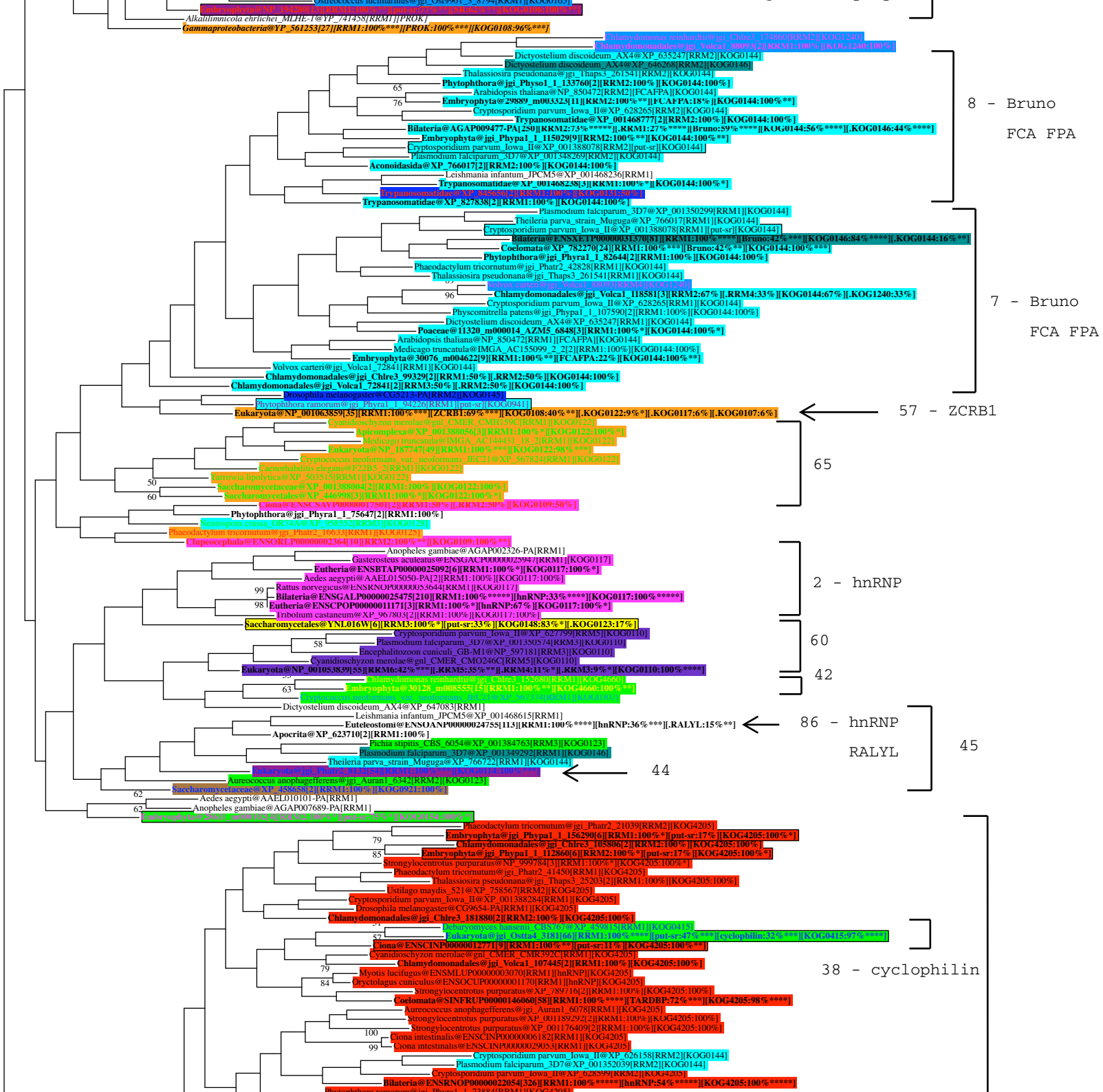

8 - Bruno
FCA FPA
7 - Bruno
FCA FPA
57 - ZCRB1
65
2 - hnRNP
60
42
86 - hnRNP
RALYL
45
44
38 - cyclophilin

Fig. S5

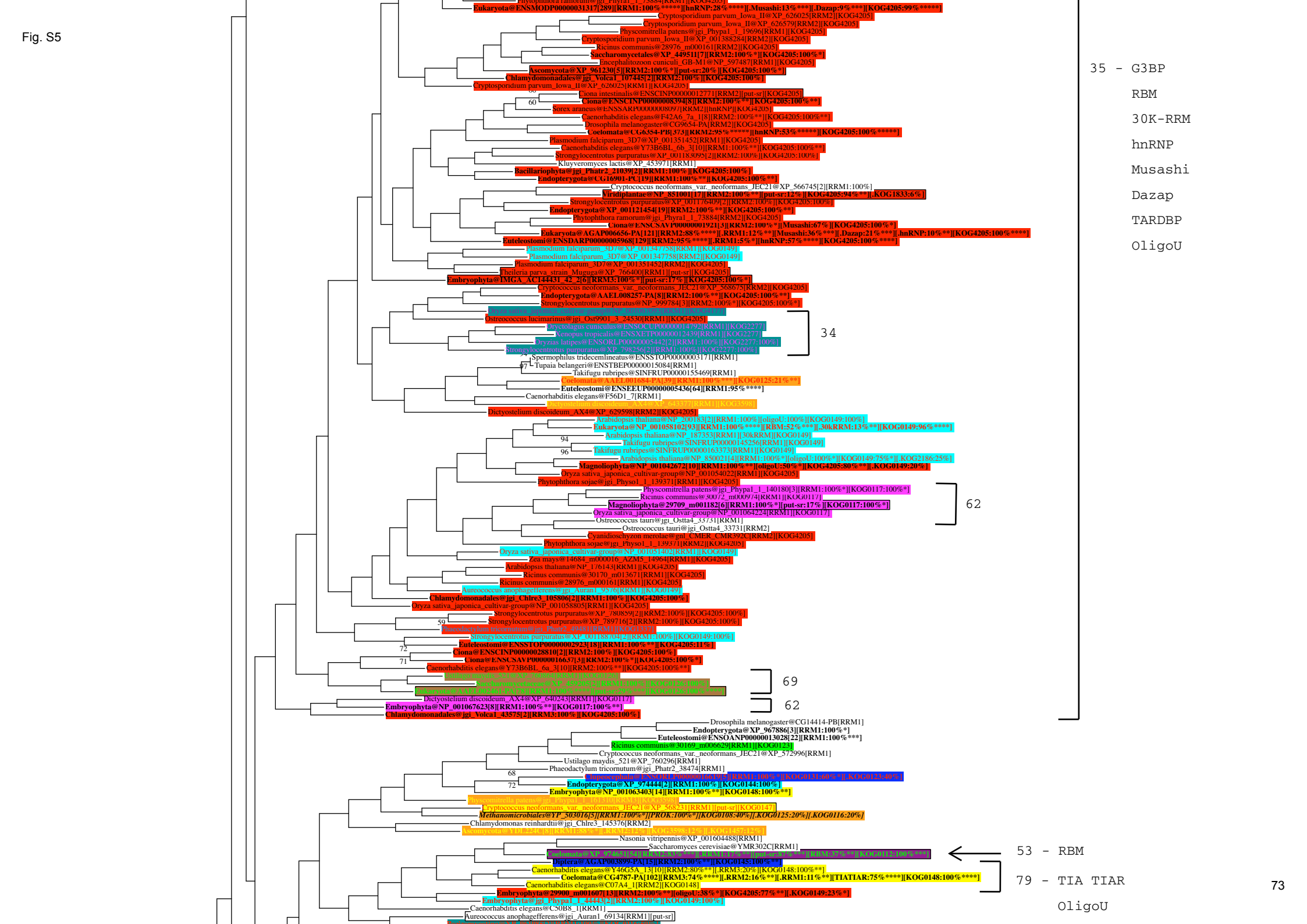

35 - G3BP
RBM
30K-RRM
hnRNP
Musashi
Dazap
TARDBP
OligoU
34
62
69
62
53 - RBM
79 - TIA TIAR
OligoU

Fig. S5

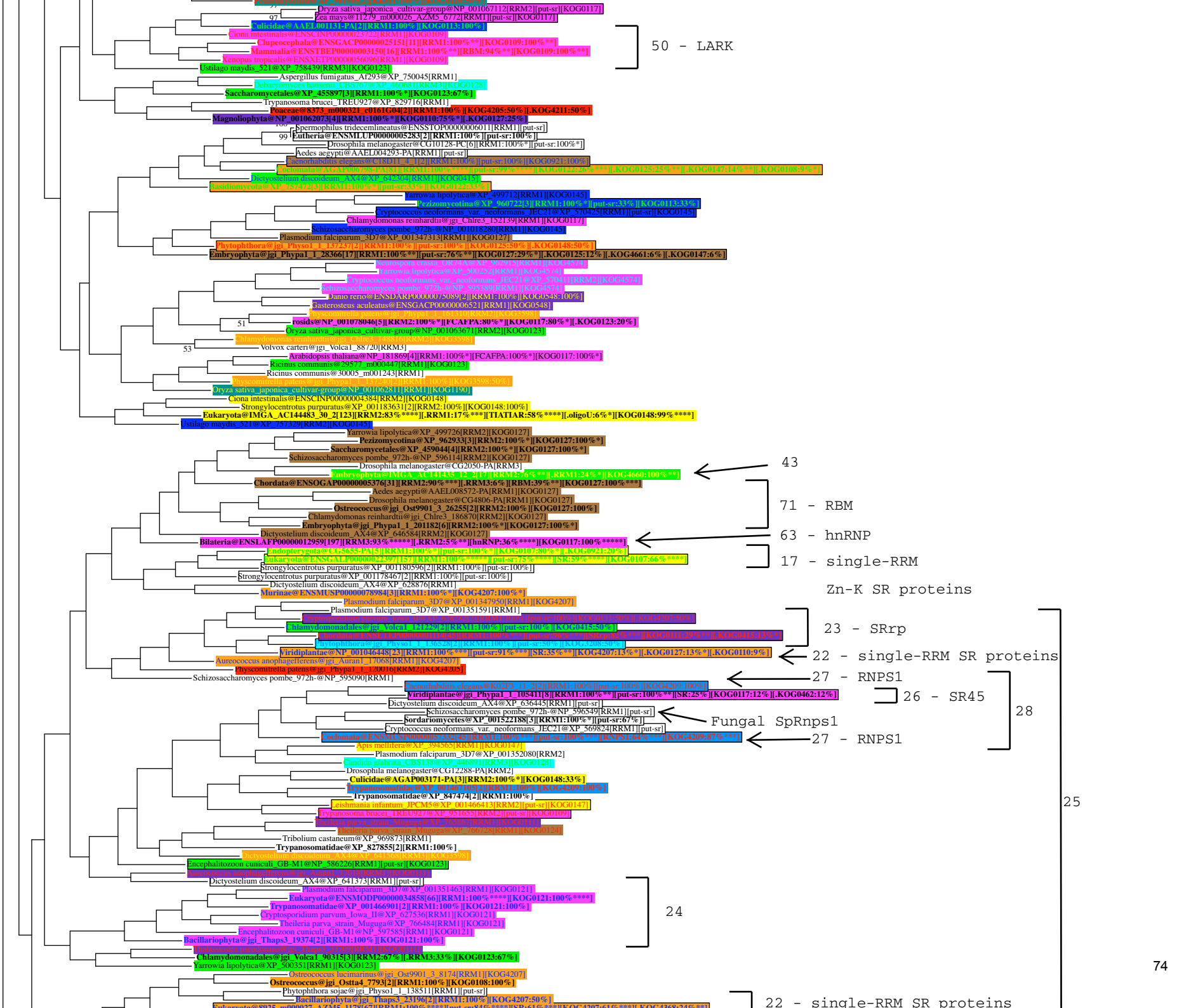

50 - LARK
43
71 - RBM
63 - hnRNP
17 - single-RRM
Zn-K SR proteins
23 - SRrp
22 - single-RRM SR proteins
27 - RNPS1
26 - SR45
28
Fungal SpRnps1
27 - RNPS1
25
24
22 - single-RRM SR proteins

Fig. S5

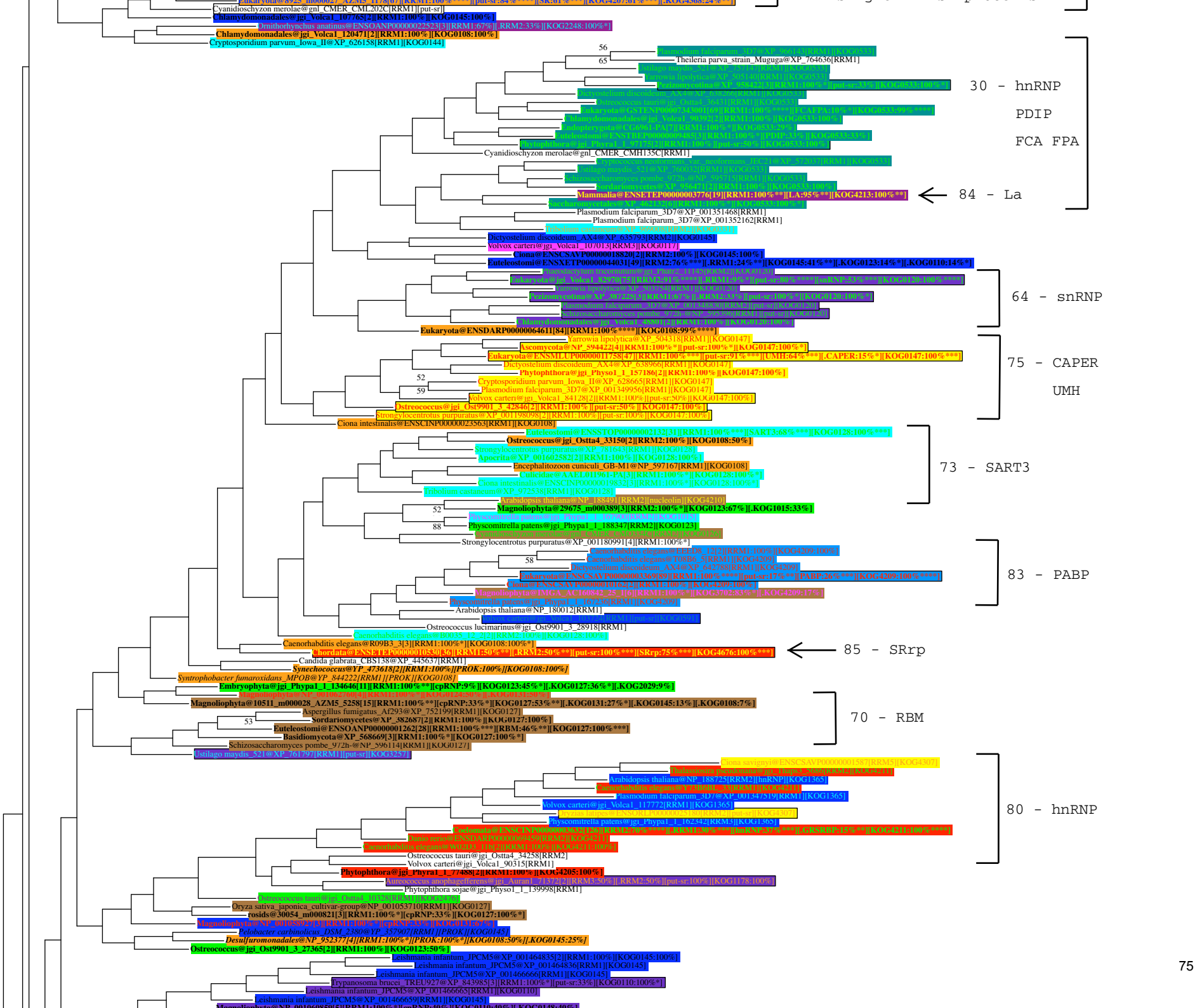

30 - hnRNP PDIP FCA FPA
84 - La
64 - snRNP
75 - CAPER UMH
73 - SART3
83 - PABP
85 - SRrp
70 - RBM
80 - hnRNP

Fig. S5

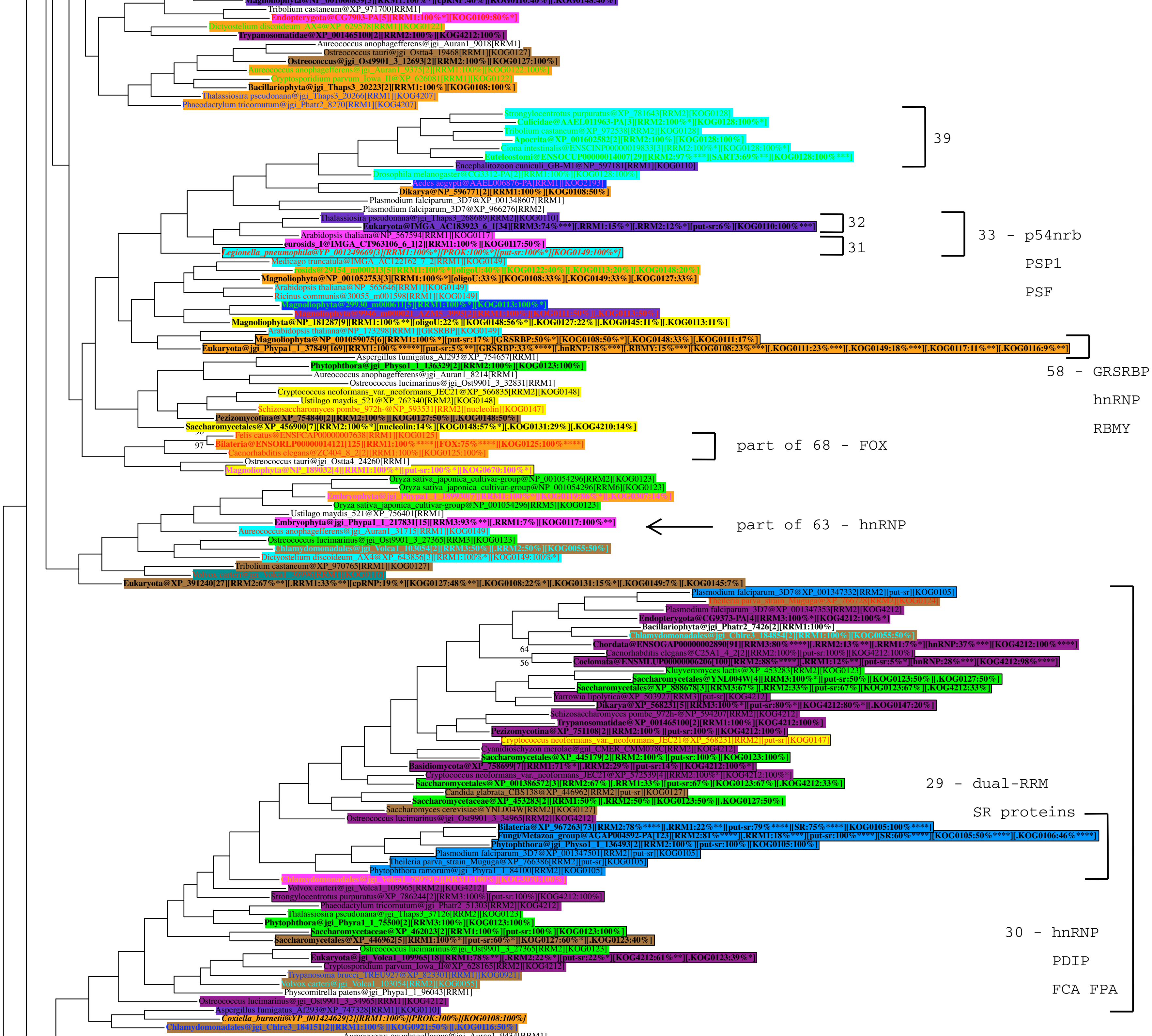

39
32
31
33 - p54nrb
PSP1
PSF
58 - GRSRBP
hnRNP
RBMY
part of 68 - FOX
part of 63 - hnRNP
29 - dual-RRM
SR proteins
30 - hnRNP
PDIP
FCA FPA

Fig. S5

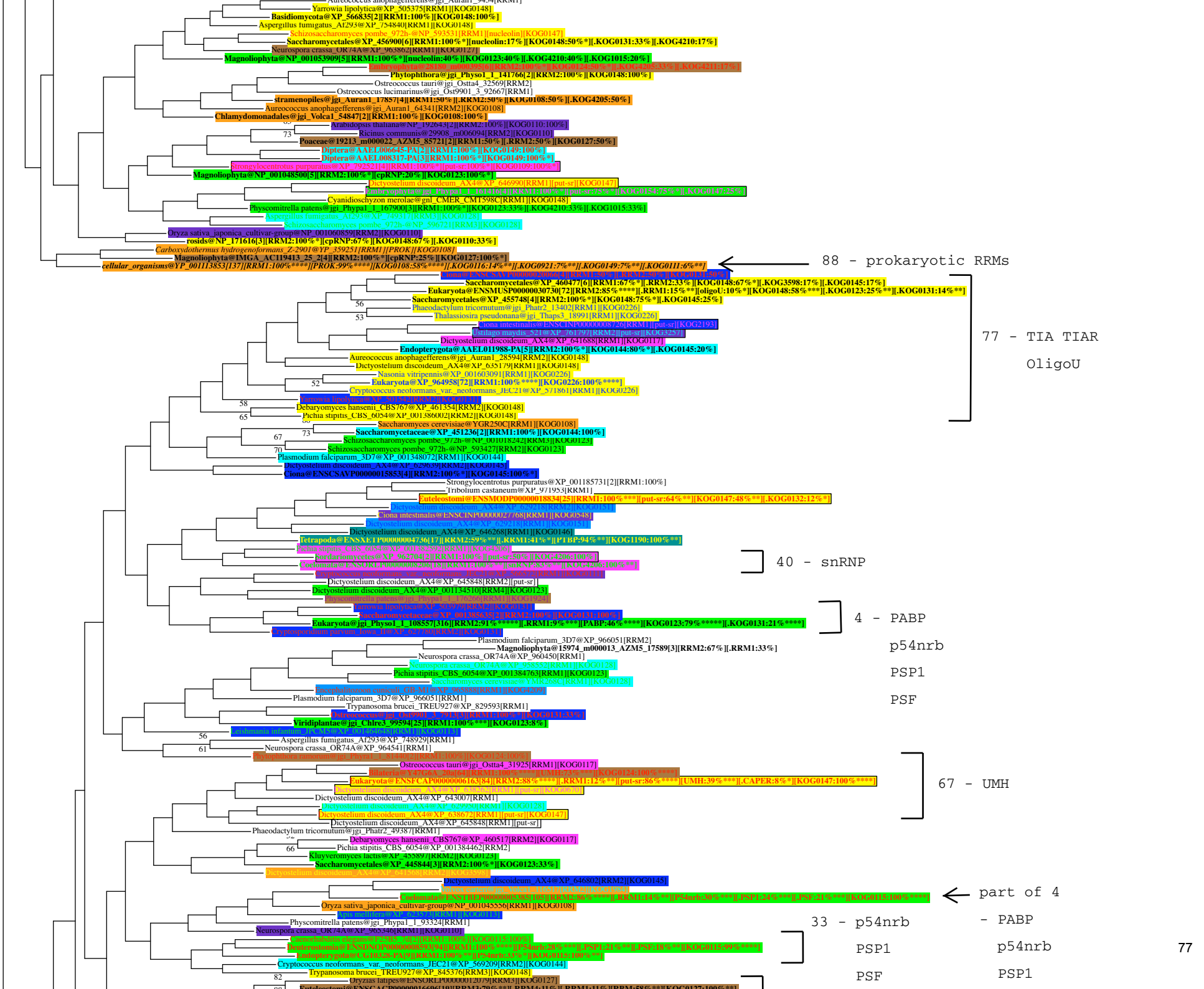

88 - prokaryotic RRMs
77 - TIA TIAR OligoU
40 - snRNP
4 - PABP p54nrb PSP1 PSF
67 - UMH
part of 4 - PABP p54nrb PSP1
33 - p54nrb PSP1 PSF

Fig. S5

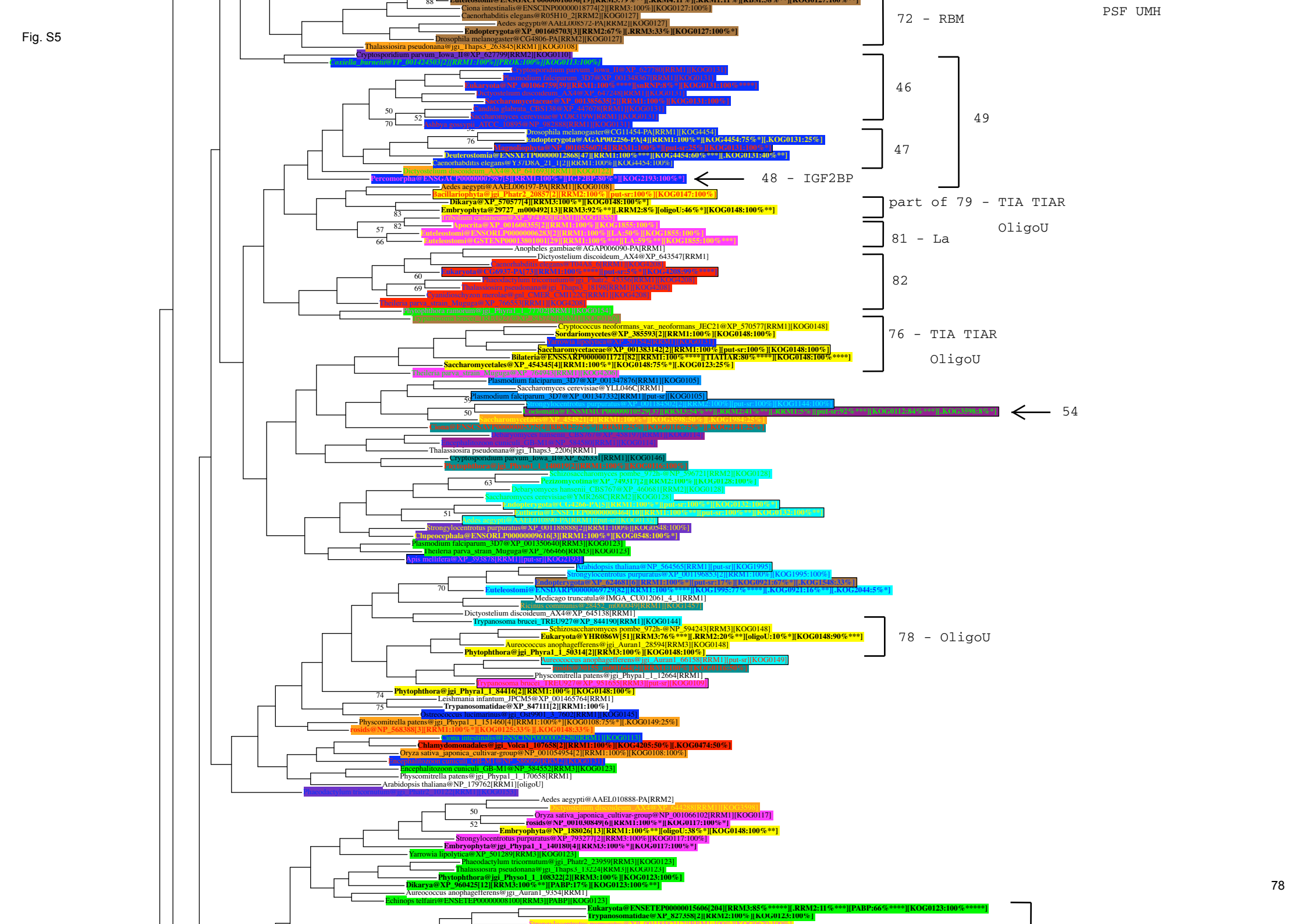

PSF UMH
72 - RBM
46
49
47
48 - IGF2BP
part of 79 - TIA TIAR
OligoU
81 - La
82
76 - TIA TIAR
OligoU
54
78 - OligoU

Fig. S5

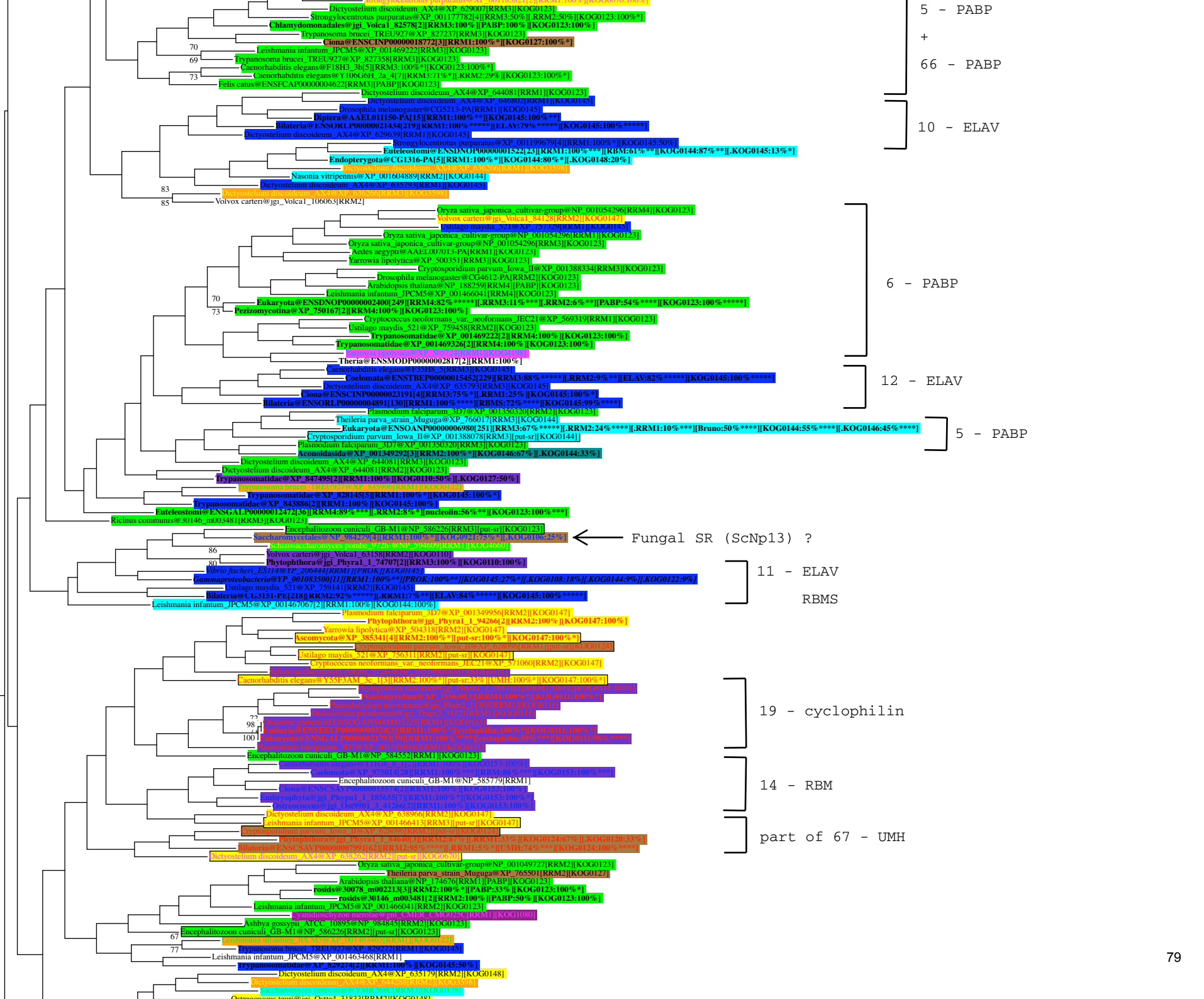

5 - PABP
+
66 - PABP
10 - ELAV
6 - PABP
12 - ELAV
5 - PABP
Fungal SR (ScNp13) ?
11 - ELAV
RBMS
19 - cyclophilin
14 - RBM
part of 67 - UMH

Fig. S5

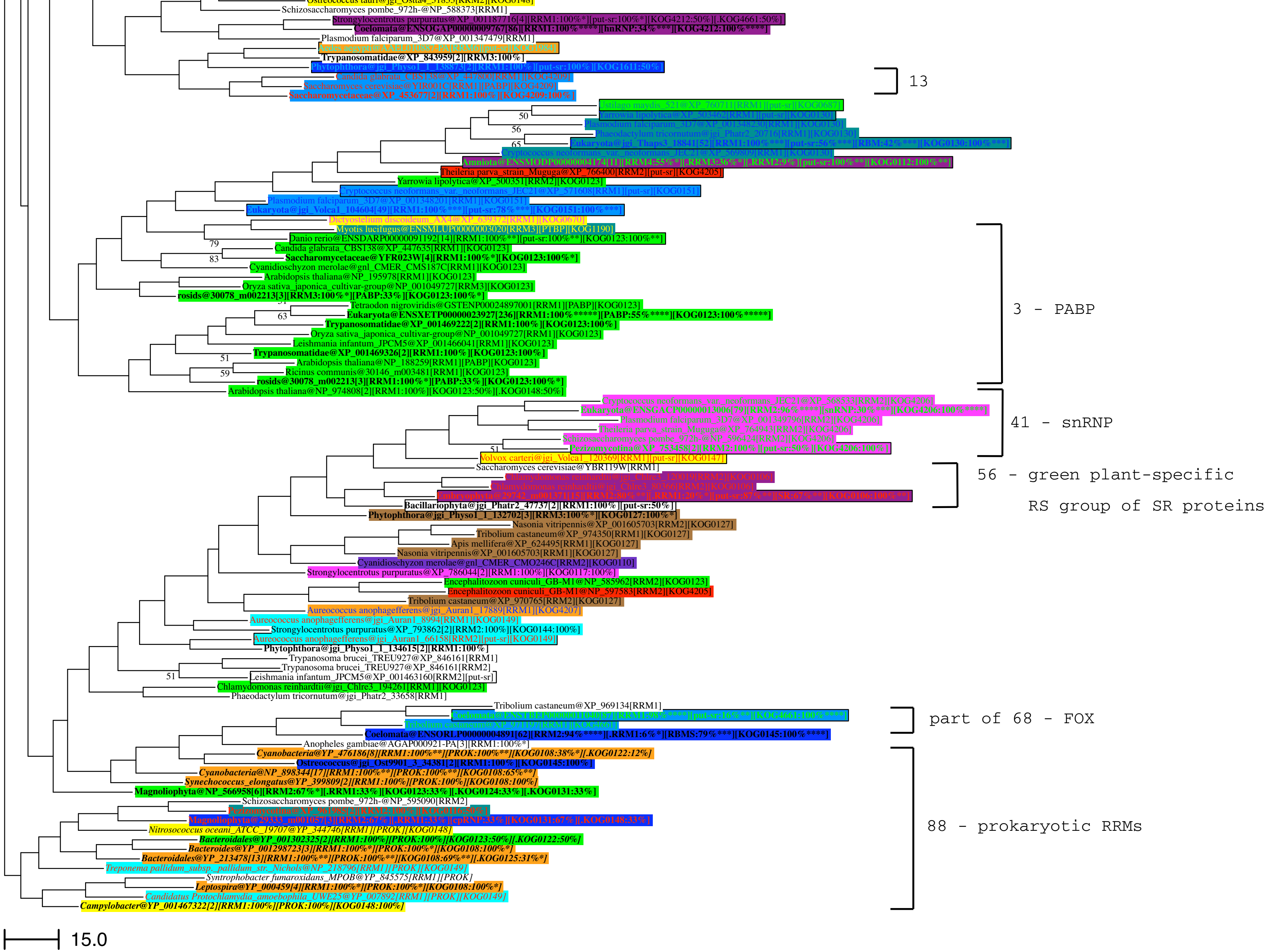
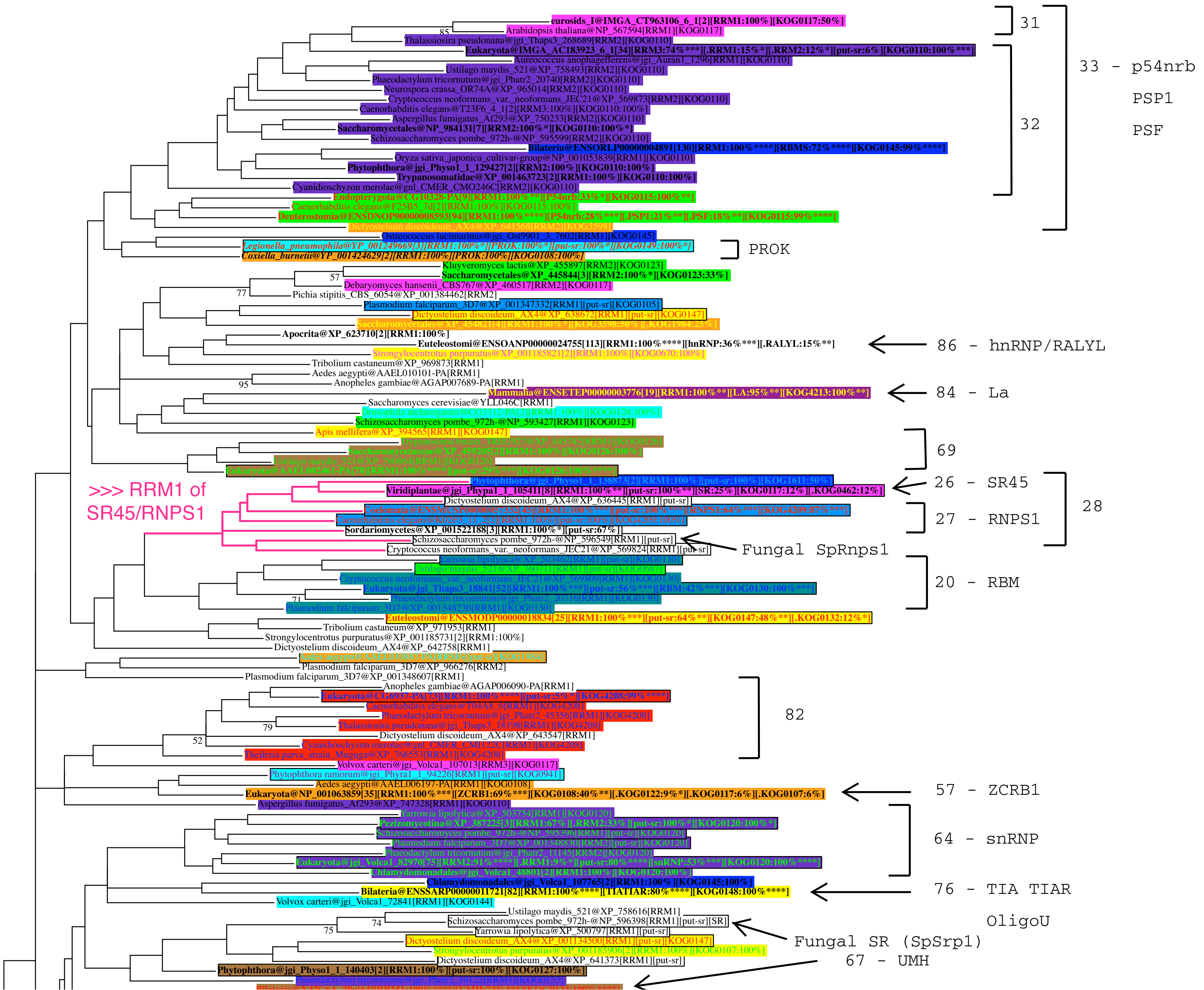

13
3 - PABP
41 - snRNP
56 - green plant-specific
RS group of SR proteins
part of 68 - FOX
88 - prokaryotic RRMs
15.0

Fig. S6



31
33 - p54nrb
PSP1
PSF
32
PROK
86 - hnRNP/RALYL
84 - La
69
>>> RRM1 of
SR45/RNPS1
26 - SR45
28
27 - RNPS1
Fungal SpRnps1
20 - RBM
82
57 - ZCRB1
64 - snRNP
76 - TIA TIAR
OligoU
Fungal SR (SpSrp1)
67 - UMH

Fig. S6

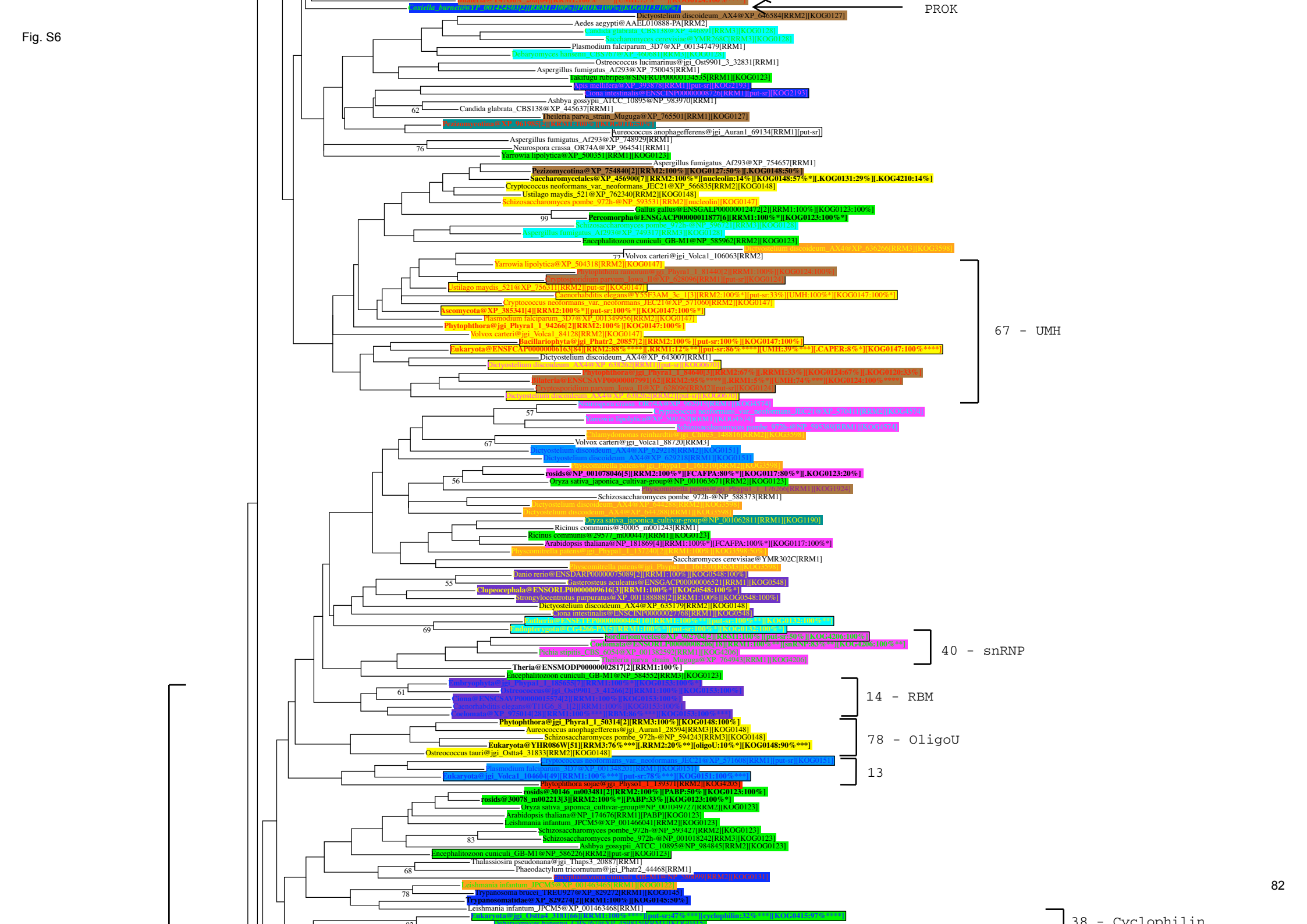
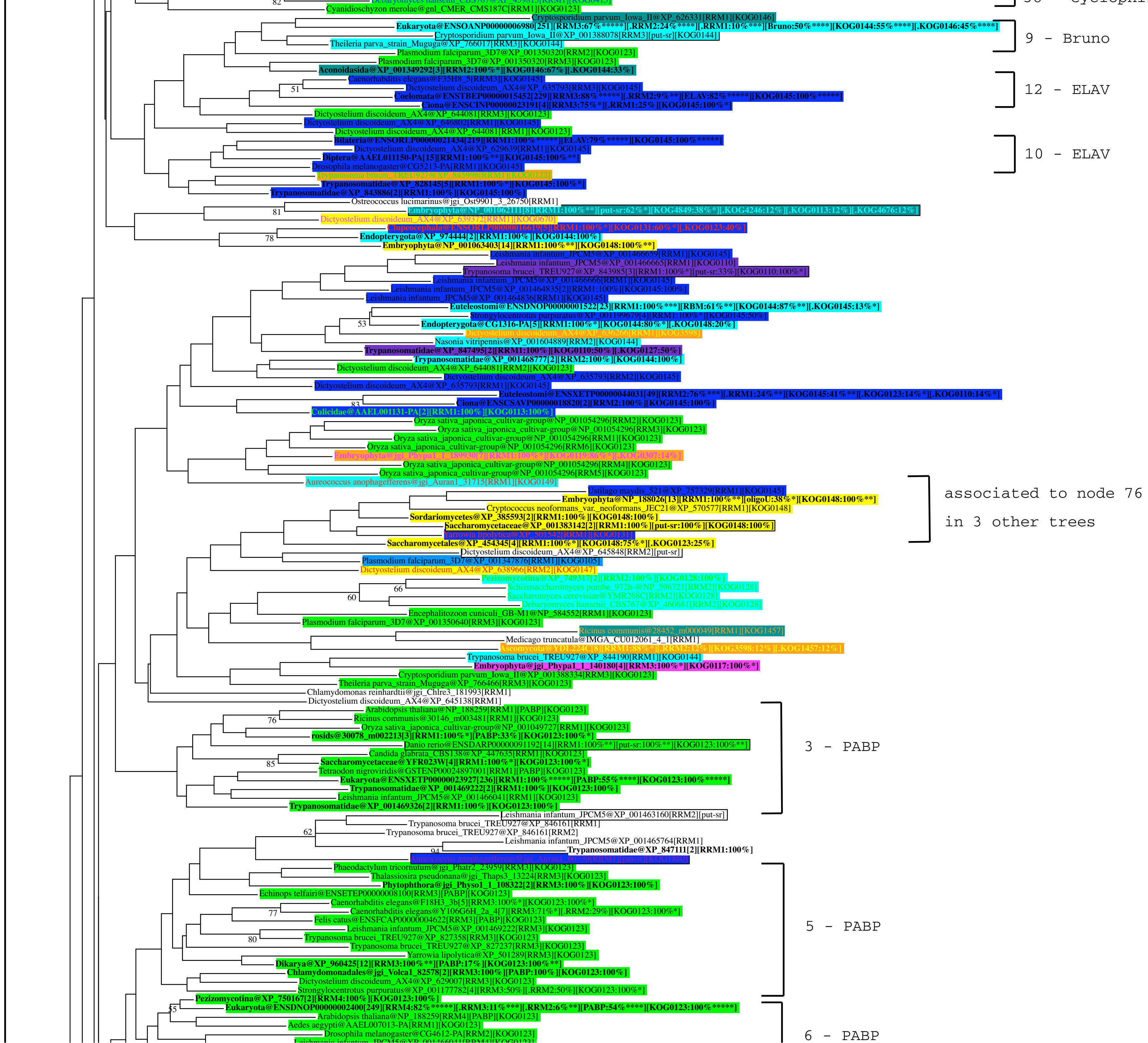


Fig. S6



9 - Bruno
12 - ELAV
10 - ELAV
associated to node 76
in 3 other trees
3 - PABP
5 - PABP
6 - PABP

Fig. S6

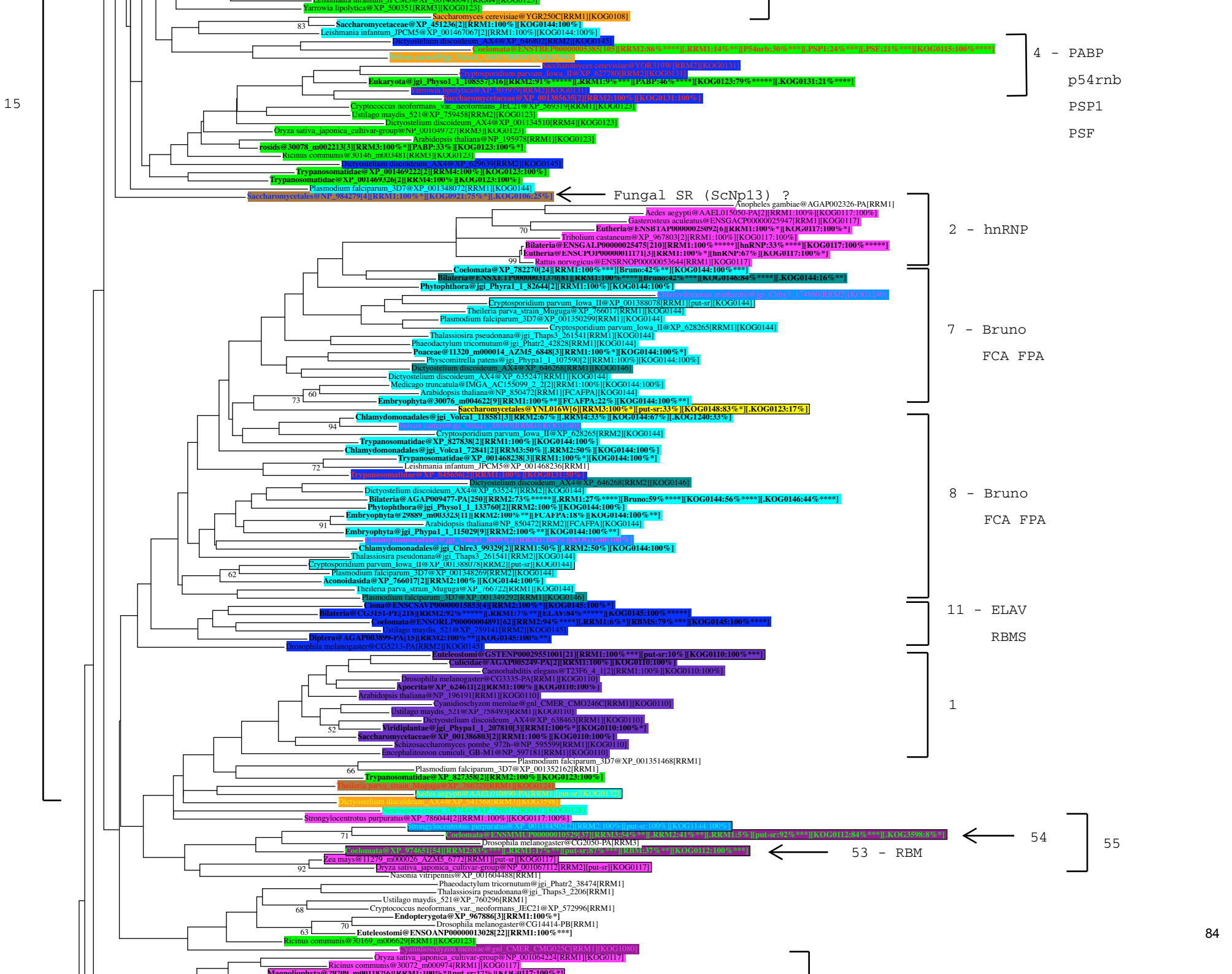

15
4 - PABP
p54rnb
PSP1
PSF
Fungal SR (ScNp13) ?
2 - hnRNP
7 - Bruno
FCA FPA
8 - Bruno
FCA FPA
11 - ELAV
RBMS
1
54
53 - RBM
55

Fig. S6

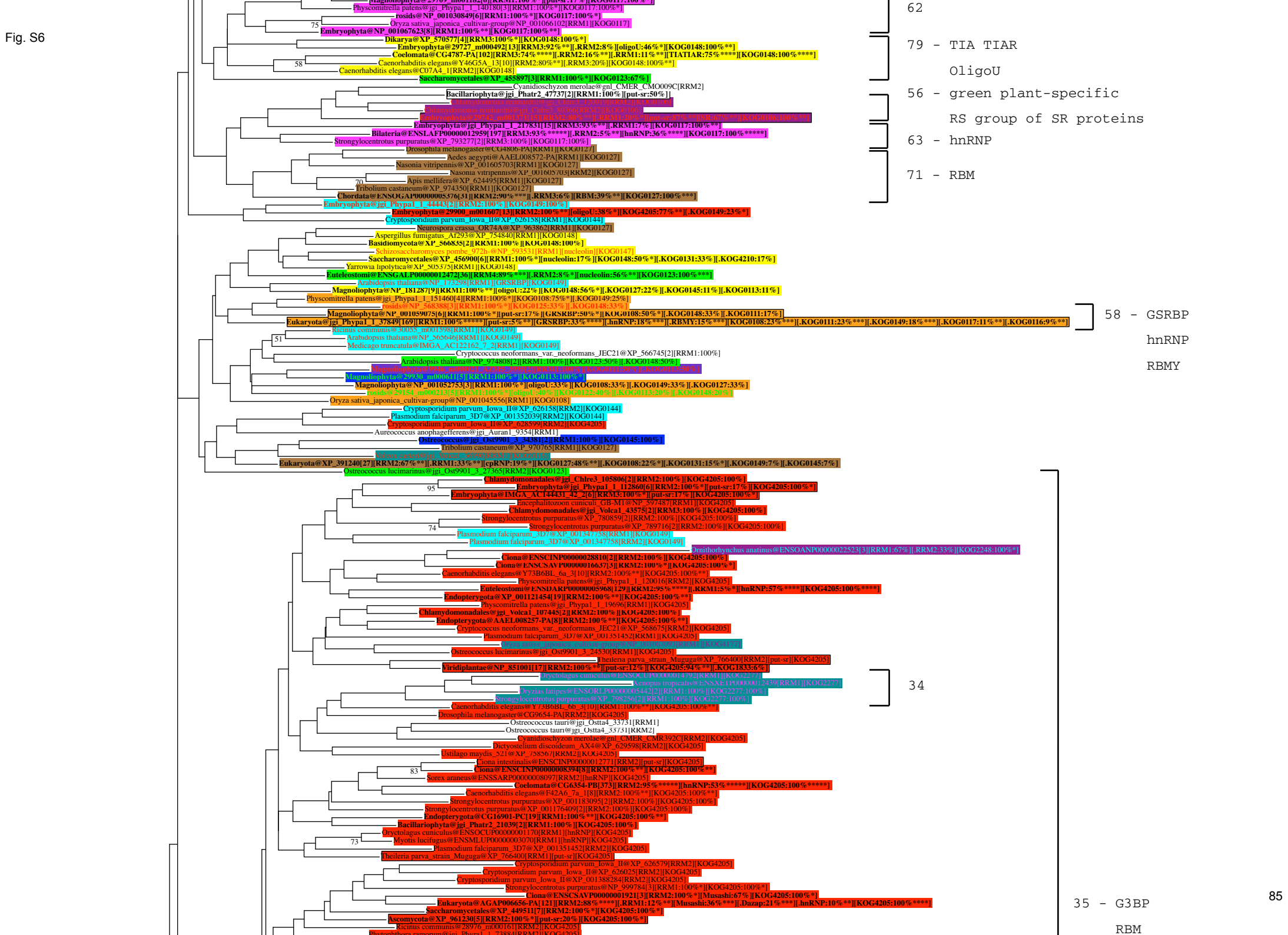

62
79 - TIA TIAR
OligoU
56 - green plant-specific
RS group of SR proteins
63 - hnRNP
71 - RBM
58 - GSRBP
hnRNP
RBMY
34
35 - G3BP
RBM

Fig. S6

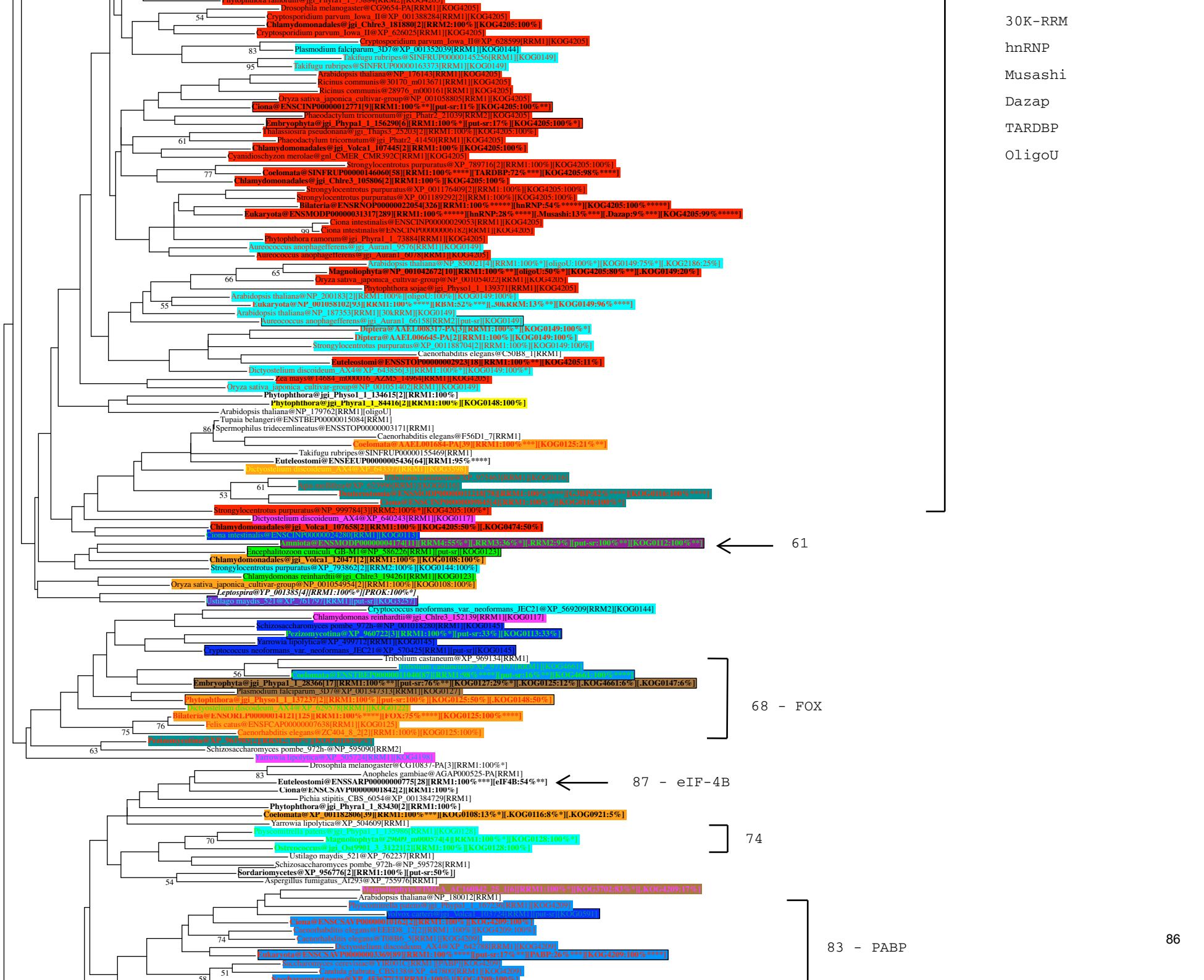

30K-RRM
hnRNP
Musashi
Dazap
TARDBP
OligoU
61
68 - FOX
87 - eIF-4B
74
83 - PABP

Fig. S6

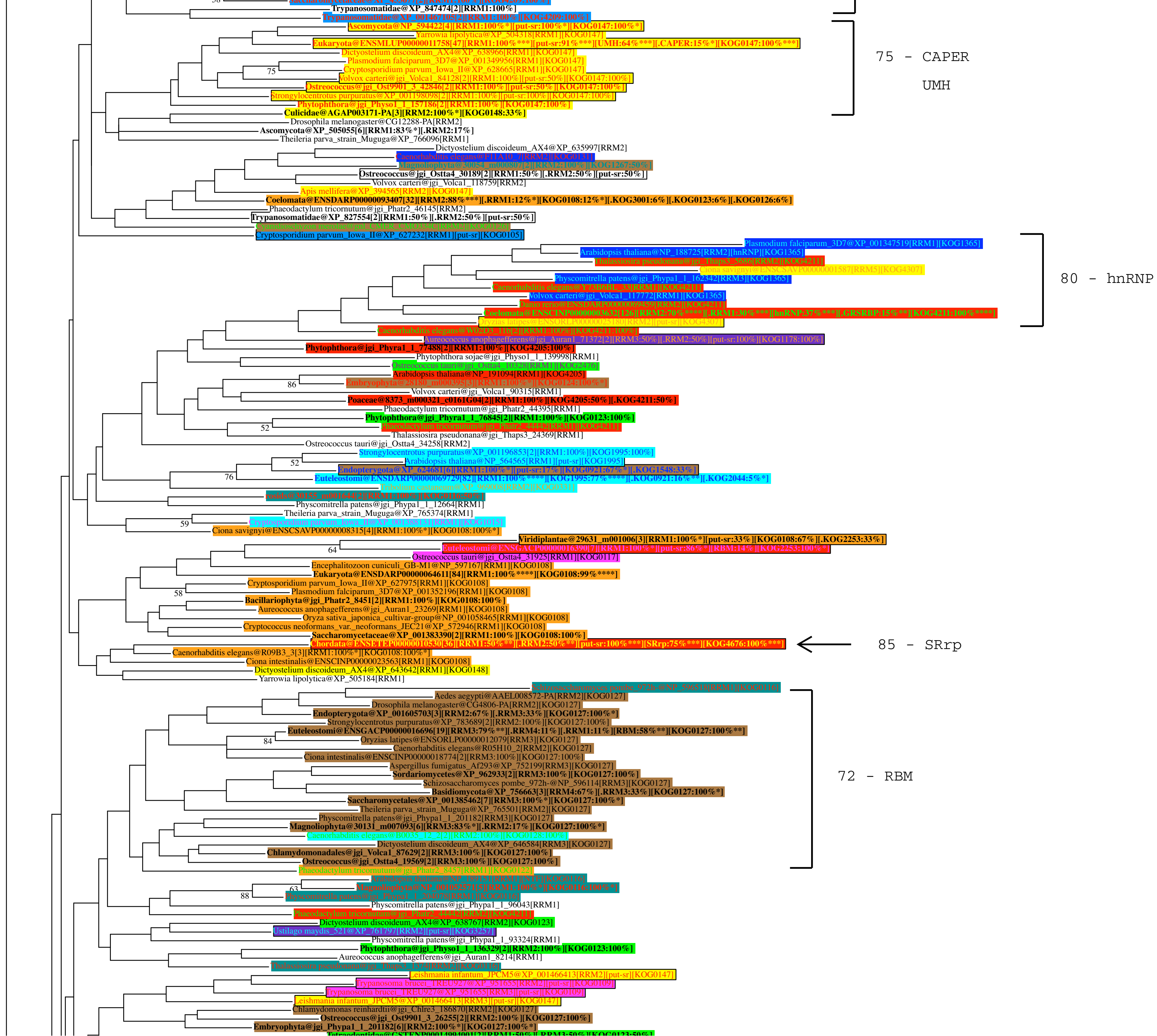


Fig. S6

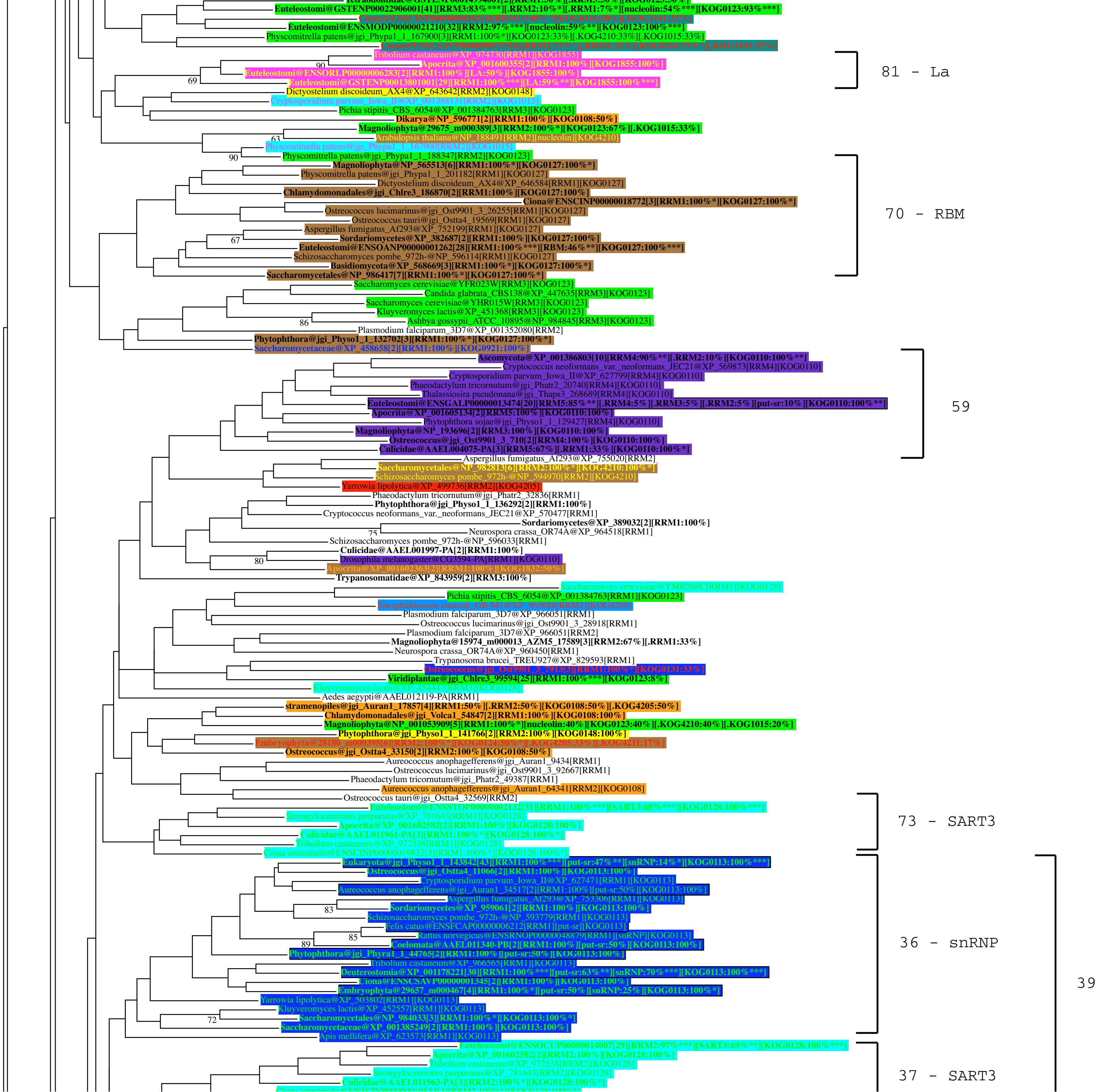


Fig. S6

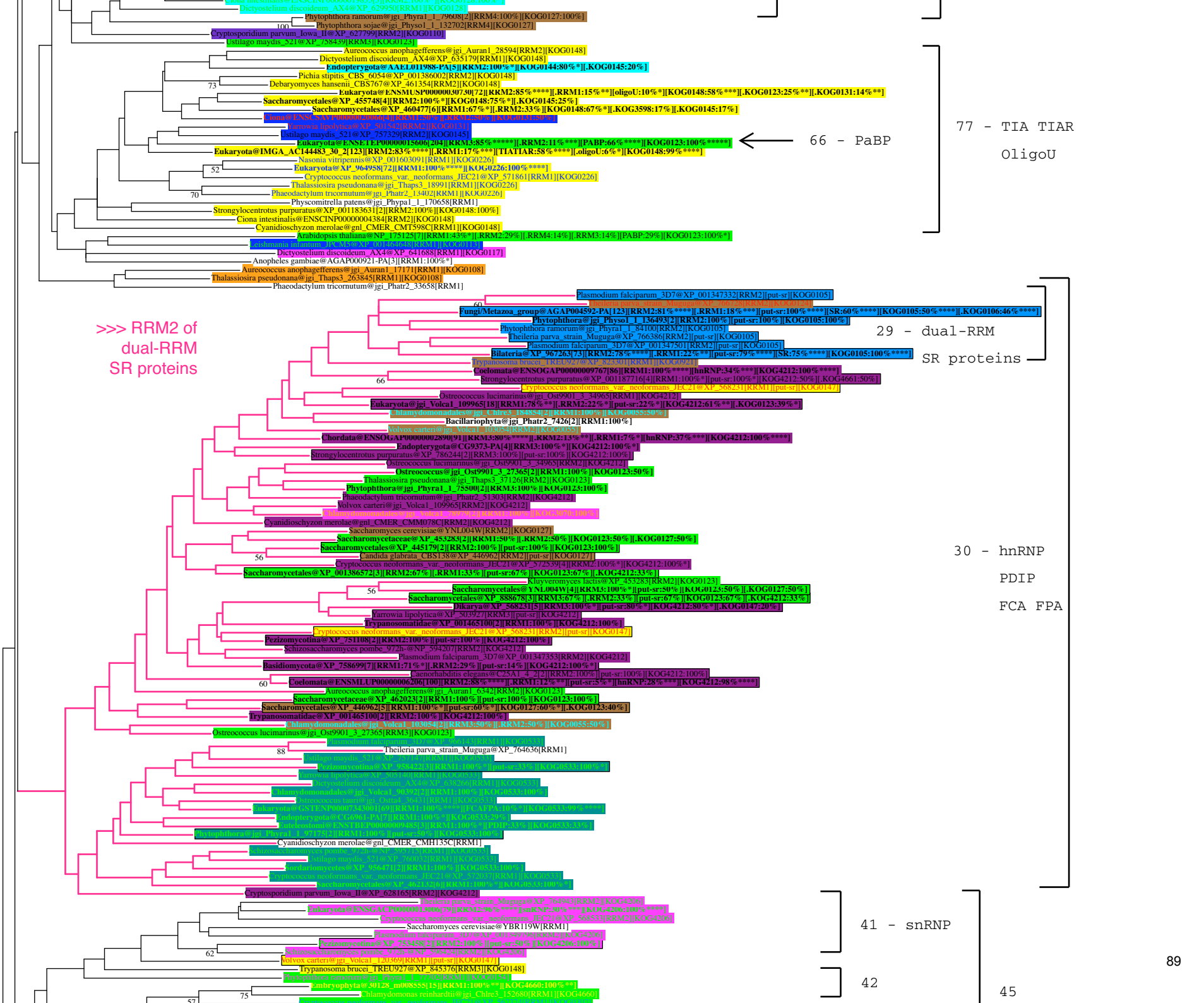

>>> RRM2 of dual-RRM SR proteins
66 - PaBP
77 - TIA TIAR OligoU
29 - dual-RRM SR proteins
30 - hnRNP PDIP FCA FPA
41 - snRNP
42
45

Fig. S6

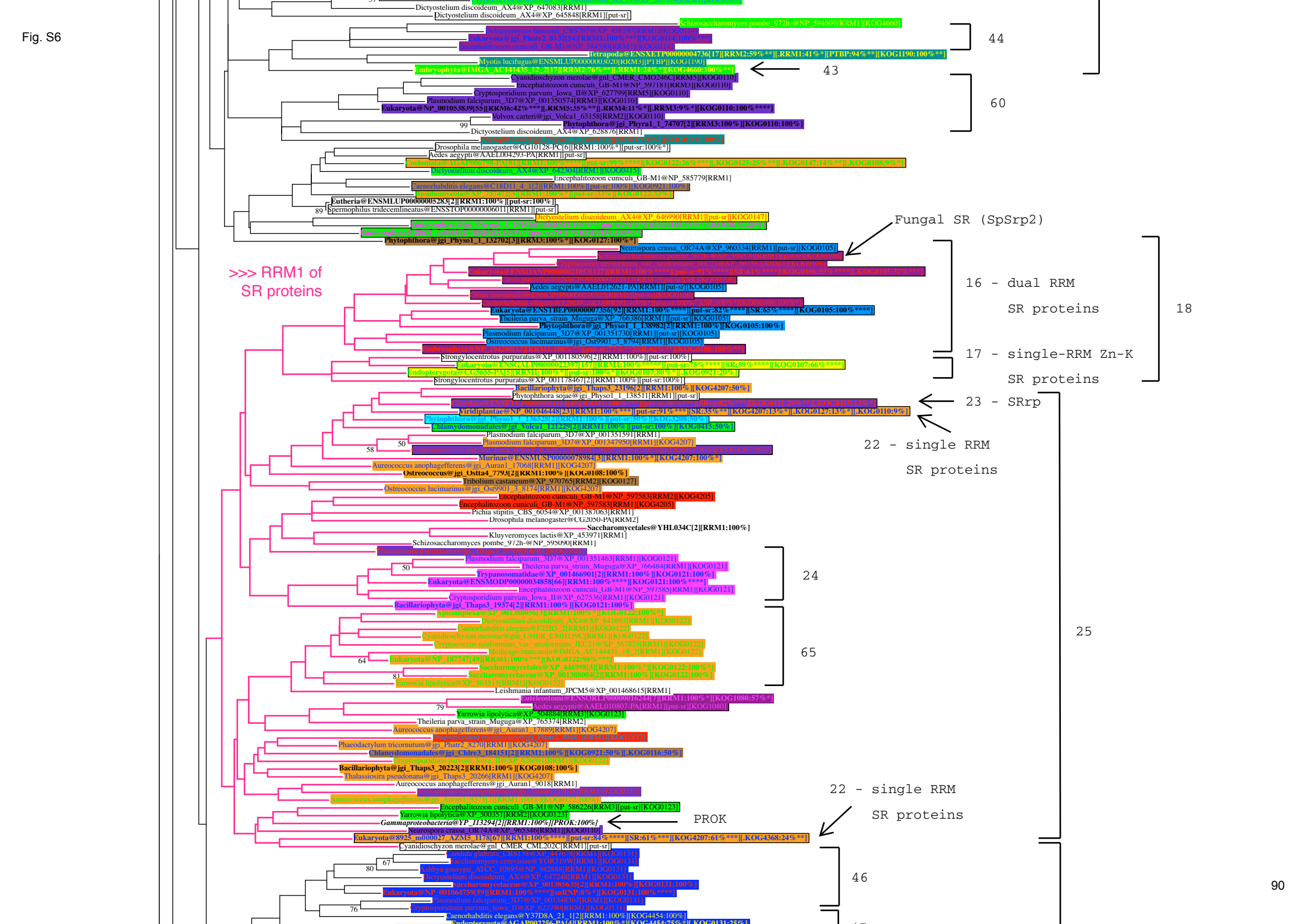

>>> RRM1 of SR proteins
Fungal SR (SpSrp2)
16 - dual RRM SR proteins
17 - single-RRM Zn-K SR proteins
18
23 - SRrp
22 - single RRM SR proteins
25
24
65
PROK
22 - single RRM SR proteins
46
44
43
60
90

Fig. S6

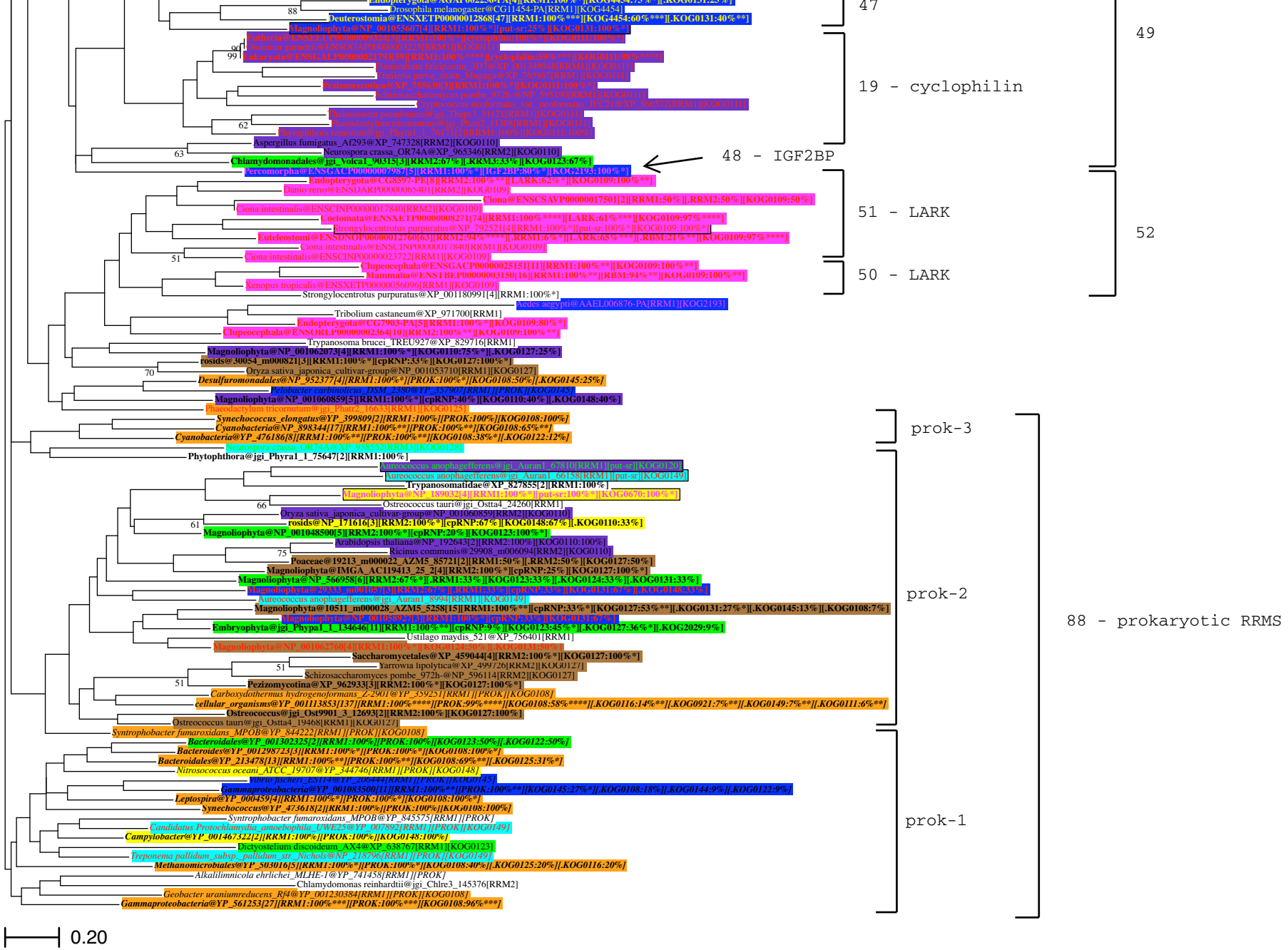

47
49
19 - cyclophilin
48 - IGF2BP
51 - LARK
52
50 - LARK
prok-3
prok-2
88 - prokaryotic RRMS
prok-1
0.20

Fig. S7

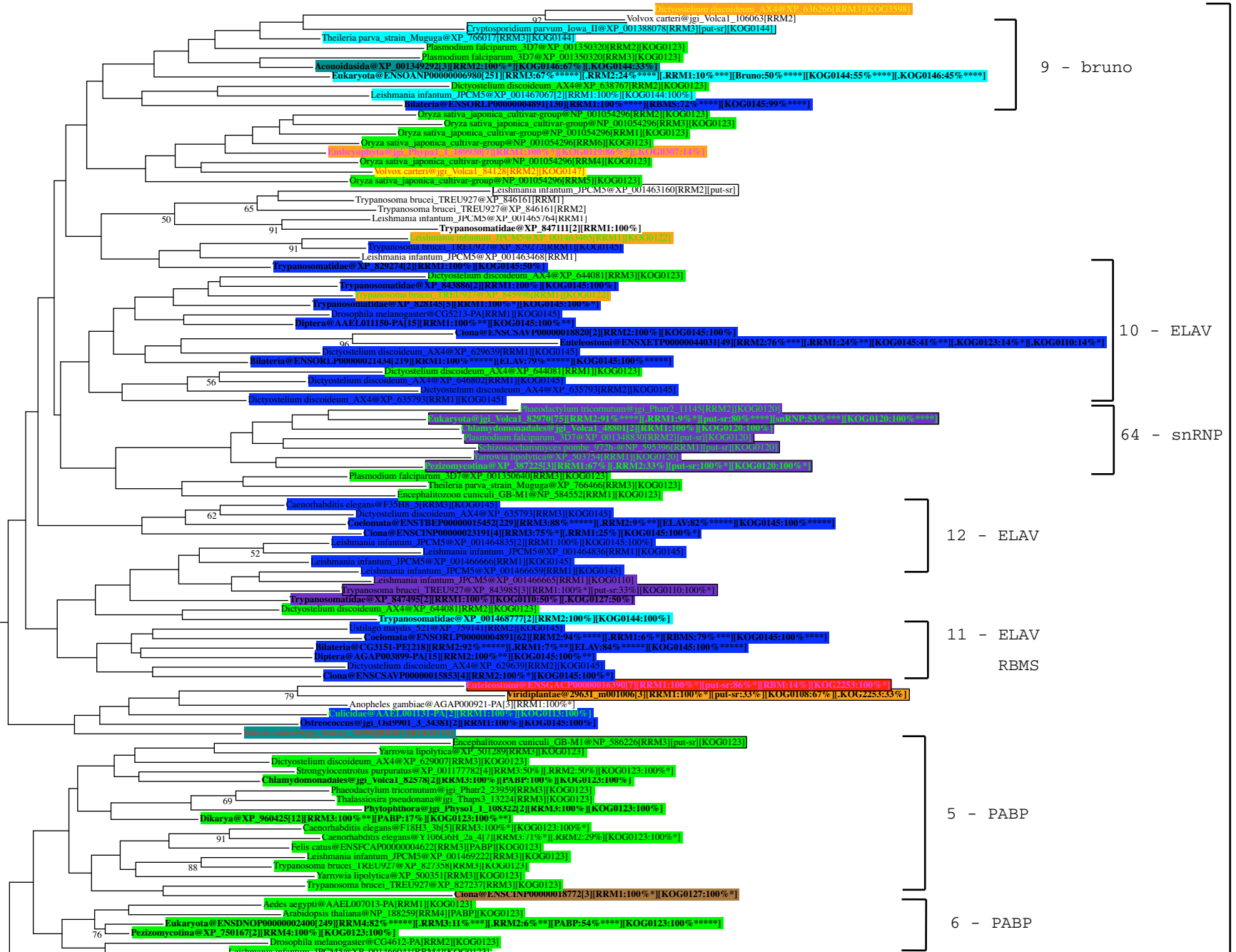


Fig. S7

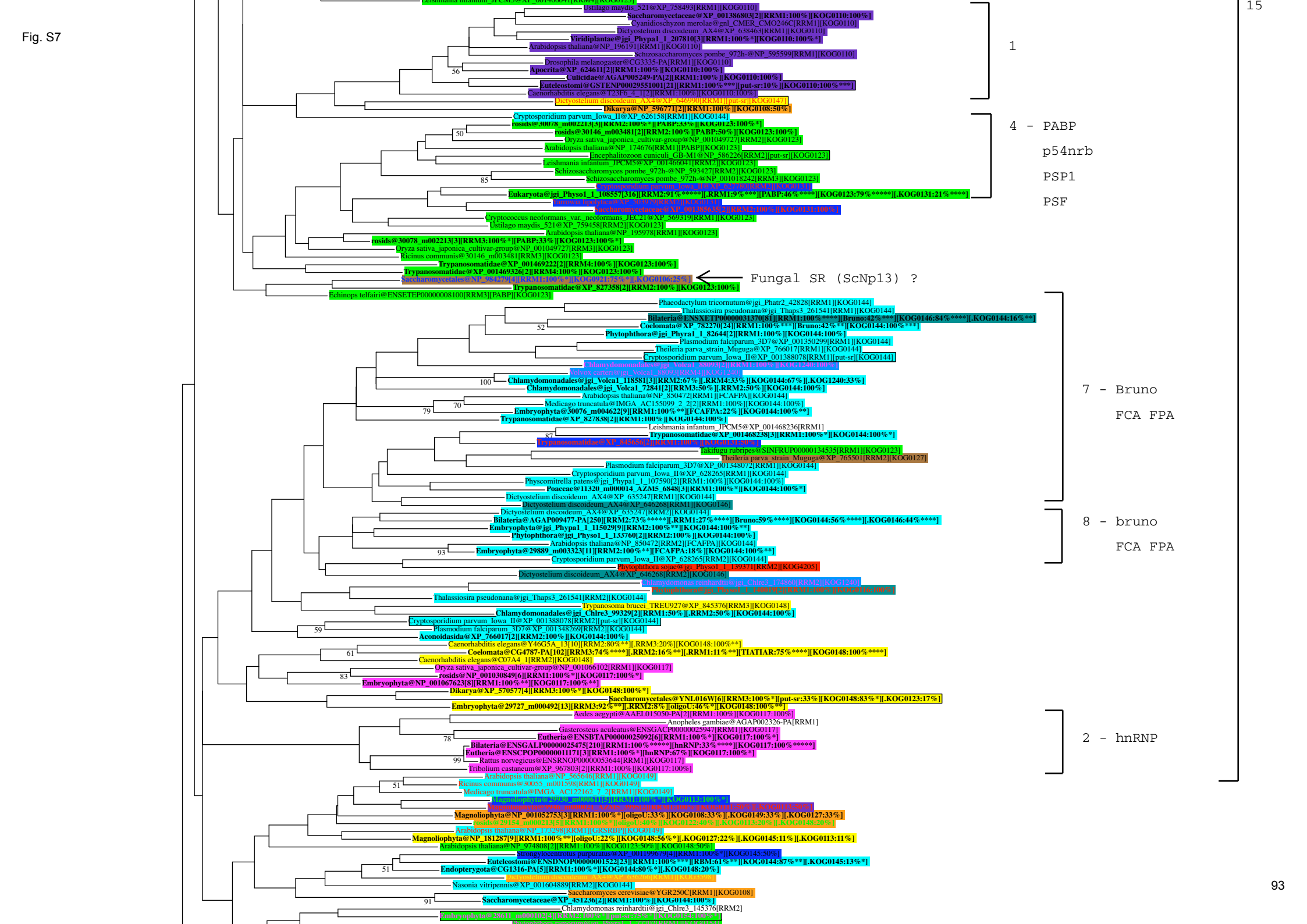

Ustilago maydis_521@XP_758493[RRM1][KOG0110]
Saccharomycetaceae@XP_001386803[2][RRM1:100%][KOG0110:100%]
Cyanidioschyzon merolae@gnl_CMER_CMO246C[RRM1][KOG0110]
Dictyostelium discoideum_AX4@XP_638463[RRM1][KOG0110]
Viridiplantae@jgi_Phypa1_1_207810[3][RRM1:100%*][KOG0110:100%*]
Arabidopsis thaliana@NP_196191[RRM1][KOG0110]
Schizosaccharomyces pombe_972h-@NP_595599[RRM1][KOG0110]
Drosophila melanogaster@CG3335-PA[RRM1][KOG0110]
Apocrita@XP_624611[2][RRM1:100%][KOG0110:100%]
Culicidae@AGAP005249-PA[2][RRM1:100%][KOG0110:100%]
Euteleostomi@GSTENP00029551001[21][RRM1:100%***][put-sr:10%][KOG0110:100%***]
Caenorhabditis elegans@T23F6_4_1[2][RRM1:100%][KOG0110:100%]
Dictyostelium discoideum_AX4@XP_646990[RRM1][put-sr][KOG0147]
Dikarya@NP_596771[2][RRM1:100%][KOG0108:50%]
Cryptosporidium parvum_Iowa_II@XP_626158[RRM1][KOG0144]
rosids@30078_m002213[5][RKM2:100%*][PABP:33%][KOG0123:100%*]
rosids@30146_m003481[2][RRM2:100%][PABP:50%][KOG0123:100%]
Oryza sativa_japonica_cultivar-group@NP_001049727[RRM2][KOG0123]
Arabidopsis thaliana@NP_174676[RRM1][PABP][KOG0123]
Encephalitozoon cuniculi_GB-M1@NP_586226[RRM2][put-sr][KOG0123]
Leishmania infantum_JPCM5@XP_001466041[RRM2][KOG0123]
Schizosaccharomyces pombe_972h-@NP_593427[RRM2][KOG0123]
Schizosaccharomyces pombe_972h-@NP_001018242[RRM3][KOG0123]
Eukaryota@jgi_Physo1_1_108557[316][RRM2:91%*****][.RRM1:9%***][PABP:46%****][KOG0123:79%*****][.KOG0131:21%****]
Saccharomycetaceae@XP_001385432[2][RRM2:100%][KOG0131:100%]
Cryptococcus neoformans_var._neoformans_JEC21@XP_569319[RRM1][KOG0123]
Ustilago maydis_521@XP_759458[RRM2][KOG0123]
Arabidopsis thaliana@NP_195978[RRM1][KOG0123]
rosids@30078_m002213[3][RRM3:100%*][PABP:33%][KOG0123:100%*]
Oryza sativa_japonica_cultivar-group@NP_001049727[RRM3][KOG0123]
Ricinus communis@30146_m003481[RRM3][KOG0123]
Trypanosomatidae@XP_001469222[2][RRM4:100%][KOG0123:100%]
Trypanosomatidae@XP_001469326[2][RRM4:100%][KOG0123:100%]
Saccharomycetales@NP_984279[4][RRM1:100%*][KOG0921:75%*][.KOG0106:25%]
Trypanosomatidae@XP_827358[2][RRM2:100%][KOG0123:100%]
Echinops telfairi@ENSETEP00000008100[RRM3][PABP][KOG0123]
Fungal SR (ScNp13) ?
Phaeodactylum tricornutum@jgi_Phatr2_42828[RRM1][KOG0144]
Thalassiosira pseudonana@jgi_Thaps3_261541[RRM1][KOG0144]
Bilateria@ENSXETP00000031370[81][RRM1:100%****][Bruno:42%***][KOG0146:84%****][.KOG0144:16%**]
Coelomata@XP_782270[24][RRM1:100%***][Bruno:42%**][KOG0144:100%***]
Phytophthora@jgi_Phyra1_1_82644[2][RRM1:100%][KOG0144:100%]
Plasmodium falciparum_3D7@XP_001350299[RRM1][KOG0144]
Theileria parva_strain_Muguga@XP_766017[RRM1][KOG0144]
Cryptosporidium parvum_Iowa_II@XP_001388078[RRM1][put-sr][KOG0144]
Chlamydomonadales@jgi_Volca1_88093[2][RRM1:100%][KOG1240:100%]
Volvox carteri@jgi_Volca1_88093[RRM3][KOG1240]
Chlamydomonadales@jgi_Volca1_118581[3][RRM2:67%][.RRM4:33%][KOG0144:67%][.KOG1240:33%]
Chlamydomonadales@jgi_Volca1_72841[2][RRM3:50%][.RRM2:50%][KOG0144:100%]
Arabidopsis thaliana@NP_850472[RRM1][FCAFPA][KOG0144]
Medicago truncatula@IMGA_AC155099_2_2[2][RRM1:100%][KOG0144:100%]
Embryophyta@30076_m004622[9][RRM1:100%**][FCAFPA:22%][KOG0144:100%**]
Trypanosomatidae@XP_827838[2][RRM1:100%][KOG0144:100%]
Leishmania infantum_JPCM5@XP_001468236[RRM1]
Trypanosomatidae@XP_001468238[3][RRM1:100%*][KOG0144:100%*]
Trypanosomatidae@XP_845656[2][RRM1:100%][KOG0144:100%]
Takifugu rubripes@SINFRUP00000134535[RRM1][KOG0123]
Theileria parva_strain_Muguga@XP_765501[RRM2][KOG0127]
Plasmodium falciparum_3D7@XP_001348072[RRM1][KOG0144]
Cryptosporidium parvum_Iowa_II@XP_628265[RRM1][KOG0144]
Physcomitrella patens@jgi_Phypa1_1_107590[2][RRM1:100%][KOG0144:100%]
Poaceae@11320_m000014_AZM5_6848[5][RRM1:100%*][KOG0144:100%*]
Dictyostelium discoideum_AX4@XP_635247[RRM1][KOG0144]
Dictyostelium discoideum_AX4@XP_646268[RRM1][KOG0146]
Dictyostelium discoideum_AX4@XP_635247[RRM2][KOG0144]
Bilateria@AGAP009477-PA[250][RRM2:73%*****][.RRM1:27%****][Bruno:59%****][KOG0144:56%****][.KOG0146:44%****]
Embryophyta@jgi_Phypa1_1_115029[9][RRM2:100%**][KOG0144:100%**]
Phytophthora@jgi_Physo1_1_133760[2][RRM2:100%][KOG0144:100%]
Arabidopsis thaliana@NP_850472[RRM2][FCAFPA][KOG0144]
Embryophyta@29889_m003323[11][RRM2:100%**][FCAFPA:18%][KOG0144:100%**]
Cryptosporidium parvum_Iowa_II@XP_628265[RRM2][KOG0144]
Phytophthora sojae@jgi_Physo1_1_139371[RRM2][KOG4205]
Dictyostelium discoideum_AX4@XP_646268[RRM2][KOG0146]
Chlamydomonas reinhardtii@jgi_Chlre3_174860[RRM2][KOG1240]
Thalassiosira pseudonana@jgi_Thaps3_261541[RRM2][KOG0144]
Trypanosoma brucei_TREU927@XP_845376[RRM3][KOG0148]
Chlamydomonadales@jgi_Chlre3_99329[2][RRM1:50%][.RRM2:50%][KOG0144:100%]
Cryptosporidium parvum_Iowa_II@XP_001388078[RRM2][put-sr][KOG0144]
Plasmodium falciparum_3D7@XP_001348269[RRM2][KOG0144]
Aconoidasida@XP_766017[2][RRM2:100%][KOG0144:100%]
Caenorhabditis elegans@Y46G5A_13[10][RRM2:80%**][.RRM3:20%][KOG0148:100%**]
Coelomata@CG4787-PA[102][RRM3:74%****][.RRM2:16%**][.RRM1:11%**][TIATIAR:75%****][KOG0148:100%****]
Caenorhabditis elegans@C07A4_1[RRM2][KOG0148]
Oryza sativa_japonica_cultivar-group@NP_001066102[RRM1][KOG0117]
rosids@NP_001030849[6][RRM1:100%*][KOG0117:100%*]
Embryophyta@NP_001067623[8][RRM1:100%**][KOG0117:100%**]
Dikarya@XP_570577[4][RRM3:100%*][KOG0148:100%*]
Saccharomycetales@YNL016W[6][RRM3:100%*][put-sr:33%][KOG0148:83%*][.KOG0123:17%]
Embryophyta@29727_m000492[13][RRM3:92%**][.RRM2:8%][oligoU:46%*][KOG0148:100%**]
Aedes aegypti@AAEL015050-PA[2][RRM1:100%][KOG0117:100%]
Anopheles gambiae@AGAP002326-PA[RRM1]
Gasterosteus aculeatus@ENSGACP00000025947[RRM1][KOG0117]
Eutheria@ENSBTAP00000025092[6][RRM1:100%*][KOG0117:100%*]
Bilateria@ENSGALP00000025475[210][RRM1:100%*****][hnRNP:33%****][KOG0117:100%*****]
Eutheria@ENSCPOP00000011171[3][RRM1:100%*][hnRNP:67%][KOG0117:100%*]
Rattus norvegicus@ENSRNOP00000053644[RRM1][KOG0117]
Tribolium castaneum@XP_967803[2][RRM1:100%][KOG0117:100%]
Arabidopsis thaliana@NP_565646[RRM1][KOG0149]
Ricinus communis@30055_m001598[RRM1][KOG0149]
Medicago truncatula@IMGA_AC122162_7_2[RRM1][KOG0149]
Magnoliophyta@29930_m000611[5][RRM1:60%*][KOG0113:100%*]
Magnoliophyta@NP_001052753[3][RRM1:100%*][oligoU:33%][KOG0108:33%][.KOG0149:33%][.KOG0127:33%]
rosids@29154_m000213[5][RRM1:100%*][oligoU:40%][KOG0122:40%][.KOG0113:20%][.KOG0148:20%]
Arabidopsis thaliana@NP_175299[RRM1][CHNRB][KOG0149]
Magnoliophyta@NP_1812879[RRM1:100%**][oligoU:22%][KOG0148:56%*][.KOG0127:22%][.KOG0145:11%][.KOG0113:11%]
Arabidopsis thaliana@NP_9748082[2][RRM1:100%][KOG0123:50%][.KOG0148:50%]
Strongylocentrotus purpuratus@XP_001179791[RRM1:100%*][KOG0145:50%]
Euteleostomi@ENSDNOP00000001522[23][RRM1:100%***][RBM:61%**][KOG0144:87%**][.KOG0145:13%*]
Endopterygota@CG1316-PA[5][RRM1:100%*][KOG0144:80%*][.KOG0148:20%]
Dictyostelium discoideum_AX4@XP_636760[RRM1][KOG0149]
Nasonia vitripennis@XP_001604889[RRM2][KOG0144]
Saccharomyces cerevisiae@YGR250C[RRM1][KOG0108]
Saccharomycetaceae@XP_451236[2][RRM1:100%][KOG0144:100%]
Chlamydomonas reinhardtii@jgi_Chlre3_145376[RRM2]
Embryophyta@28611_m001021[4][RRM2:100%][put-sr:58%][KOG0154:100%]
1
4 - PABP
p54nrb
PSP1
PSF
7 - Bruno
FCA FPA
8 - bruno
FCA FPA
2 - hnRNP
15

Fig. S7

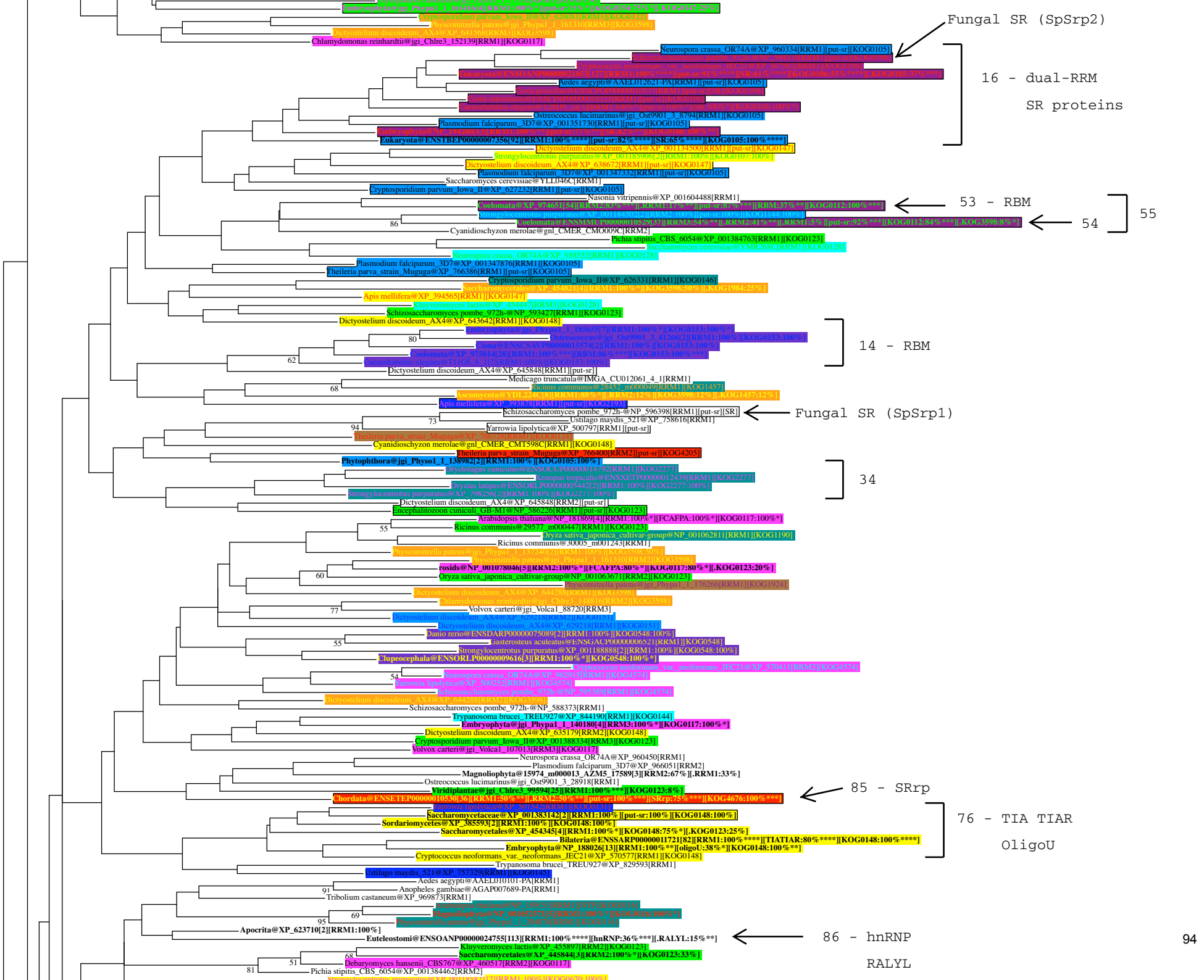

Fungal SR (SpSrp2)
16 - dual-RRM SR proteins
53 - RBM
54
55
14 - RBM
Fungal SR (SpSrp1)
34
85 - SRrp
76 - TIA TIAR OligoU
86 - hnRNP RALYL

Fig. S7

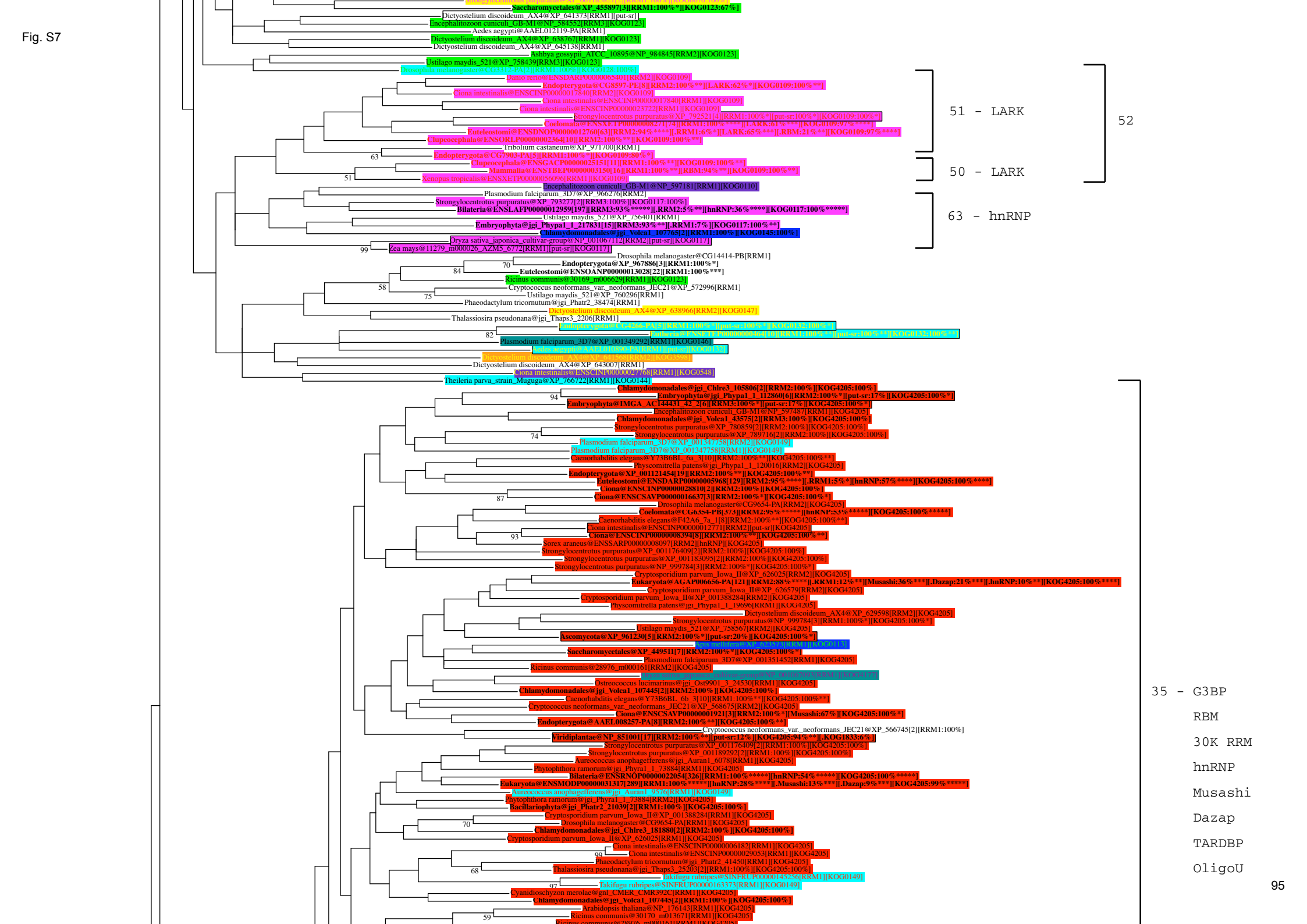

Saccharomycetales@XP_455897[3][RRM1:100%*][KOG0123:67%*]
Dictyostelium discoideum_AX4@XP_643373[RRM1][put-sr]
Encephalitozoon cuniculi_GB-M1@NP_584552[RRM3][KOG0123]
Aedes aegypti@AAEL012119-PA[RRM1]
Dictyostelium discoideum_AX4@XP_638767[RRM1][KOG0123]
Dictyostelium discoideum_AX4@XP_645138[RRM1]
Ashbya gossypii_ATCC_10895@NP_984845[RRM2][KOG0123]
Ustilago maydis_521@XP_758439[RRM3][KOG0123]
Drosophila melanogaster@CG33122-PA[2][RRM1:100%][KOG0128:100%]
Danio rerio@ENSDARP00000065401[RRM2][KOG0109]
Endopterygota@CG8597-PE[8][RRM2:100%**][LARK:62%*][KOG0109:100%**]
Ciona intestinalis@ENSCINP00000017840[RRM2][KOG0109]
Ciona intestinalis@ENSCINP00000017840[RRM1][KOG0109]
Ciona intestinalis@ENSCINP00000023722[RRM1][KOG0109]
Strongylocentrotus purpuratus@XP_792521[4][RRM1:100%*][put-sr:100%*][KOG0109:100%*]
Coelomata@ENSXETP00000005271[74][RRM1:100%****][LARK:61%***][KOG0109:97%****]
Euteleostomi@ENSDNOP00000012760[63][RRM2:94%****][.RRM1:6%*][LARK:65%***][.RBM:21%*][KOG0109:97%****]
Clupeocephala@ENSORLP00000002364[10][RRM2:100%**][KOG0109:100%**]
Tribolium castaneum@XP_971700[RRM1]
Endopterygota@CG7903-PA[5][RRM1:100%*][KOG0109:80%*]
Clupeocephala@ENSGACP00000025151[11][RRM1:100%**][KOG0109:100%**]
Mammalia@ENSBTAP00000031510[16][RRM1:100%**][RBM:94%**][KOG0109:100%**]
Xenopus tropicalis@ENSXETP00000056096[RRM1][KOG0109]
Encephalitozoon cuniculi_GB-M1@NP_597181[RRM1][KOG0110]
Plasmodium falciparum_3D7@XP_966276[RRM2]
Strongylocentrotus purpuratus@XP_793277[2][RRM3:100%][KOG0117:100%]
Bilateria@ENSLAFP00000012959[197][RRM3:93%*****][.RRM2:5%**][hnRNP:36%****][KOG0117:100%*****]
Ustilago maydis_521@XP_756401[RRM1]
Embryophyta@jgi_Phypa1_1_217831[15][RRM3:93%**][.RRM1:7%][KOG0117:100%**]
Chlamydomonadales@jgi_Volca1_107765[2][RRM1:100%][KOG0145:100%]
Oryza sativa_japonica_cultivar-group@NP_001067112[RRM2][put-sr][KOG0117]
Zea mays@11279_m000026_AZM5_6772[RRM1][put-sr][KOG0117]
Drosophila melanogaster@CG14414-PB[RRM1]
Endopterygota@XP_967886[3][RRM1:100%*]
Euteleostomi@ENSOANP00000013028[22][RRM1:100%***]
Ricinus communis@30169_m006629[RRM1][KOG0123]
Cryptococcus neoformans_var._neoformans_JEC21@XP_572996[RRM1]
Ustilago maydis_521@XP_760296[RRM1]
Phaeodactylum tricornutum@jgi_Phatr2_38474[RRM1]
Dictyostelium discoideum_AX4@XP_638966[RRM2][KOG0147]
Thalassiosira pseudonana@jgi_Thaps3_2206[RRM1]
Endopterygota@CG4266-PA[2][RRM1:100%*][put-sr:100%*][KOG0132:100%**]
Coelomata@ENSXETP00000034410[8][RRM1:100%**][put-sr:100%**][KOG0132:100%**]
Plasmodium falciparum_3D7@XP_001349292[RRM1][KOG0146]
Aedes aegypti@AAEL010990-PA[RRM1][put-sr][KOG0132]
Dictyostelium discoideum_AX4@XP_641960[RRM1][KOG0595]
Dictyostelium discoideum_AX4@XP_643007[RRM1]
Ciona intestinalis@ENSCINP00000027768[RRM1][KOG0548]
Theileria parva_strain_Muguga@XP_766722[RRM1][KOG0144]
Chlamydomonadales@jgi_Chlre3_105806[2][RRM2:100%][KOG4205:100%]
Embryophyta@jgi_Phypa1_1_112860[6][RRM2:100%*][put-sr:17%][KOG4205:100%*]
Embryophyta@IMGA_AC144431_42_2[6][RRM3:100%*][put-sr:17%][KOG4205:100%**]
Encephalitozoon cuniculi_GB-M1@NP_597487[RRM1][KOG4205]
Chlamydomonadales@jgi_Volca1_43575[2][RRM3:100%][KOG4205:100%]
Strongylocentrotus purpuratus@XP_780859[2][RRM2:100%][KOG4205:100%]
Strongylocentrotus purpuratus@XP_789716[2][RRM2:100%][KOG4205:100%]
Plasmodium falciparum_3D7@XP_001347758[RRM2][KOG0149]
Plasmodium falciparum_3D7@XP_001347758[RRM1][KOG0149]
Caenorhabditis elegans@Y73B6BL_6a_3[10][RRM2:100%**][KOG4205:100%**]
Physcomitrella patens@jgi_Phypa1_1_120016[RRM2][KOG4205]
Endopterygota@XP_001121454[19][RRM2:100%**][KOG4205:100%**]
Euteleostomi@ENSDARP00000005968[129][RRM2:95%****][.RRM1:5%*][hnRNP:57%****][KOG4205:100%****]
Ciona@ENSCINP00000028810[2][RRM2:100%][KOG4205:100%]
Ciona@ENSCSAVP00000016637[3][RRM2:100%*][KOG4205:100%*]
Drosophila melanogaster@CG9654-PA[RRM2][KOG4205]
Coelomata@CG6554-PB[175][RRM2:95%*****][hnRNP:53%*****][KOG4205:100%*****]
Caenorhabditis elegans@F42A6_7a_1[8][RRM2:100%**][KOG4205:100%**]
Ciona intestinalis@ENSCINP00000012771[RRM2][put-sr][KOG4205]
Ciona@ENSCINP00000008394[8][RRM2:100%**][KOG4205:100%**]
Sorex araneus@ENSSARP00000008097[RRM2][hnRNP][KOG4205]
Strongylocentrotus purpuratus@XP_001176409[2][RRM2:100%][KOG4205:100%]
Strongylocentrotus purpuratus@XP_001183095[2][RRM2:100%][KOG4205:100%]
Strongylocentrotus purpuratus@NP_999784[3][RRM2:100%*][KOG4205:100%*]
Cryptosporidium parvum_Iowa_II@XP_626025[RRM2][KOG4205]
Eukaryota@AGAP006656-PA[121][RRM2:88%****][.RRM1:12%**][Musashi:36%***][.Dazap:21%***][.hnRNP:10%**][KOG4205:100%****]
Cryptosporidium parvum_Iowa_II@XP_626579[RRM2][KOG4205]
Cryptosporidium parvum_Iowa_II@XP_001388284[RRM2][KOG4205]
Physcomitrella patens@jgi_Phypa1_1_19696[RRM1][KOG4205]
Dictyostelium discoideum_AX4@XP_629598[RRM2][KOG4205]
Strongylocentrotus purpuratus@NP_999784[3][RRM1:100%*][KOG4205:100%*]
Ustilago maydis_521@XP_758567[RRM2][KOG4205]
Ascomycota@XP_961230[5][RRM2:100%*][put-sr:20%][KOG4205:100%*]
Saccharomycetales@XP_449511[7][RRM2:100%*][KOG4205:100%*]
Plasmodium falciparum_3D7@XP_001351452[RRM1][KOG4205]
Ricinus communis@28976_m000161[RRM2][KOG4205]
Ostreococcus lucimarinus@jgi_Ost9901_3_24530[RRM1][KOG4205]
Chlamydomonadales@jgi_Volca1_107445[2][RRM2:100%][KOG4205:100%]
Caenorhabditis elegans@Y73B6BL_6b_3[10][RRM1:100%**][KOG4205:100%**]
Cryptococcus neoformans_var._neoformans_JEC21@XP_568675[RRM2][KOG4205]
Ciona@ENSCSAVP00000001921[3][RRM2:100%*][Musashi:67%][KOG4205:100%*]
Endopterygota@AAEL008257-PA[8][RRM2:100%**][KOG4205:100%**]
Cryptococcus neoformans_var._neoformans_JEC21@XP_566745[2][RRM1:100%]
Viridiplantae@NP_851001[17][RRM2:100%**][put-sr:12%][KOG4205:94%**][.KOG1833:6%]
Strongylocentrotus purpuratus@XP_001176409[2][RRM1:100%][KOG4205:100%]
Strongylocentrotus purpuratus@XP_001189292[2][RRM1:100%][KOG4205:100%]
Aureococcus anophagefferens@jgi_Auran1_6078[RRM1][KOG4205]
Phytophthora ramorum@jgi_Phyra1_1_73884[RRM1][KOG4205]
Bilateria@ENSRNOP00000022054[326][RRM1:100%*****][hnRNP:54%*****][KOG4205:100%*****]
Eukaryota@ENSMODP00000031317[289][RRM1:100%*****][hnRNP:28%****][.Musashi:13%***][.Dazap:9%***][KOG4205:99%*****]
Aureococcus anophagefferens@jgi_Auran1_9576[RRM1][KOG0149]
Phytophthora ramorum@jgi_Phyra1_1_73884[RRM2][KOG4205]
Bacillariophyta@jgi_Phatr2_21039[2][RRM1:100%][KOG4205:100%]
Cryptosporidium parvum_Iowa_II@XP_001388284[RRM1][KOG4205]
Drosophila melanogaster@CG9654-PA[RRM1][KOG4205]
Chlamydomonadales@jgi_Chlre3_181880[2][RRM2:100%][KOG4205:100%]
Cryptosporidium parvum_Iowa_II@XP_626025[RRM1][KOG4205]
Ciona intestinalis@ENSCINP00000006182[RRM1][KOG4205]
Ciona intestinalis@ENSCINP00000029053[RRM1][KOG4205]
Phaeodactylum tricornutum@jgi_Phatr2_41450[RRM1][KOG4205]
Thalassiosira pseudonana@jgi_Thaps3_25203[2][RRM1:100%][KOG4205:100%]
Takifugu rubripes@SINFRUP00000145256[RRM1][KOG0149]
Takifugu rubripes@SINFRUP00000163373[RRM1][KOG0149]
Cyanidioschyzon merolae@gnl_CMER_CMR392C[RRM1][KOG4205]
Chlamydomonadales@jgi_Volca1_107445[2][RRM1:100%][KOG4205:100%]
Arabidopsis thaliana@NP_176143[RRM1][KOG4205]
Ricinus communis@30170_m013671[RRM1][KOG4205]
51 - LARK
50 - LARK
52
63 - hnRNP
35 - G3BP
RBM
30K RRM
hnRNP
Musashi
Dazap
TARDBP
OligoU
95

Fig. S7

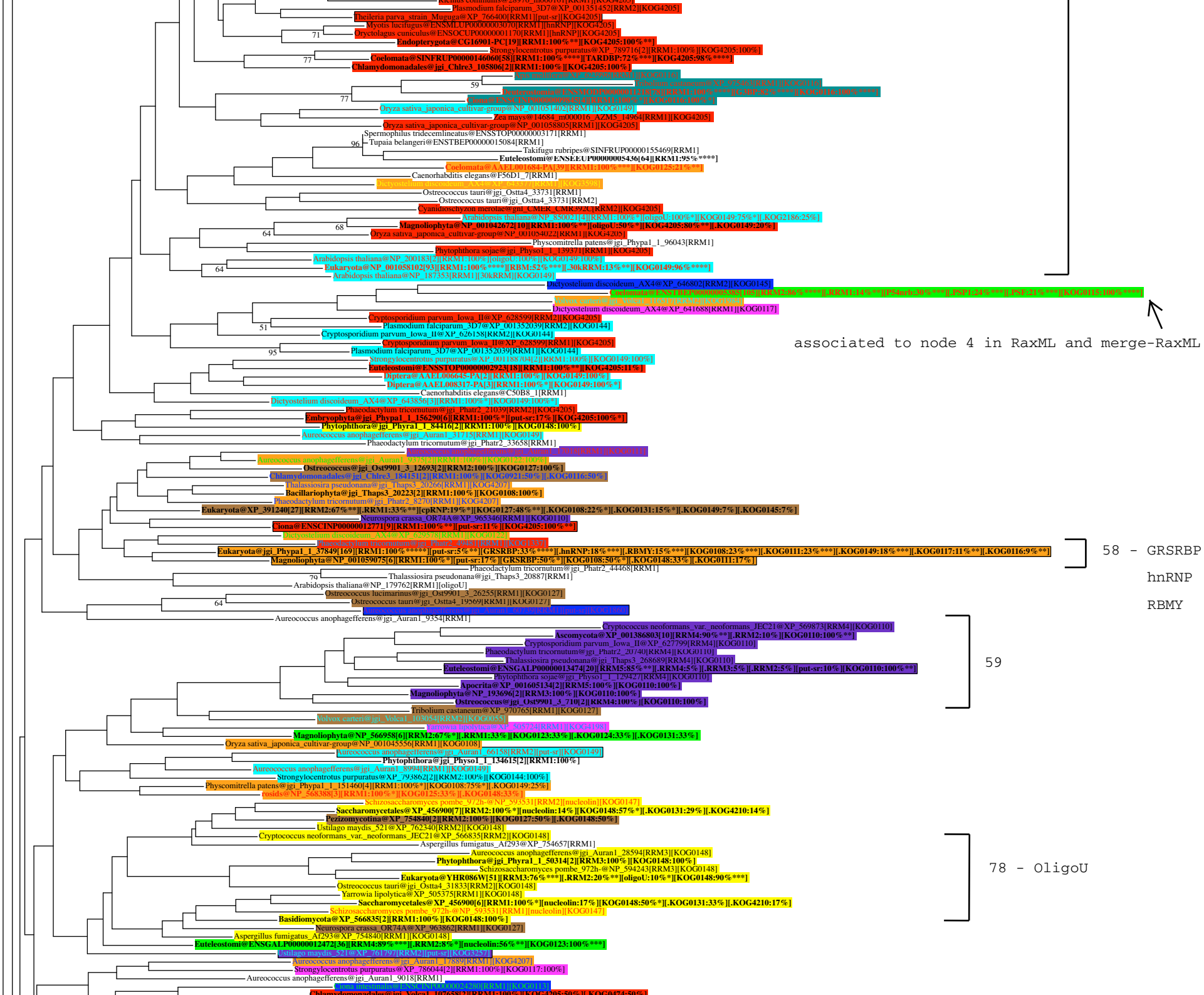

Plasmodium falciparum_3D7@XP_001351452[RRM2][KOG4205]
Theileria parva_strain_Muguga@XP_766400[RRM1][put-sr][KOG4205]
Myotis lucifugus@ENSMLUP00000003070[RRM1][hnRNP][KOG4205]
Oryctolagus cuniculus@ENSOCUP00000001170[RRM1/hnRNP][KOG4205]
Endopterygota@CG16901-PC[19][RRM1:100%**][KOG4205:100%**]
Strongylocentrotus purpuratus@XP_789716[2][RRM1:100%][KOG4205:100%]
Coelomata@SINFRUP00000146060[58][RRM1:100%****][TARDBP:72%***][KOG4205:98%****]
Chlamydomonadales@jgi_Chlre3_105806[2][RRM1:100%][KOG4205:100%]
Oryza sativa_japonica_cultivar-group@NP_001051402[RRM1][KOG0149]
Zea mays@14684_m000016_AZM5_14964[RRM1][KOG4205]
Oryza sativa_japonica_cultivar-group@NP_001058805[RRM1][KOG4205]
Spermophilus tridecemlineatus@ENSSTOP00000003171[RRM1]
Tupaia belangeri@ENSTBEP00000015084[RRM1]
Takifugu rubripes@SINFRUP00000155469[RRM1]
Euteleostomi@ENSEEUP00000005436[64][RRM1:95%****]
Caenorhabditis elegans@F56D1_7[RRM1]
Ostreococcus tauri@jgi_Ostta4_33731[RRM1]
Ostreococcus tauri@jgi_Ostta4_33731[RRM2]
Cyanidioschyzon merolae@gnl_CMER_CMR392C[RRM2][KOG4205]
Magnoliophyta@NP_001042672[10][RRM1:100%**][oligoU:50%*][KOG4205:80%**][.KOG0149:20%]
Oryza sativa_japonica_cultivar-group@NP_001054022[RRM1][KOG4205]
Physcomitrella patens@jgi_Phypa1_1_96043[RRM1]
Phytophthora sojae@jgi_Physo1_1_139371[RRM1][KOG4205]
Eukaryota@NP_001058102[93][RRM1:100%****][RBM:52%***][.30kRRM:13%**][KOG0149:96%****]
Arabidopsis thaliana@NP_187353[RRM1][30kRRM][KOG0149]
Dictyostelium discoideum_AX4@XP_646802[RRM2][KOG0145]
Dictyostelium discoideum_AX4@XP_641688[RRM1][KOG0117]
Cryptosporidium parvum_Iowa_II@XP_628599[RRM2][KOG4205]
Plasmodium falciparum_3D7@XP_001352039[RRM2][KOG0144]
Cryptosporidium parvum_Iowa_II@XP_626158[RRM2][KOG0144]
Cryptosporidium parvum_Iowa_II@XP_628599[RRM1][KOG4205]
Plasmodium falciparum_3D7@XP_001352039[RRM1][KOG0144]
Euteleostomi@ENSSTOP00000002923[15][RRM1:100%**][KOG4205:11%]
Diptera@AAEL006645-PA2[RRM1:100%*][KOG0149:100%]
Diptera@AAEL008317-PA3[RRM1:100%*][KOG0149:100%*]
Caenorhabditis elegans@C50B8_1[RRM1]
Dictyostelium discoideum_AX4@XP_643856[3][RRM1:100%*][KOG0149:100%*]
Phaeodactylum tricornutum@jgi_Phatr2_21039[RRM2][KOG4205]
Embryophyta@jgi_Phypa1_1_156290[6][RRM1:100%*][put-sr:17%][KOG4205:100%*]
Phytophthora@jgi_Phyra1_1_84416[2][RRM1:100%][KOG0148:100%]
Aureococcus anophagefferens@jgi_Auran1_31715[RRM1][KOG0149]
Phaeodactylum tricornutum@jgi_Phatr2_33658[RRM1]
Aureococcus anophagefferens@jgi_Auran1_9375[2][RRM1:100%][KOG0122:100%]
Ostreococcus@jgi_Ost9901_3_12693[2][RRM2:100%][KOG0127:100%]
Chlamydomonadales@jgi_Chlre3_184151[2][RRM1:100%][.KOG0921:50%][.KOG0116:50%]
Thalassiosira pseudonana@jgi_Thaps3_20266[RRM1][KOG4207]
Bacillariophyta@jgi_Thaps3_20223[2][RRM1:100%][KOG0108:100%]
Phaeodactylum tricornutum@jgi_Phatr2_8270[RRM1][KOG4207]
Eukaryota@XP_391240[27][RRM2:67%**][.RRM1:33%**][cpRNP:19%*][KOG0127:48%**][.KOG0108:22%*][.KOG0131:15%*][.KOG0149:7%][.KOG0145:7%]
Neurospora crassa_OR74A@XP_965346[RRM1][KOG0110]
Ciona@ENSCINP00000012771[9][RRM1:100%***][put-sr:11%][KOG4205:100%**]
Dictyostelium discoideum_AX4@XP_629978[RRM1][KOG0122]
Eukaryota@jgi_Phypa1_1_37849[169][RRM1:100%****][put-sr:5%**][GRSRBP:33%****][.hnRNP:18%***][.RBMY:15%***][KOG0108:23%***][.KOG0111:23%***][.KOG0149:18%***][.KOG0117:11%**][.KOG0116:9%**]
Magnoliophyta@NP_001059075[6][RRM1:100%*][put-sr:17%][GRSRBP:50%*][KOG0108:50%*][.KOG0148:33%][.KOG0111:17%]
Phaeodactylum tricornutum@jgi_Phatr2_44468[RRM1]
Thalassiosira pseudonana@jgi_Thaps3_20887[RRM1]
Arabidopsis thaliana@NP_179762[RRM1][oligoU]
Ostreococcus lucimarinus@jgi_Ost9901_3_26255[RRM1][KOG0127]
Ostreococcus tauri@jgi_Ostta4_19569[RRM1][KOG0127]
Aureococcus anophagefferens@jgi_Auran1_9354[RRM1]
Cryptococcus neoformans_var._neoformans_JEC21@XP_569873[RRM4][KOG0110]
Ascomycota@XP_001386803[10][RRM4:90%**][.RRM2:10%][KOG0110:100%**]
Cryptosporidium parvum_Iowa_II@XP_627799[RRM4][KOG0110]
Phaeodactylum tricornutum@jgi_Phatr2_20740[RRM4][KOG0110]
Thalassiosira pseudonana@jgi_Thaps3_268689[RRM4][KOG0110]
Euteleostomi@ENSGALP00000013474[20][RRM5:85%**][.RRM4:5%][.RRM3:5%][.RRM2:5%][put-sr:10%][KOG0110:100%**]
Phytophthora sojae@jgi_Physo1_1_129427[RRM4][KOG0110]
Apocrita@XP_001605134[2][RRM5:100%][KOG0110:100%]
Magnoliophyta@NP_193696[2][RRM3:100%][KOG0110:100%]
Ostreococcus@jgi_Ost9901_3_710[2][RRM4:100%][KOG0110:100%]
Tribolium castaneum@XP_970765[RRM1][KOG0127]
Volvox carteri@jgi_Volca1_106857[RRM2][KOG0145]
Yarrowia lipolytica@XP_505724[RRM1][KOG4198]
Magnoliophyta@NP_566958[6][RRM2:67%**][.RRM1:33%][KOG0123:33%][.KOG0124:33%][.KOG0131:33%]
Oryza sativa_japonica_cultivar-group@NP_001045556[RRM1][KOG0108]
Aureococcus anophagefferens@jgi_Auran1_66158[RRM2][put-sr][KOG0149]
Phytophthora@jgi_Physo1_1_134615[2][RRM1:100%]
Aureococcus anophagefferens@jgi_Auran1_8994[RRM1][KOG0149]
Strongylocentrotus purpuratus@XP_793862[2][RRM2:100%][KOG0144:100%]
Physcomitrella patens@jgi_Phypa1_1_151460[4][RRM1:100%*][KOG0108:75%*][.KOG0149:25%]
rosids@NP_568388[3][RRM1:100%*][KOG0125:33%][.KOG0148:33%]
Schizosaccharomyces pombe_972h-@NP_593531[RRM2][nucleolin][KOG0147]
Saccharomycetales@XP_456900[7][RRM2:100%*][nucleolin:14%][KOG0148:57%*][.KOG0131:29%][.KOG4210:14%]
Pezizomycotina@XP_754840[2][RRM2:100%][KOG0127:50%][.KOG0148:50%]
Ustilago maydis_521@XP_762340[RRM2][KOG0148]
Cryptococcus neoformans_var._neoformans_JEC21@XP_568835[RRM2][KOG0148]
Aspergillus fumigatus_Af293@XP_754657[RRM1]
Aureococcus anophagefferens@jgi_Auran1_28594[RRM3][KOG0148]
Phytophthora@jgi_Phyra1_1_50314[2][RRM3:100%][KOG0148:100%]
Schizosaccharomyces pombe_972h-@NP_594243[RRM3][KOG0148]
Eukaryota@YHR086W[51][RRM3:76%***][.RRM2:20%**][oligoU:10%*][KOG0148:90%***]
Ostreococcus tauri@jgi_Ostta4_31833[RRM2][KOG0148]
Yarrowia lipolytica@XP_505375[RRM1][KOG0148]
Saccharomycetales@XP_456900[6][RRM1:100%*][nucleolin:17%][KOG0148:50%*][.KOG0131:33%][.KOG4210:17%]
Schizosaccharomyces pombe_972h-@NP_593531[RRM1][nucleolin][KOG0147]
Basidiomycota@XP_566835[2][RRM1:100%][KOG0148:100%]
Neurospora crassa_OR74A@XP_963862[RRM1][KOG0127]
Aspergillus fumigatus_Af293@XP_754840[RRM1][KOG0148]
Euteleostomi@ENSGALP00000012472[36][RRM4:89%***][.RRM2:8%*][nucleolin:56%**][KOG0123:100%***]
Ustilago maydis_521@XP_761707[RRM2][put-sr][KOG4257]
Aureococcus anophagefferens@jgi_Auran1_17889[RRM1][KOG4207]
Strongylocentrotus purpuratus@XP_786044[2][RRM1:100%][KOG0117:100%]
Aureococcus anophagefferens@jgi_Auran1_9018[RRM1]
Danio rerio@ENSDARP00000024280[RRM1][KOG0113]
Chlamydomonadales@jgi_Volca1_107658[2][RRM1:100%][KOG4205:50%][.KOG0474:50%]
associated to node 4 in RaxML and merge-RaxML
58 - GRSRBP
hnRNP
RBMY
59
78 - OligoU

Fig. S7

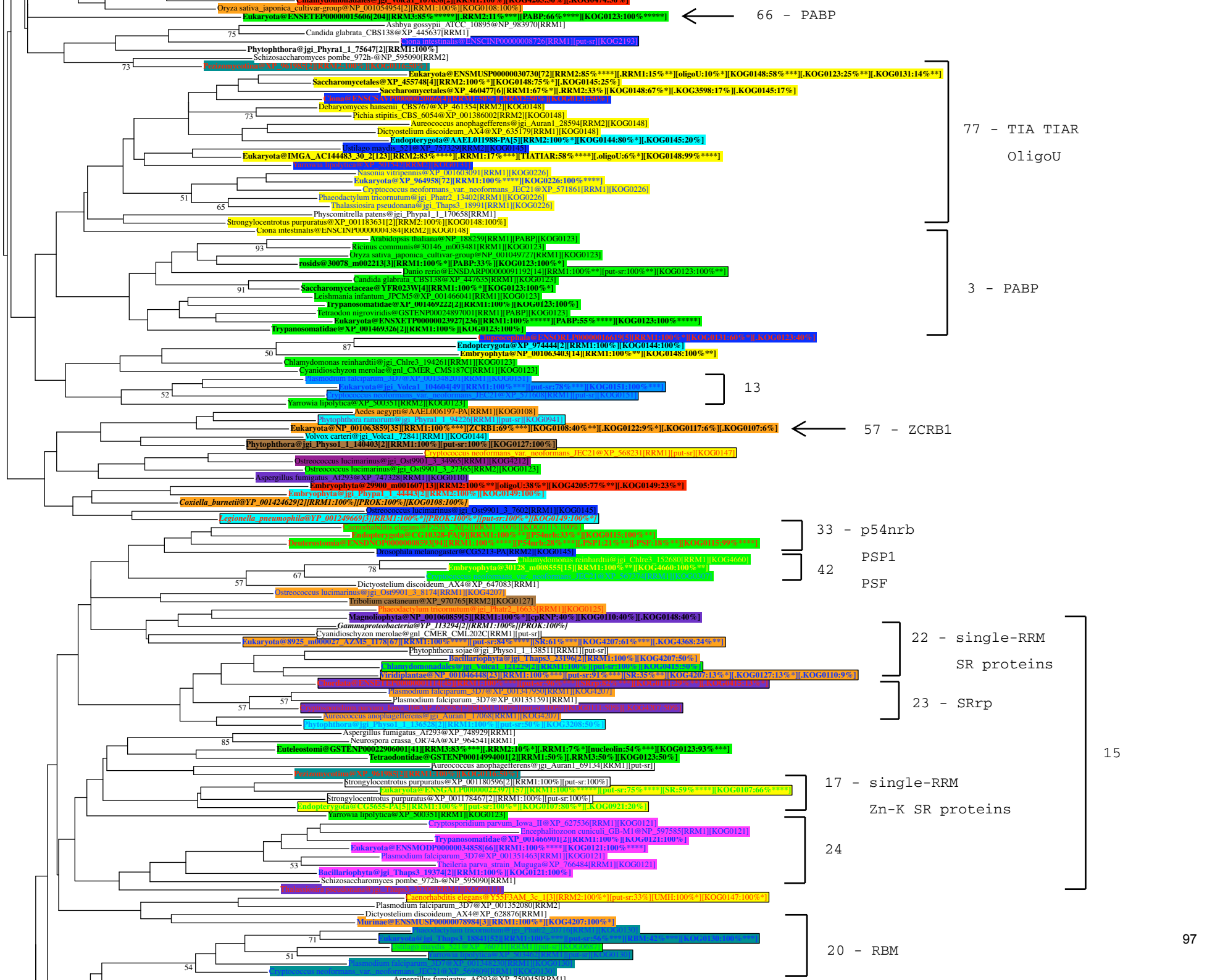
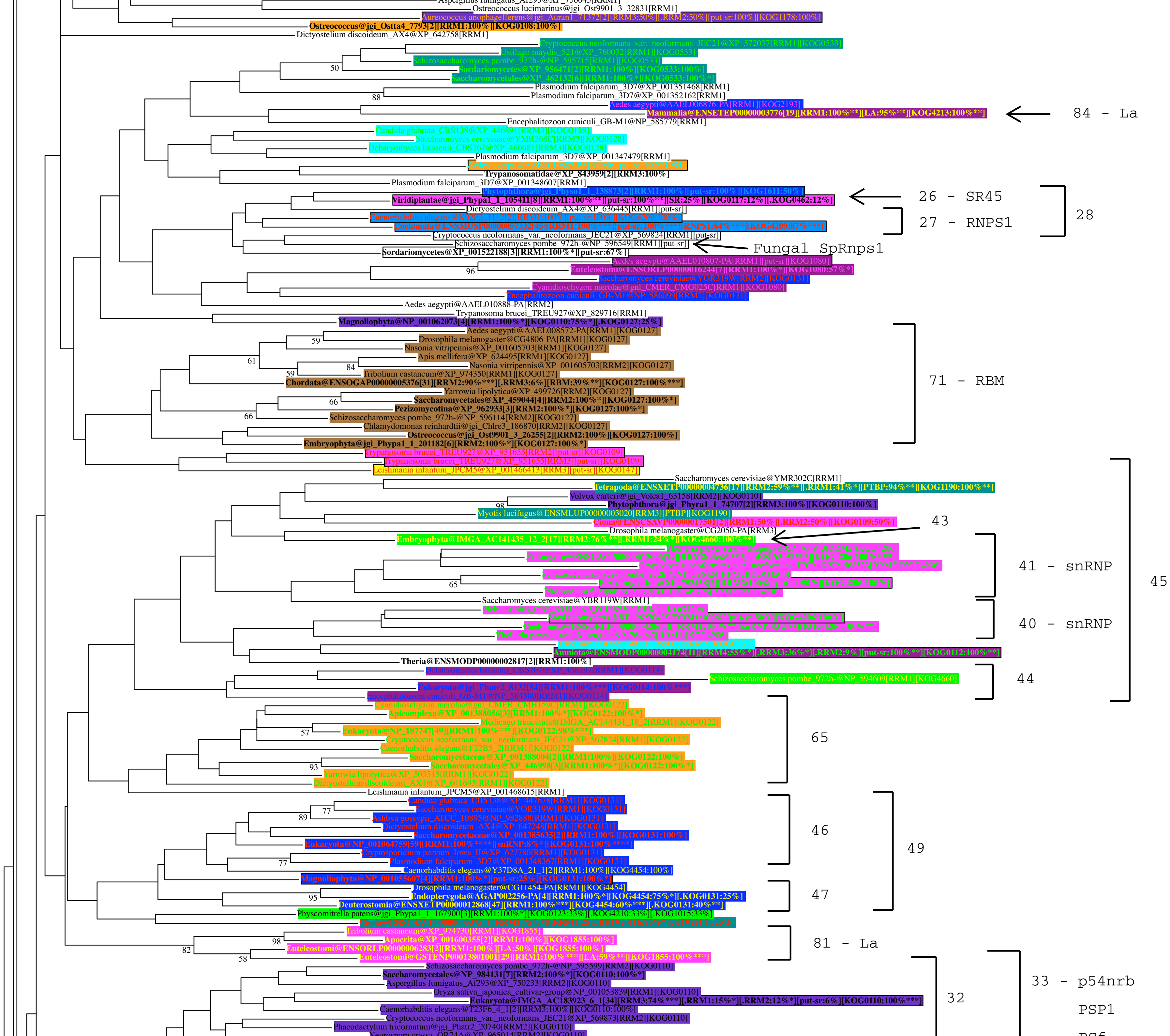

66 - PABP
77 - TIA TIAR
OligoU
3 - PABP
13
57 - ZCRB1
33 - p54nrb
PSP1
42
PSF
22 - single-RRM
SR proteins
23 - SRrp
15
17 - single-RRM
Zn-K SR proteins
24
20 - RBM
97

Fig. S7



Ostreococcus lucimarinus@jgi_Ost9901_3_32831[RRM1]
Aureococcus anophagefferens@jgi_Auran1_71372[2][RRM3:50%][.RRM2:50%][put-sr:100%][KOG1178:100%]
Ostreococcus@jgi_Ostta4_7793[2][RRM1:100%][KOG0108:100%]
Dictyostelium discoideum_AX4@XP_642758[RRM1]
Cryptococcus neoformans_var._neoformans_JEC21@XP_572037[RRM1][KOG0533]
Ustilago maydis_521@XP_760032[RRM1][KOG0533]
Schizosaccharomyces pombe_972h-@NP_595715[RRM1][KOG0533]
Sordariomycetes@XP_956471[2][RRM1:100%][KOG0533:100%]
Saccharomycetales@XP_462132[6][RRM1:100%*][KOG0533:100%*]
Plasmodium falciparum_3D7@XP_001351468[RRM1]
Plasmodium falciparum_3D7@XP_001352162[RRM1]
Aedes aegypti@AAEL006876-PA[RRM1][KOG2193]
Mammalia@ENSETEP00000003776[19][RRM1:100%**][LA:95%**][KOG4213:100%**]
84 - La
Encephalitozoon cuniculi_GB-M1@NP_585779[RRM1]
Candida glabrata_CBS138@XP_446891[RRM3][KOG0128]
Saccharomyces cerevisiae@YMR268C[RRM3][KOG0128]
Debaryomyces hansenii_CBS767@XP_460681[RRM3][KOG0128]
Plasmodium falciparum_3D7@XP_001347479[RRM1]
Trypanosomatidae@XP_843959[2][RRM3:100%]
Plasmodium falciparum_3D7@XP_001348607[RRM1]
Phytophthora@jgi_Physo1_1_138873[2][RRM1:100%][put-sr:100%][KOG1611:50%]
Viridiplantae@jgi_Phypa1_1_105411[8][RRM1:100%**][put-sr:100%**][SR:25%][KOG0117:12%][.KOG0462:12%]
26 - SR45
Dictyostelium discoideum_AX4@XP_636445[RRM1][put-sr]
27 - RNPS1
28
Cryptococcus neoformans_var._neoformans_JEC21@XP_569824[RRM1][put-sr]
Schizosaccharomyces pombe_972h-@NP_596549[RRM1][put-sr]
Fungal SpRnps1
Sordariomycetes@XP_001522188[3][RRM1:100%*][put-sr:67%]
Aedes aegypti@AAEL010807-PA[RRM1][put-sr][KOG1080]
Euteleostomi@ENSORLP00000016244[7][RRM1:100%*][KOG1080:57%*]
Saccharomyces cerevisiae@YOR319W[RRM2][KOG0131]
Cyanidioschyzon merolae@gnl_CMER_CMG025C[RRM1][KOG1080]
Aedes aegypti@AAEL010888-PA[RRM2]
Trypanosoma brucei_TREU927@XP_829716[RRM1]
Magnoliophyta@NP_001062073[4][RRM1:100%*][KOG0110:75%*][.KOG0127:25%]
Aedes aegypti@AAEL008572-PA[RRM1][KOG0127]
Drosophila melanogaster@CG4806-PA[RRM1][KOG0127]
Nasonia vitripennis@XP_001605703[RRM1][KOG0127]
Apis mellifera@XP_624495[RRM1][KOG0127]
Nasonia vitripennis@XP_001605703[RRM2][KOG0127]
Tribolium castaneum@XP_974350[RRM1][KOG0127]
Chordata@ENSOGAP00000005376[31][RRM2:90%***][.RRM3:6%][RBM:39%**][KOG0127:100%***]
71 - RBM
Yarrowia lipolytica@XP_499726[RRM2][KOG0127]
Saccharomycetales@XP_459044[4][RRM2:100%*][KOG0127:100%*]
Pezizomycotina@XP_962933[3][RRM2:100%*][KOG0127:100%*]
Schizosaccharomyces pombe_972h-@NP_596114[RRM2][KOG0127]
Chlamydomonas reinhardtii@jgi_Chlre3_186870[RRM2][KOG0127]
Ostreococcus@jgi_Ost9901_3_26255[2][RRM2:100%][KOG0127:100%]
Embryophyta@jgi_Phypa1_1_201182[6][RRM2:100%*][KOG0127:100%*]
Trypanosoma brucei_TREU927@XP_951655[RRM2][put-sr][KOG0109]
Trypanosoma brucei_TREU927@XP_951655[RRM3][put-sr][KOG0109]
Leishmania infantum_JPCM5@XP_001466413[RRM3][put-sr][KOG0147]
Saccharomyces cerevisiae@YMR302C[RRM1]
Tetrapoda@ENSXETP00000004736[17][RRM2:59%**][.RRM1:41%*][PTBP:94%**][KOG1190:100%**]
Volvox carteri@jgi_Volca1_63158[RRM2][KOG0110]
Phytophthora@jgi_Phyra1_1_74707[2][RRM3:100%][KOG0110:100%]
Myotis lucifugus@ENSMLUP00000003020[RRM3][PTBP][KOG1190]
Ciona@ENSCSAVP00000017501[2][RRM1:50%][.RRM2:50%][KOG0109:50%]
43
Drosophila melanogaster@CG2050-PA[RRM3]
Embryophyta@IMGA_AC141435_12_2[17][RRM2:76%**][.RRM1:24%*][KOG4660:100%**]
41 - snRNP
45
Saccharomyces cerevisiae@YBR119W[RRM1]
40 - snRNP
Theria@ENSMODP00000002817[2][RRM1:100%]
Debaryomyces hansenii_CBS767@XP_458197[RRM1][KOG0114]
Schizosaccharomyces pombe_972h-@NP_594609[RRM1][KOG4660]
44
Eukaryota@jgi_Phatr2_8132[54][RRM1:100%***][KOG0114:100%***]
Encephalitozoon cuniculi_GB-M1@NP_584580[RRM1][KOG0114]
Cyanidioschyzon merolae@gnl_CMER_CMH189C[RRM1][KOG0122]
Apicomplexa@XP_001388056[3][RRM1:100%*][KOG0122:100%*]
Medicago truncatula@IMGA_AC144431_18_2[RRM1][KOG0122]
Eukaryota@NP_187747[49][RRM1:100%***][KOG0122:98%***]
Cryptococcus neoformans_var._neoformans_JEC21@XP_567824[RRM1][KOG0122]
Caenorhabditis elegans@F22B5_2[RRM1][KOG0122]
Saccharomycetaceae@XP_001388004[2][RRM1:100%][KOG0122:100%]
Saccharomycetales@XP_446998[3][RRM1:100%*][KOG0122:100%*]
Yarrowia lipolytica@XP_503515[RRM1][KOG0122]
Dictyostelium discoideum_AX4@XP_641693[RRM1][KOG0122]
65
Leishmania infantum_JPCM5@XP_001468615[RRM1]
Candida glabrata_CBS138@XP_447678[RRM1][KOG0131]
Saccharomyces cerevisiae@YOR319W[RRM1][KOG0131]
Ashbya gossypii_ATCC_10895@NP_982888[RRM1][KOG0131]
Dictyostelium discoideum_AX4@XP_647248[RRM1][KOG0131]
Saccharomycetaceae@XP_001385635[2][RRM1:100%][KOG0131:100%]
Eukaryota@NP_001064759[59][RRM1:100%****][snRNP:8%*][KOG0131:100%****]
Cryptosporidium parvum_Iowa_II@XP_627780[RRM1][KOG0131]
Plasmodium falciparum_3D7@XP_001348367[RRM1][KOG0131]
Caenorhabditis elegans@Y37D8A_21_1[2][RRM1:100%][KOG4454:100%]
Magnoliophyta@NP_001055607[4][RRM1:100%*][put-sr:25%][KOG0131:100%*]
46
49
Drosophila melanogaster@CG11454-PA[RRM1][KOG4454]
Endopterygota@AGAP002256-PA[4][RRM1:100%*][KOG4454:75%*][.KOG0131:25%]
Deuterostomia@ENSXETP00000012868[47][RRM1:100%***][KOG4454:60%***][.KOG0131:40%***]
Physcomitrella patens@jgi_Phypa1_1_167900[3][RRM1:100%*][KOG0123:33%][.KOG4210:33%][.KOG1015:33%]
47
Tribolium castaneum@XP_974730[RRM1][KOG1855]
Apocrita@XP_001600355[2][RRM1:100%][KOG1855:100%]
Euteleostomi@ENSORLP00000006283[2][RRM1:100%][LA:50%][KOG1855:100%]
Euteleostomi@GSTENP00013801001[29][RRM1:100%***][LA:59%**][KOG1855:100%***]
81 - La
Schizosaccharomyces pombe_972h-@NP_595599[RRM2][KOG0110]
Saccharomycetales@NP_984131[7][RRM2:100%*][KOG0110:100%*]
Aspergillus fumigatus_Af293@XP_750233[RRM2][KOG0110]
Oryza sativa_japonica_cultivar-group@NP_001053839[RRM1][KOG0110]
Eukaryota@IMGA_AC183923_6_1[34][RRM3:74%***][.RRM1:15%*][.RRM2:12%*][put-sr:6%][KOG0110:100%***]
Caenorhabditis elegans@T23F6_4_1[2][RRM3:100%][KOG0110:100%]
Cryptococcus neoformans_var._neoformans_JEC21@XP_569873[RRM2][KOG0110]
Phaeodactylum tricornutum@jgi_Phatr2_20740[RRM2][KOG0110]
32
33 - p54nrb
PSP1

Fig. S7

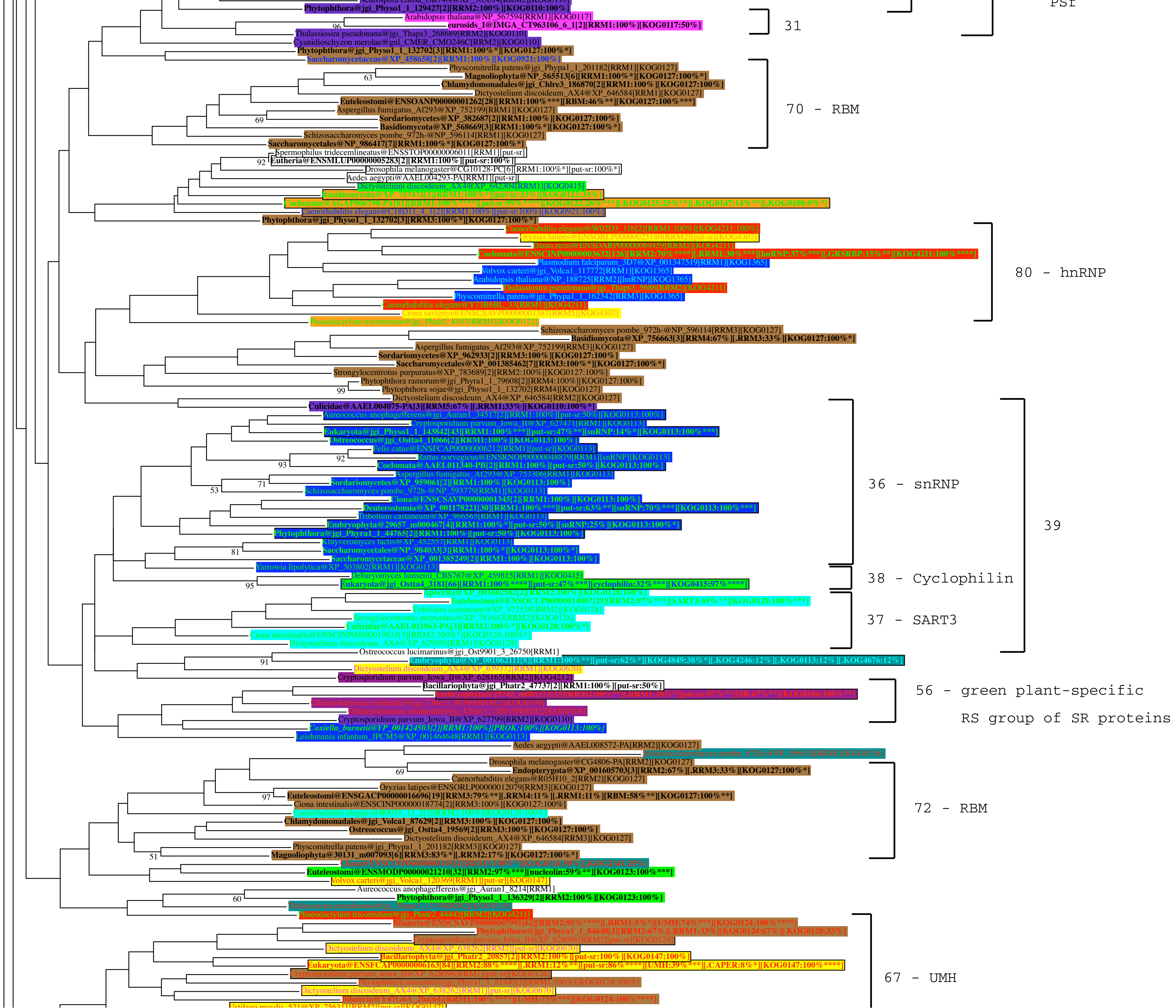
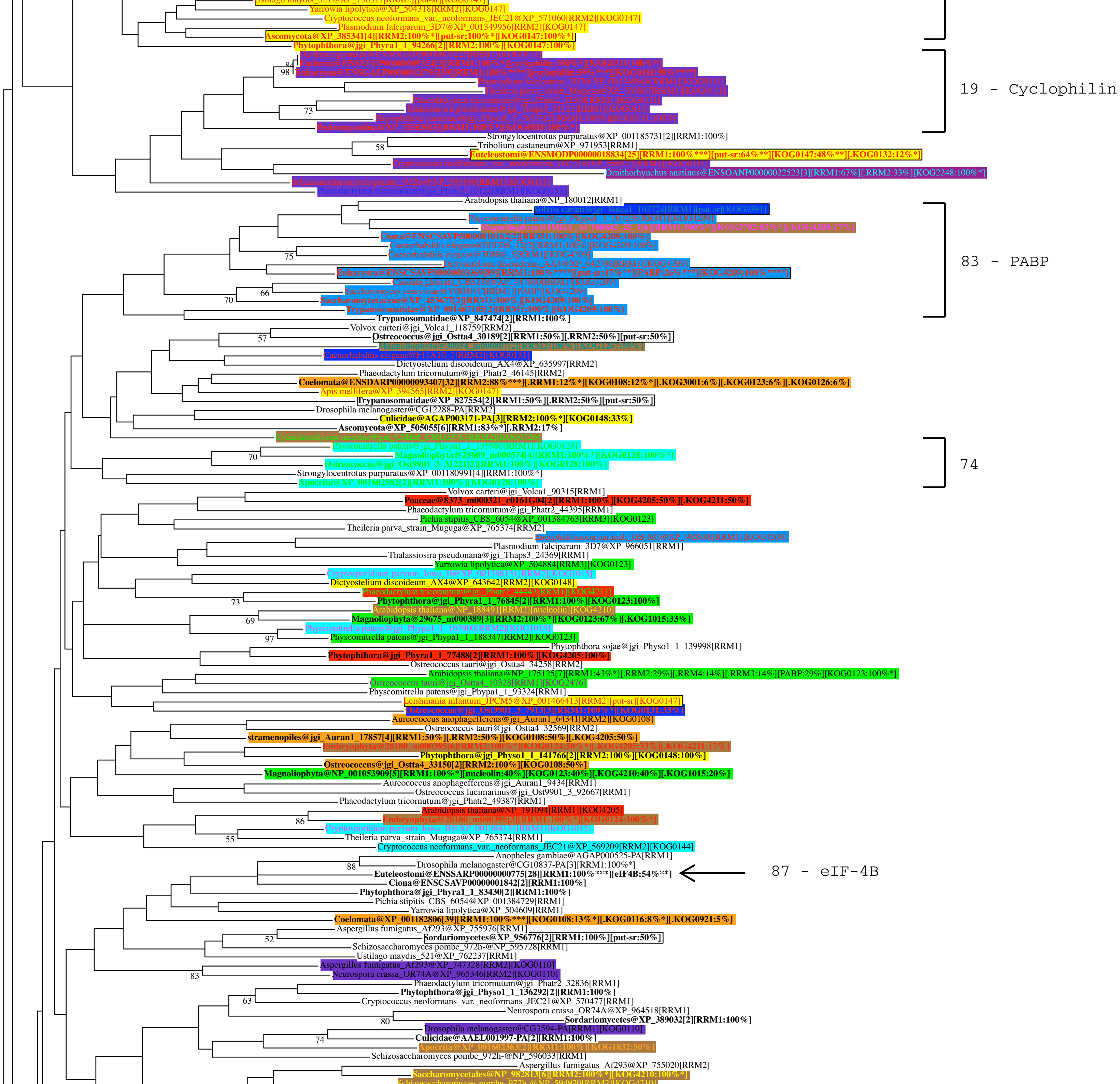
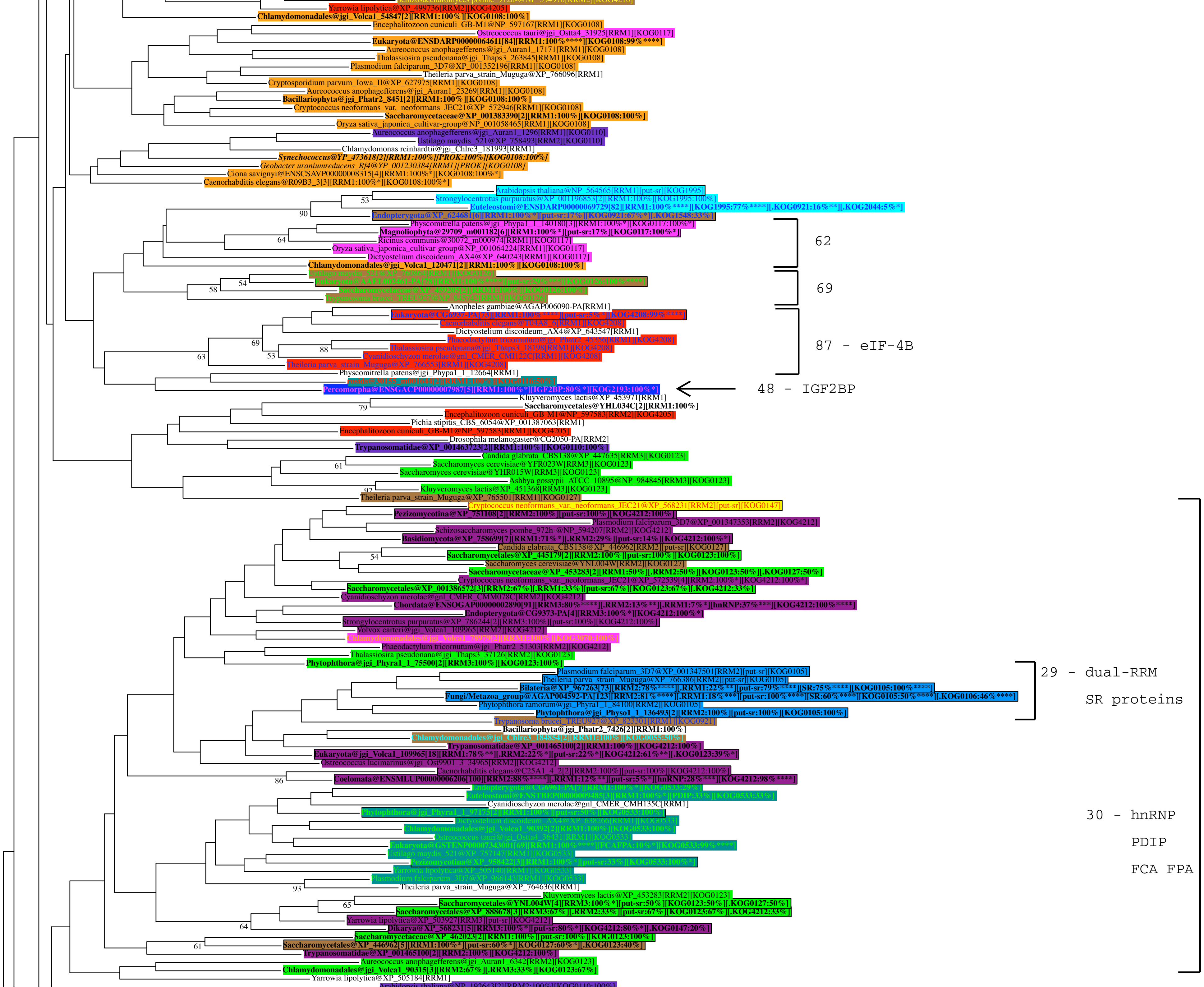

PSF
31
70 - RBM
80 - hnRNP
36 - snRNP
39
38 - Cyclophilin
37 - SART3
56 - green plant-specific
RS group of SR proteins
72 - RBM
67 - UMH

Fig. S7



Yarrowia lipolytica@XP_504318[RRM2][KOG0147]
Cryptococcus neoformans_var._neoformans_JEC21@XP_571060[RRM2][KOG0147]
Plasmodium falciparum_3D7@XP_001349956[RRM2][KOG0147]
Ascomycota@XP_385341[4][RRM2:100%*][put-sr:100%*][KOG0147:100%*]
Phytophthora@jgi_Phyra1_1_94266[2][RRM2:100%][KOG0147:100%]
19 - Cyclophilin
98
73
Strongylocentrotus purpuratus@XP_001185731[2][RRM1:100%]
Tribolium castaneum@XP_971953[RRM1]
58
Euteleostomi@ENSMODP00000018834[25][RRM1:100%***][put-sr:64%**][KOG0147:48%**][.KOG0132:12%*]
Ornithorhynchus anatinus@ENSOANP00000022523[3][RRM1:67%][.RRM2:33%][KOG2248:100%*]
Arabidopsis thaliana@NP_180012[RRM1]
83 - PABP
75
70
66
Trypanosomatidae@XP_847474[2][RRM1:100%]
Volvox carteri@jgi_Volca1_118759[RRM2]
57
Ostreococcus@jgi_Ostta4_30189[2][RRM1:50%][.RRM2:50%][put-sr:50%]
Dictyostelium discoideum_AX4@XP_635997[RRM2]
Phaeodactylum tricornutum@jgi_Phatr2_46145[RRM2]
Coelomata@ENSDARP00000093407[32][RRM2:88%***][.RRM1:12%*][KOG0108:12%*][.KOG3001:6%][.KOG0123:6%][.KOG0126:6%]
Apis mellifera@XP_394565[RRM2][KOG0147]
Trypanosomatidae@XP_827554[2][RRM1:50%][.RRM2:50%][put-sr:50%]
Drosophila melanogaster@CG12288-PA[RRM2]
Culicidae@AGAP003171-PA[3][RRM2:100%*][KOG0148:33%]
Ascomycota@XP_505055[6][RRM1:83%*][.RRM2:17%]
74
Physcomitrella patens@jgi_Phypa1_1_135980[RRM1][KOG0128]
70
Magnoliophyta@29609_m000574[4][RRM1:100%*][KOG0128:100%*]
Ostreococcus@jgi_Ost9901_3_31221[2][RRM1:100%][KOG0128:100%]
Strongylocentrotus purpuratus@XP_001180991[4][RRM1:100%*]
Apocrita@XP_001602582[2][RRM1:100%][KOG0128:100%]
Volvox carteri@jgi_Volca1_90315[RRM1]
Poaceae@8373_m000321_c0161G04[2][RRM1:100%][KOG4205:50%][.KOG4211:50%]
Phaeodactylum tricornutum@jgi_Phatr2_44395[RRM1]
Pichia stipitis_CBS_6054@XP_001384763[RRM3][KOG0123]
Theileria parva_strain_Muguga@XP_765374[RRM2]
Encephalitozoon cuniculi_GB-M1@XP_965888[RRM1][KOG4209]
Plasmodium falciparum_3D7@XP_966051[RRM1]
Thalassiosira pseudonana@jgi_Thaps3_24369[RRM1]
Yarrowia lipolytica@XP_504884[RRM3][KOG0123]
Cryptosporidium parvum_Iowa_II@XP_001388131[RRM2][KOG1015]
Dictyostelium discoideum_AX4@XP_643642[RRM2][KOG0148]
Phaeodactylum tricornutum@jgi_Phatr2_44442[RRM1][KOG4211]
73
Phytophthora@jgi_Phyra1_1_76845[2][RRM1:100%][KOG0123:100%]
Arabidopsis thaliana@NP_188491[RRM2][nucleolin][KOG4210]
69
Magnoliophyta@29675_m000389[3][RRM2:100%*][KOG0123:67%][.KOG1015:33%]
Physcomitrella patens@jgi_Phypa1_1_167900[RRM2][KOG1015]
97
Physcomitrella patens@jgi_Phypa1_1_188347[RRM2][KOG0123]
Phytophthora sojae@jgi_Physo1_1_139998[RRM1]
Phytophthora@jgi_Phyra1_1_77488[2][RRM1:100%][KOG4205:100%]
Ostreococcus tauri@jgi_Ostta4_34258[RRM2]
Arabidopsis thaliana@NP_175125[7][RRM1:43%*][.RRM2:29%][.RRM4:14%][.RRM3:14%][PABP:29%][KOG0123:100%*]
Ostreococcus tauri@jgi_Ostta4_10328[RRM1][KOG2476]
Physcomitrella patens@jgi_Phypa1_1_93324[RRM1]
Leishmania infantum_JPCM5@XP_001466413[RRM2][put-sr][KOG0147]
Aureococcus anophagefferens@jgi_Auran1_64341[RRM2][KOG0108]
Ostreococcus tauri@jgi_Ostta4_32569[RRM2]
stramenopiles@jgi_Auran1_17857[4][RRM1:50%][.RRM2:50%][KOG0108:50%][.KOG4205:50%]
Phytophthora@jgi_Physo1_1_141766[2][RRM2:100%][KOG0148:100%]
Ostreococcus@jgi_Ostta4_33150[2][RRM2:100%][KOG0108:50%]
Magnoliophyta@NP_001053909[5][RRM1:100%*][nucleolin:40%][KOG0123:40%][.KOG4210:40%][.KOG1015:20%]
Aureococcus anophagefferens@jgi_Auran1_9434[RRM1]
Ostreococcus lucimarinus@jgi_Ost9901_3_92667[RRM1]
Phaeodactylum tricornutum@jgi_Phatr2_49387[RRM1]
Arabidopsis thaliana@NP_191094[RRM1][KOG4205]
86
Cryptosporidium parvum_Iowa_II@XP_001388131[RRM1][KOG1015]
55
Theileria parva_strain_Muguga@XP_765374[RRM1]
Cryptococcus neoformans_var._neoformans_JEC21@XP_569209[RRM2][KOG0144]
Anopheles gambiae@AGAP000525-PA[RRM1]
88
Drosophila melanogaster@CG10837-PA[3][RRM1:100%*]
Euteleostomi@ENSSARP00000000775[28][RRM1:100%***][eIF4B:54%**]
87 - eIF-4B
Ciona@ENSCSAVP00000001842[2][RRM1:100%]
Phytophthora@jgi_Phyra1_1_83430[2][RRM1:100%]
Pichia stipitis_CBS_6054@XP_001384729[RRM1]
Yarrowia lipolytica@XP_504609[RRM1]
Coelomata@XP_001182806[39][RRM1:100%***][KOG0108:13%*][.KOG0116:8%*][.KOG0921:5%]
Aspergillus fumigatus_Af293@XP_755976[RRM1]
52
Sordariomycetes@XP_956776[2][RRM1:100%][put-sr:50%]
Schizosaccharomyces pombe_972h-@NP_595728[RRM1]
Ustilago maydis_521@XP_762237[RRM1]
Aspergillus fumigatus_Af293@XP_747328[RRM2][KOG0110]
83
Neurospora crassa_OR74A@XP_965346[RRM2][KOG0110]
Phaeodactylum tricornutum@jgi_Phatr2_32836[RRM1]
Phytophthora@jgi_Physo1_1_136292[2][RRM1:100%]
63
Cryptococcus neoformans_var._neoformans_JEC21@XP_570477[RRM1]
Neurospora crassa_OR74A@XP_964518[RRM1]
80
Sordariomycetes@XP_389032[2][RRM1:100%]
Drosophila melanogaster@CG3594-PA[RRM1][KOG0110]
74
Culicidae@AAEL001997-PA[2][RRM1:100%]
Schizosaccharomyces pombe_972h-@NP_596033[RRM1]
Aspergillus fumigatus_Af293@XP_755020[RRM2]
Saccharomycetales@NP_982813[6][RRM2:100%*][KOG4210:100%*]

Fig. S7



62
69
87 - eIF-4B
48 - IGF2BP
29 - dual-RRM
SR proteins
30 - hnRNP
PDIP
FCA FPA

Fig. S7

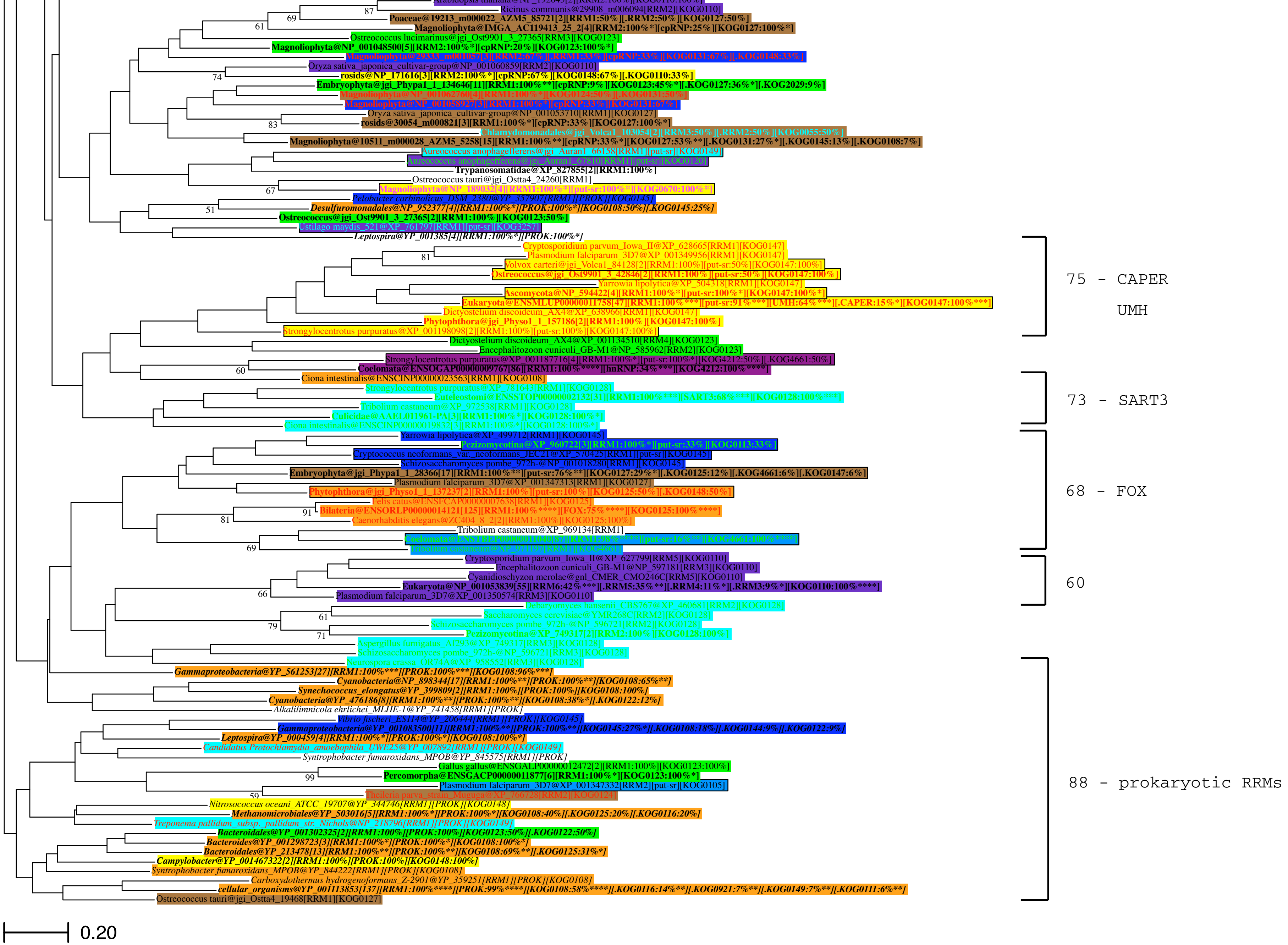


Fig. S8

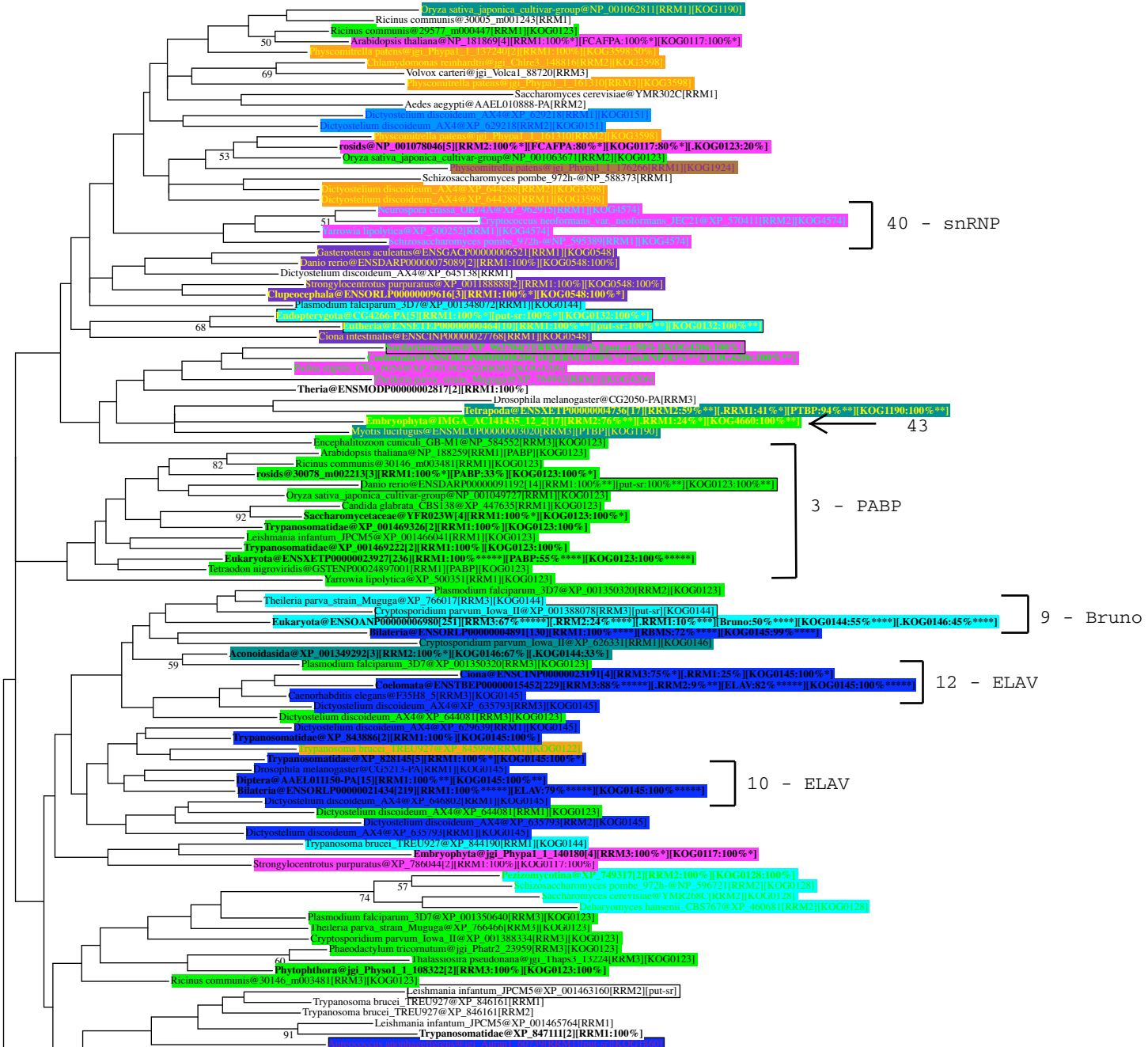

Oryza sativa_japonica_cultivar-group@NP_001062811[RRM1][KOG1190]
Ricinus communis@30005_m001243[RRM1]
Ricinus communis@29577_m000447[RRM1][KOG0123]
Arabidopsis thaliana@NP_181869[4][RRM1:100%*][FCAFPA:100%*][KOG0117:100%*]
Physcomitrella patens@jgi_Phypa1_1_137240[2][RRM1:100%*][KOG3598:50%*]
Chlamydomonas reinhardtii@jgi_Chlre3_148816[RRM2][KOG3598]
Volvox carteri@jgi_Volca1_88720[RRM3]
Physcomitrella patens@jgi_Phypa1_1_161310[RRM3][KOG3598]
Saccharomyces cerevisiae@YMR302C[RRM1]
Aedes aegypti@AAEL010888-PA[RRM2]
Dictyostelium discoideum_AX4@XP_629218[RRM1][KOG0151]
Dictyostelium discoideum_AX4@XP_629218[RRM2][KOG0151]
rosids@NP_001078046[5][RRM2:100%*][FCAFPA:80%*][KOG0117:80%*][.KOG0123:20%*]
Oryza sativa_japonica_cultivar-group@NP_001063671[RRM2][KOG0123]
Schizosaccharomyces pombe_972h-@NP_588373[RRM1]
Dictyostelium discoideum_AX4@XP_644288[RRM2][KOG3598]
Dictyostelium discoideum_AX4@XP_644288[RRM1][KOG3598]
Gasterosteus aculeatus@ENSGACP00000006521[RRM1][KOG0548]
Danio rerio@ENSDARP00000075089[2][RRM1:100%][KOG0548:100%]
Dictyostelium discoideum_AX4@XP_645138[RRM1]
Strongylocentrotus purpuratus@XP_001188888[2][RRM1:100%][KOG0548:100%]
Clupeocephala@ENSORLP00000009616[3][RRM1:100%*][KOG0548:100%*]
Plasmodium falciparum_3D7@XP_001348072[RRM1][KOG0144]
Endopterygota@CG4266-PA[5][RRM1:100%*][put-sr:100%*][KOG0132:100%*]
Euteria@ENSLLEP00000003464[19][RRM1:100%*][put-sr:100%**][KOG0132:100%**]
Ciona intestinalis@ENSCINP00000027668[RRM1][KOG0548]
Sordariomycetes@XP_962704[2][RRM1:100%][put-sr:50%][KOG4206:100%]
Coelomata@ENSORLP00000008268[18][RRM1:100%***][snRNP:83%***][KOG4206:100%***]
Pichia stipitis_CBS_6054@XP_001382592[RRM1][KOG4206]
Theileria parva_strain_Muguga@XP_764943[RRM1][KOG4206]
Theria@ENSMODP00000002817[2][RRM1:100%]
Drosophila melanogaster@CG2050-PA[RRM3]
Tetrapoda@ENSXETP00000004736[17][RRM2:59%**][.RRM1:41%*][PTBP:94%**][KOG1190:100%**]
Embryophyta@IMGA_AC141435_12_2[17][RRM2:76%**][.RRM1:24%*][KOG4660:100%**]
Myotis lucifugus@ENSMLUP00000003020[RRM3][PTBP][KOG1190]
Encephalitozoon cuniculi_GB-M1@NP_584552[RRM3][KOG0123]
Arabidopsis thaliana@NP_188259[RRM1][PABP][KOG0123]
Ricinus communis@30146_m003481[RRM1][KOG0123]
rosids@30078_m002213[3][RRM1:100%*][PABP:33%][KOG0123:100%*]
Danio rerio@ENSDARP00000091192[14][RRM1:100%**][put-sr:100%**][KOG0123:100%**]
Oryza sativa_japonica_cultivar-group@NP_001049727[RRM1][KOG0123]
Candida glabrata_CBS138@XP_447635[RRM1][KOG0123]
Saccharomycetaceae@YFR023W[4][RRM1:100%*][KOG0123:100%*]
Trypanosomatidae@XP_001469326[2][RRM1:100%][KOG0123:100%]
Leishmania infantum_JPCM5@XP_001466041[RRM1][KOG0123]
Trypanosomatidae@XP_001469222[2][RRM1:100%][KOG0123:100%]
Eukaryota@ENSXETP00000023927[236][RRM1:100%*****][PABP:55%****][KOG0123:100%*****]
Tetraodon nigroviridis@GSTENP00024897001[RRM1][PABP][KOG0123]
Yarrowia lipolytica@XP_500351[RRM1][KOG0123]
Plasmodium falciparum_3D7@XP_001350320[RRM2][KOG0123]
Theileria parva_strain_Muguga@XP_766017[RRM3][KOG0144]
Cryptosporidium parvum_Iowa_II@XP_001388078[RRM3][put-sr][KOG0144]
Eukaryota@ENSOANP00000006980[251][RRM3:67%*****][.RRM2:24%****][.RRM1:10%***][Bruno:50%****][KOG0144:55%****][.KOG0146:45%****]
Bilateria@ENSORLP00000004891[130][RRM1:100%*****][RBMS:72%*****][KOG0145:99%*****]
Cryptosporidium parvum_Iowa_II@XP_626331[RRM1][KOG0146]
Aconoidasida@XP_001349292[3][RRM2:100%*][KOG0146:67%][.KOG0144:33%]
Plasmodium falciparum_3D7@XP_001350320[RRM3][KOG0123]
Ciona@ENSCINP00000022191[4][RRM3:75%*][.RRM1:25%][KOG0145:100%*]
Coelomata@ENSTBEP00000015452[229][RRM3:88%*****][.RRM2:9%**][ELAV:82%*****][KOG0145:100%*****]
Caenorhabditis elegans@F35H8_5[RRM3][KOG0145]
Dictyostelium discoideum_AX4@XP_635793[RRM3][KOG0145]
Dictyostelium discoideum_AX4@XP_644081[RRM3][KOG0123]
Dictyostelium discoideum_AX4@XP_629639[RRM1][KOG0145]
Trypanosomatidae@XP_843886[2][RRM1:100%][KOG0145:100%]
Trypanosoma brucei_TREU927@XP_845996[RRM1][KOG0122]
Trypanosomatidae@XP_828145[5][RRM1:100%*][KOG0145:100%*]
Drosophila melanogaster@CG5213-PA[RRM1][KOG0145]
Diptera@AAEL011150-PA[15][RRM1:100%**][KOG0145:100%**]
Bilateria@ENSORLP00000021434[219][RRM1:100%*****][ELAV:79%*****][KOG0145:100%*****]
Dictyostelium discoideum_AX4@XP_646802[RRM1][KOG0145]
Dictyostelium discoideum_AX4@XP_644081[RRM1][KOG0123]
Dictyostelium discoideum_AX4@XP_635793[RRM2][KOG0145]
Dictyostelium discoideum_AX4@XP_635793[RRM1][KOG0145]
Trypanosoma brucei_TREU927@XP_844190[RRM1][KOG0144]
Embryophyta@jgi_Phypa1_1_140180[4][RRM3:100%*][KOG0117:100%*]
Strongylocentrotus purpuratus@XP_786044[2][RRM1:100%][KOG0117:100%]
Pezizomycotina@XP_749317[2][RRM2:100%][KOG0128:100%]
Schizosaccharomyces pombe_972h-@NP_596721[RRM2][KOG0128]
Saccharomyces cerevisiae@YMR268C[RRM2][KOG0128]
Debaryomyces hansenii_CBS767@XP_460681[RRM2][KOG0128]
Plasmodium falciparum_3D7@XP_001350640[RRM3][KOG0123]
Theileria parva_strain_Muguga@XP_766466[RRM3][KOG0123]
Cryptosporidium parvum_Iowa_II@XP_001388334[RRM3][KOG0123]
Phaeodactylum tricornutum@jgi_Phatr2_23959[RRM3][KOG0123]
Thalassiosira pseudonana@jgi_Thaps3_13224[RRM3][KOG0123]
Phytophthora@jgi_Physo1_1_108322[2][RRM3:100%][KOG0123:100%]
Ricinus communis@30146_m003481[RRM3][KOG0123]
Leishmania infantum_JPCM5@XP_001463160[RRM2][put-sr]
Trypanosoma brucei_TREU927@XP_846161[RRM1]
Trypanosoma brucei_TREU927@XP_846161[RRM2]
Leishmania infantum_JPCM5@XP_001465764[RRM1]
Trypanosomatidae@XP_847111[2][RRM1:100%]
40 - snRNP
43
3 - PABP
9 - Bruno
12 - ELAV
10 - ELAV

Fig. S8

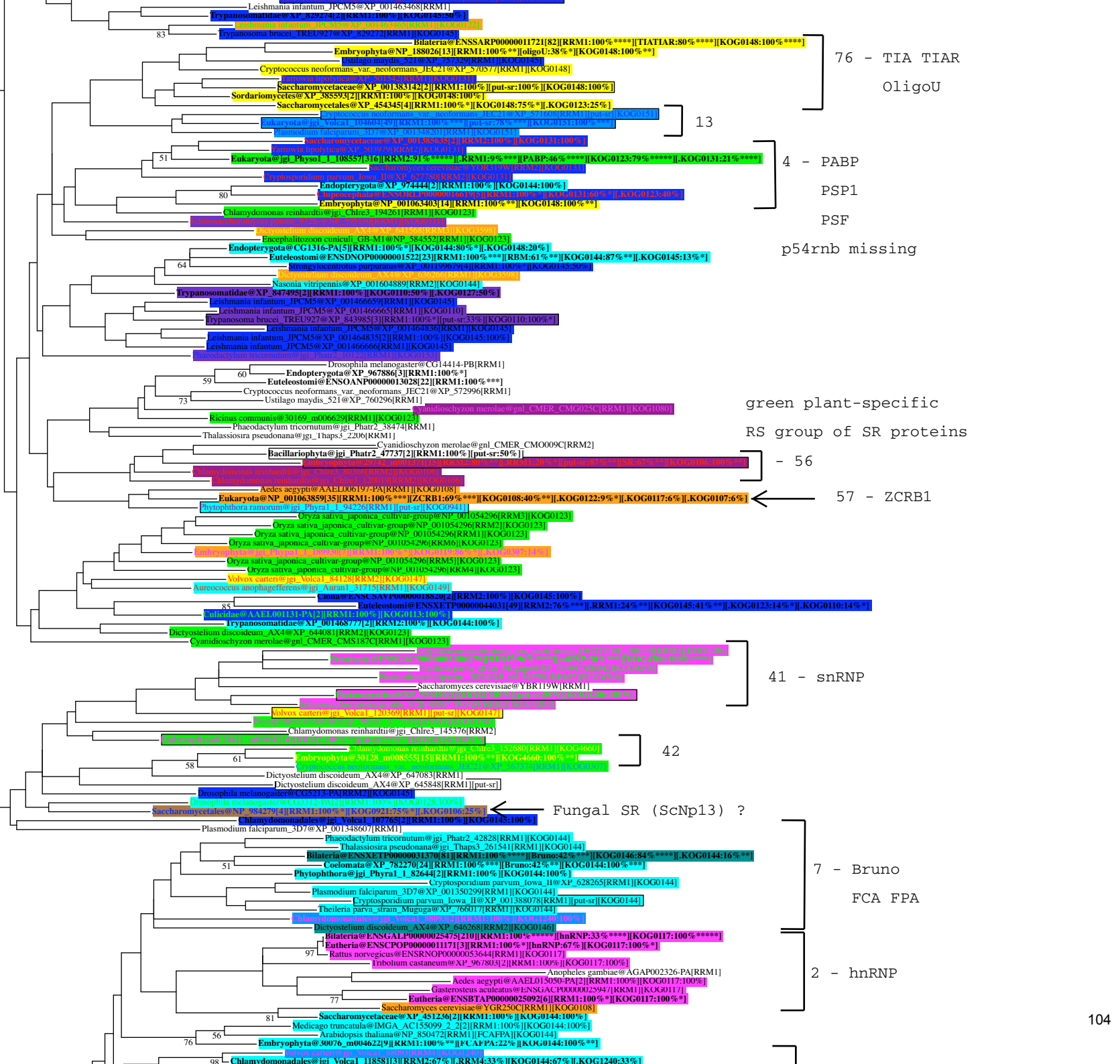

76 - TIA TIAR
OligoU
13
4 - PABP
PSP1
PSF
p54rnb missing
green plant-specific
RS group of SR proteins
- 56
57 - ZCRB1
41 - snRNP
Fungal SR (ScNp13) ?
42
7 - Bruno
FCA FPA
2 - hnRNP

Fig. S8

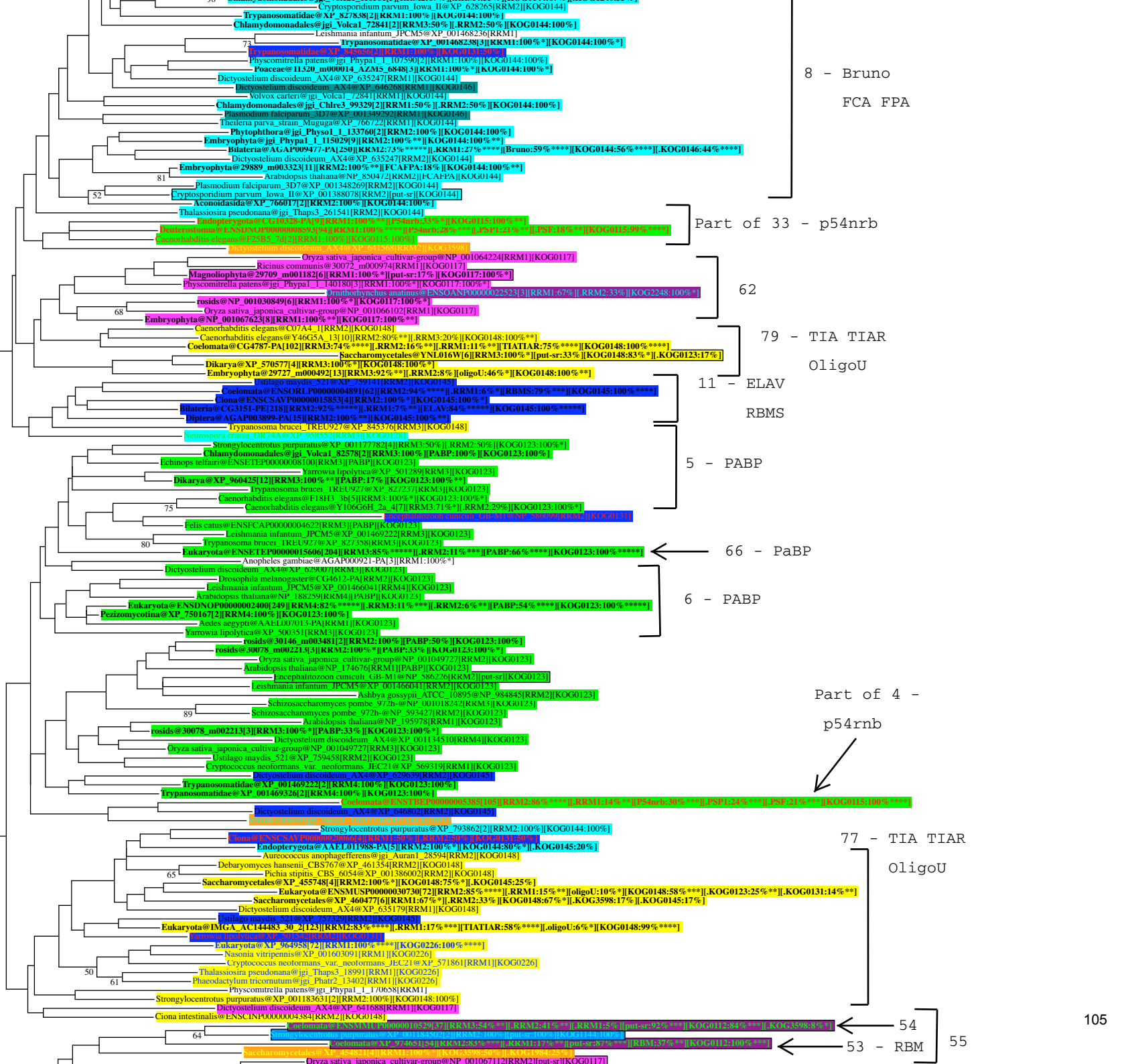
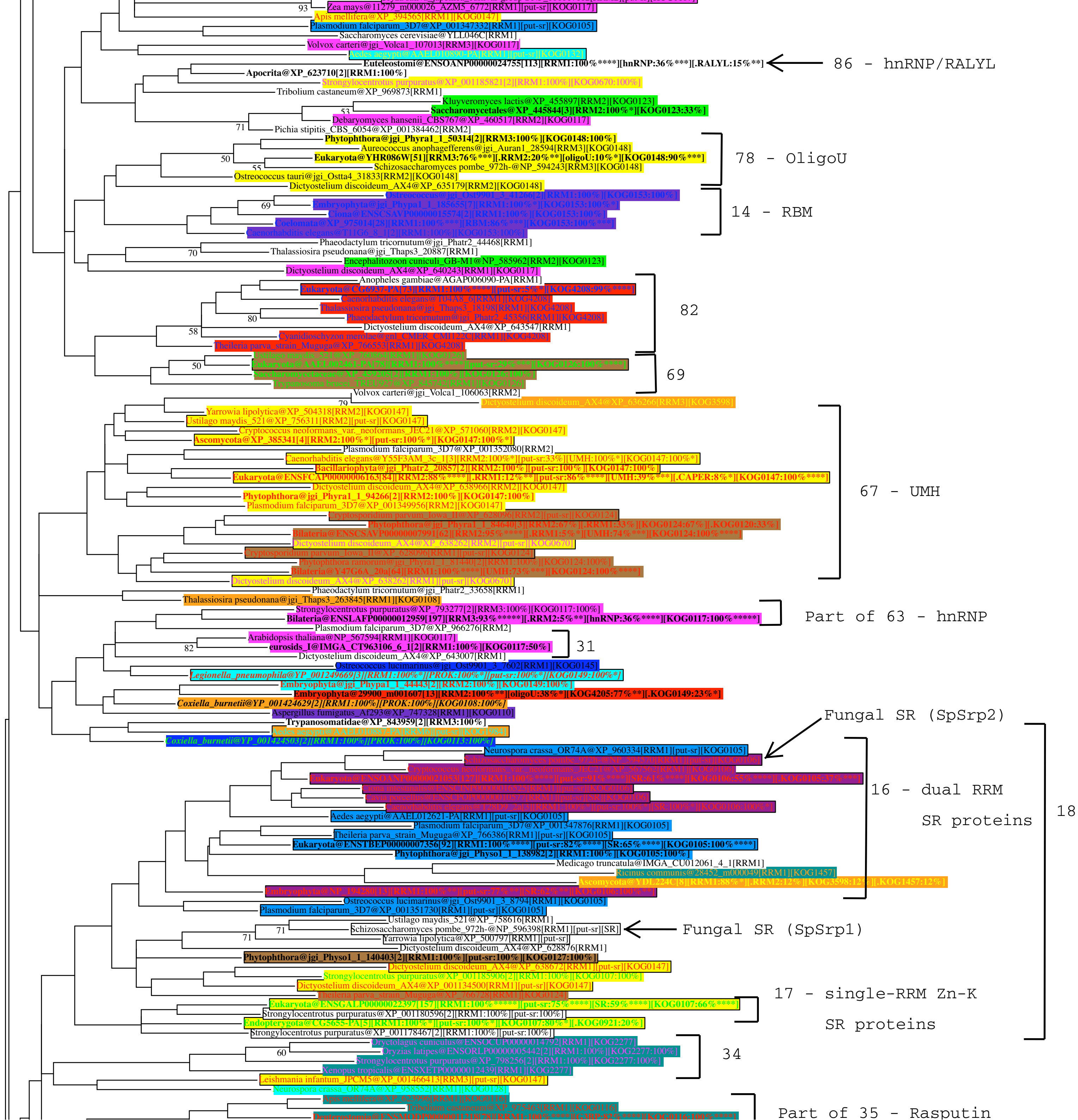

8 - Bruno
FCA FPA
Part of 33 - p54nrb
62
79 - TIA TIAR
OligoU
11 - ELAV
RBMS
5 - PABP
66 - PaBP
6 - PABP
Part of 4 -
p54rnb
77 - TIA TIAR
OligoU
54
53 - RBM
55

Fig. S8



86 - hnRNP/RALYL
78 - OligoU
14 - RBM
82
69
67 - UMH
Part of 63 - hnRNP
31
Fungal SR (SpSrp2)
16 - dual RRM
SR proteins
18
Fungal SR (SpSrp1)
17 - single-RRM Zn-K
SR proteins
34
Part of 35 - Rasputin

Fig. S8

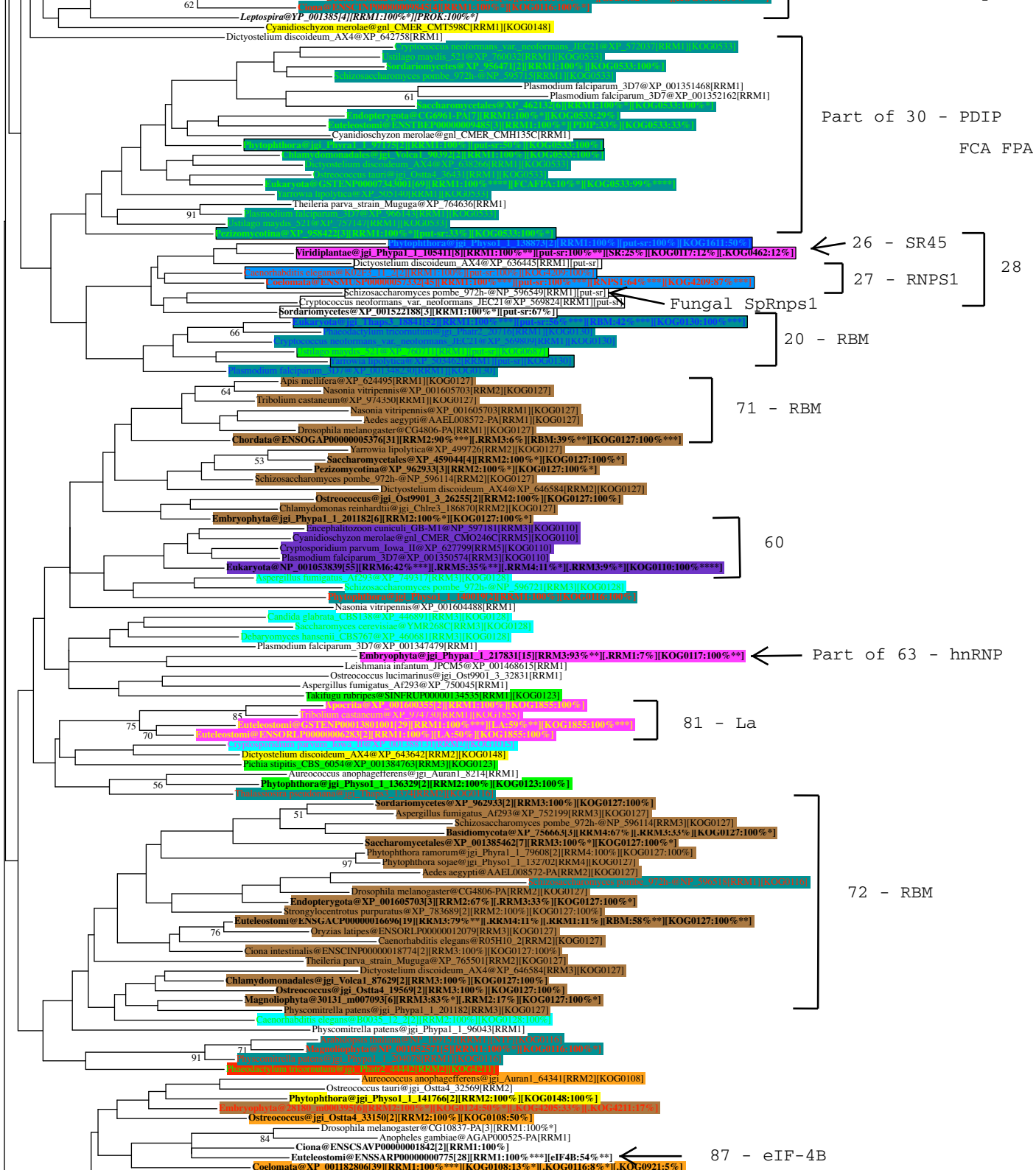

Part of 30 - PDIP
FCA FPA
26 - SR45
27 - RNPS1
28
Fungal SpRnps1
20 - RBM
71 - RBM
60
Part of 63 - hnRNP
81 - La
72 - RBM
87 - eIF-4B

Fig. S8

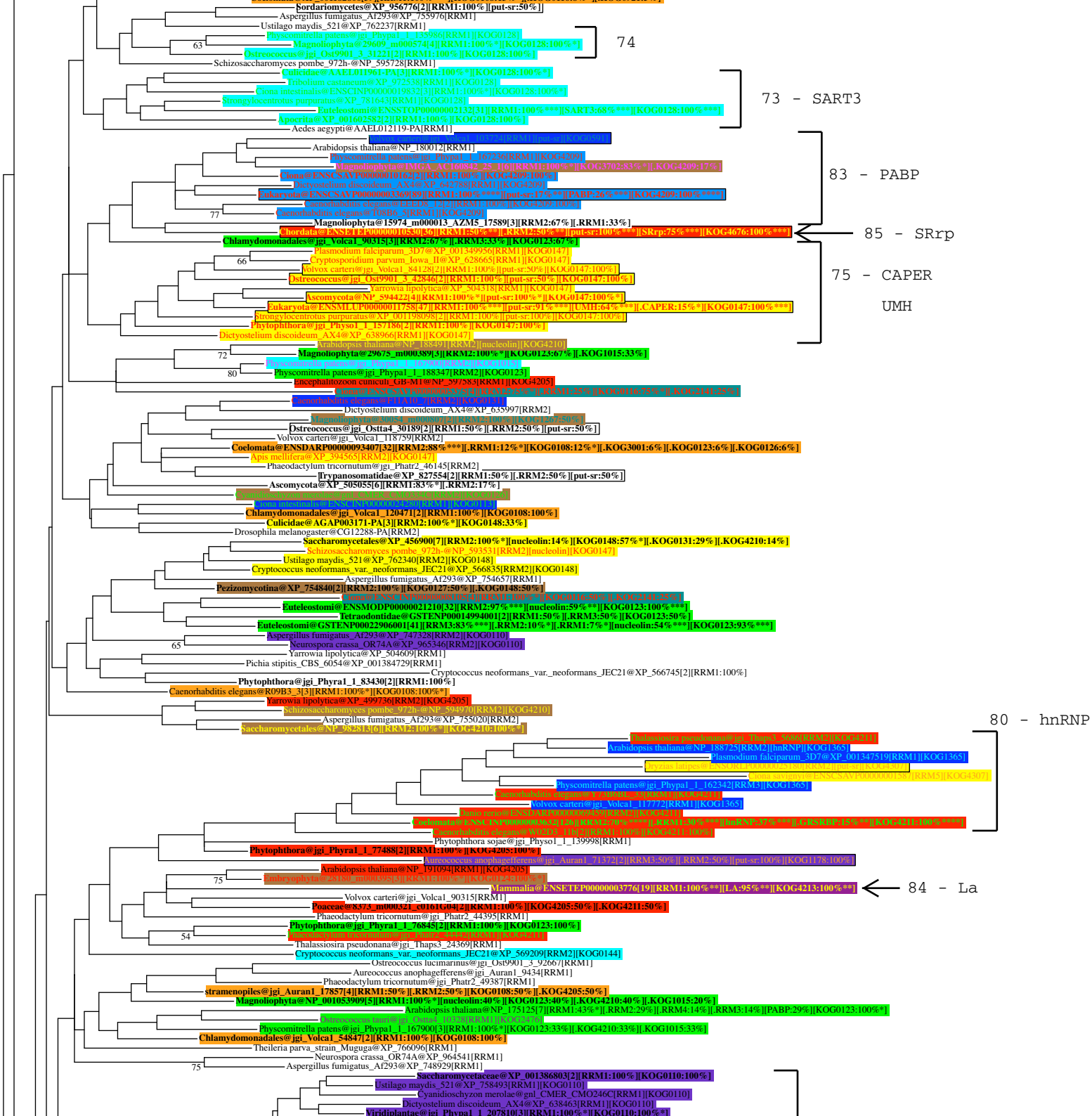


Fig. S8

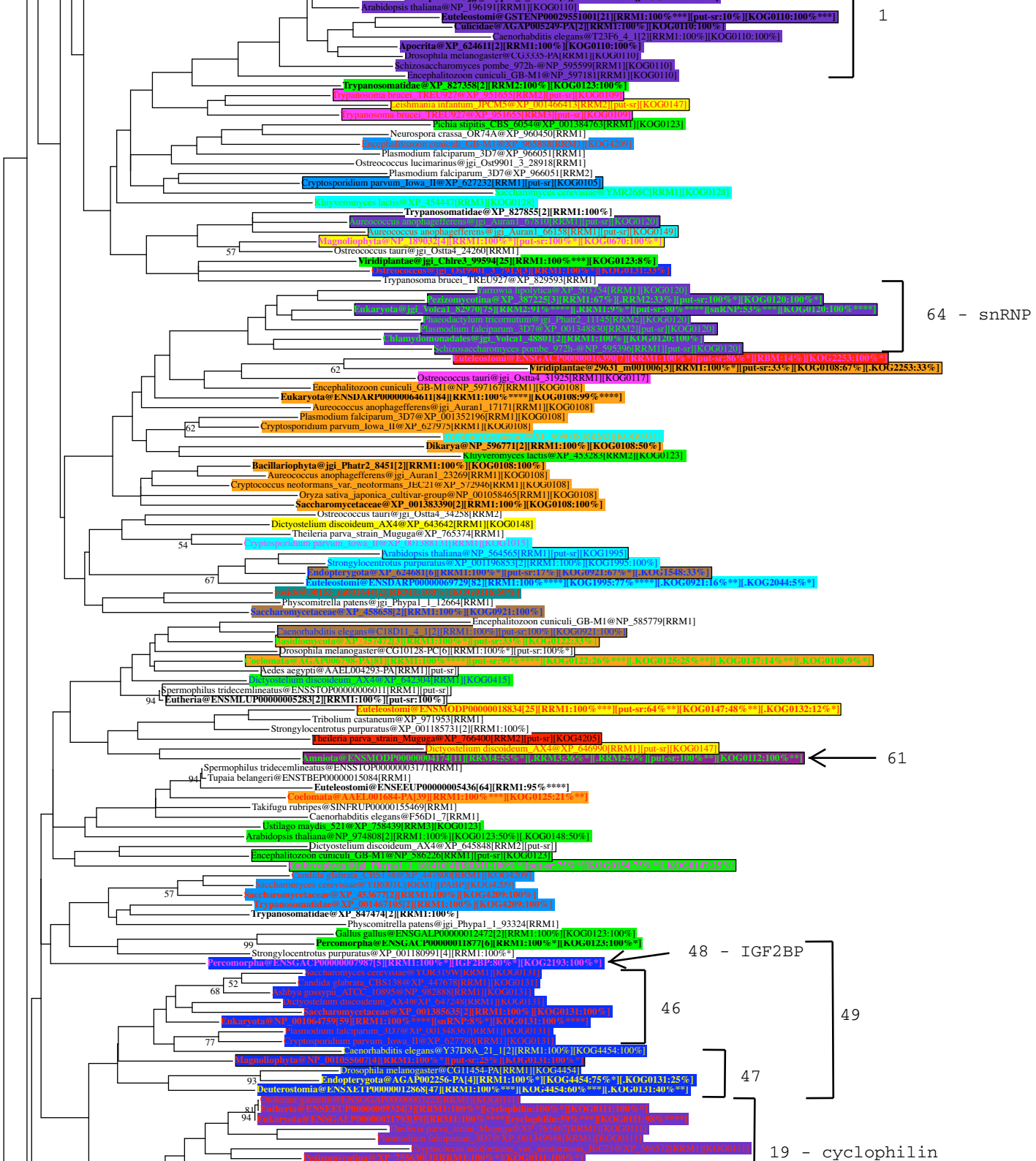
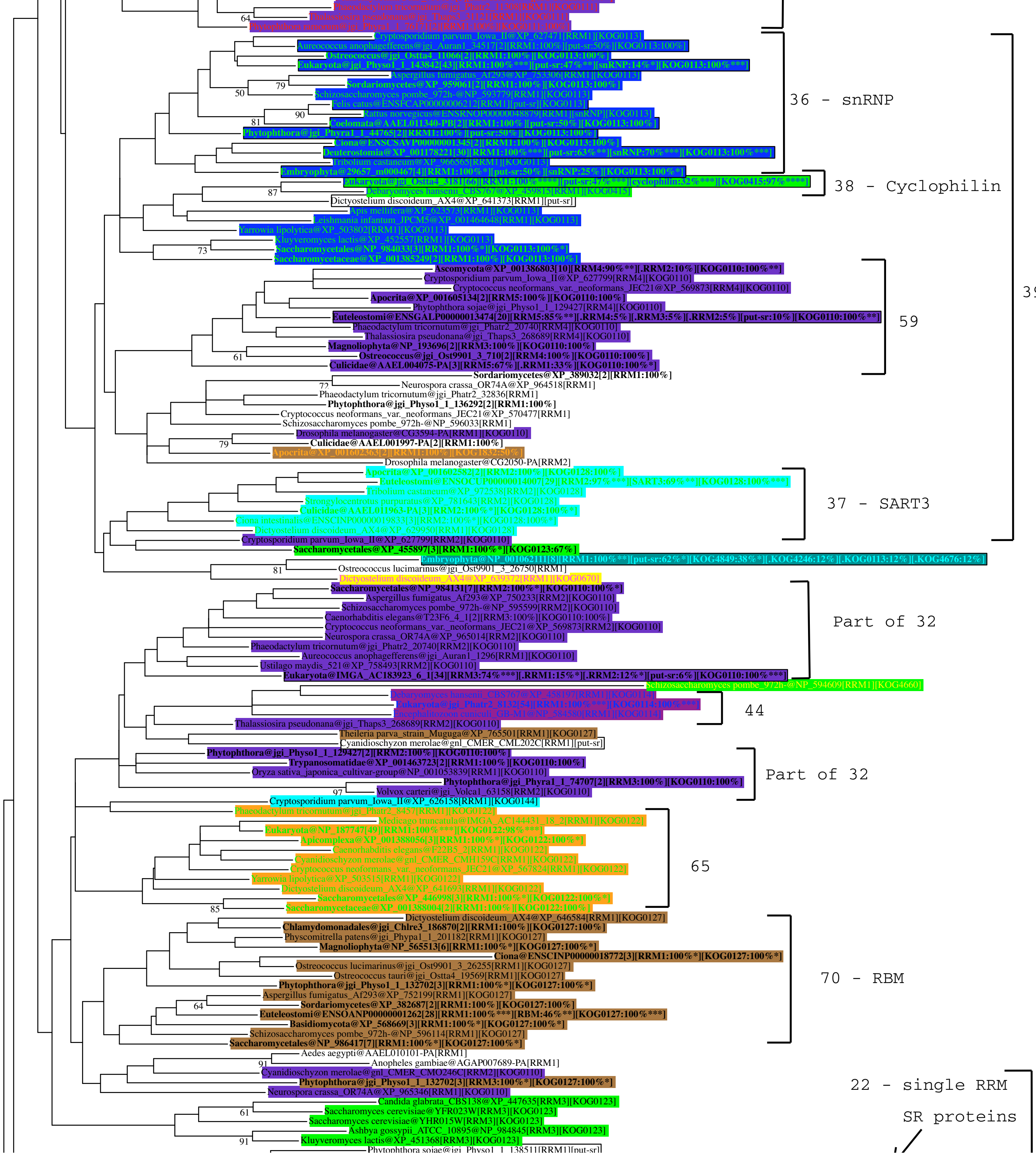


Fig. S8



36 - snRNP
38 - Cyclophilin
39
59
37 - SART3
Part of 32
44
Part of 32
65
70 - RBM
22 - single RRM
SR proteins

Fig. S8

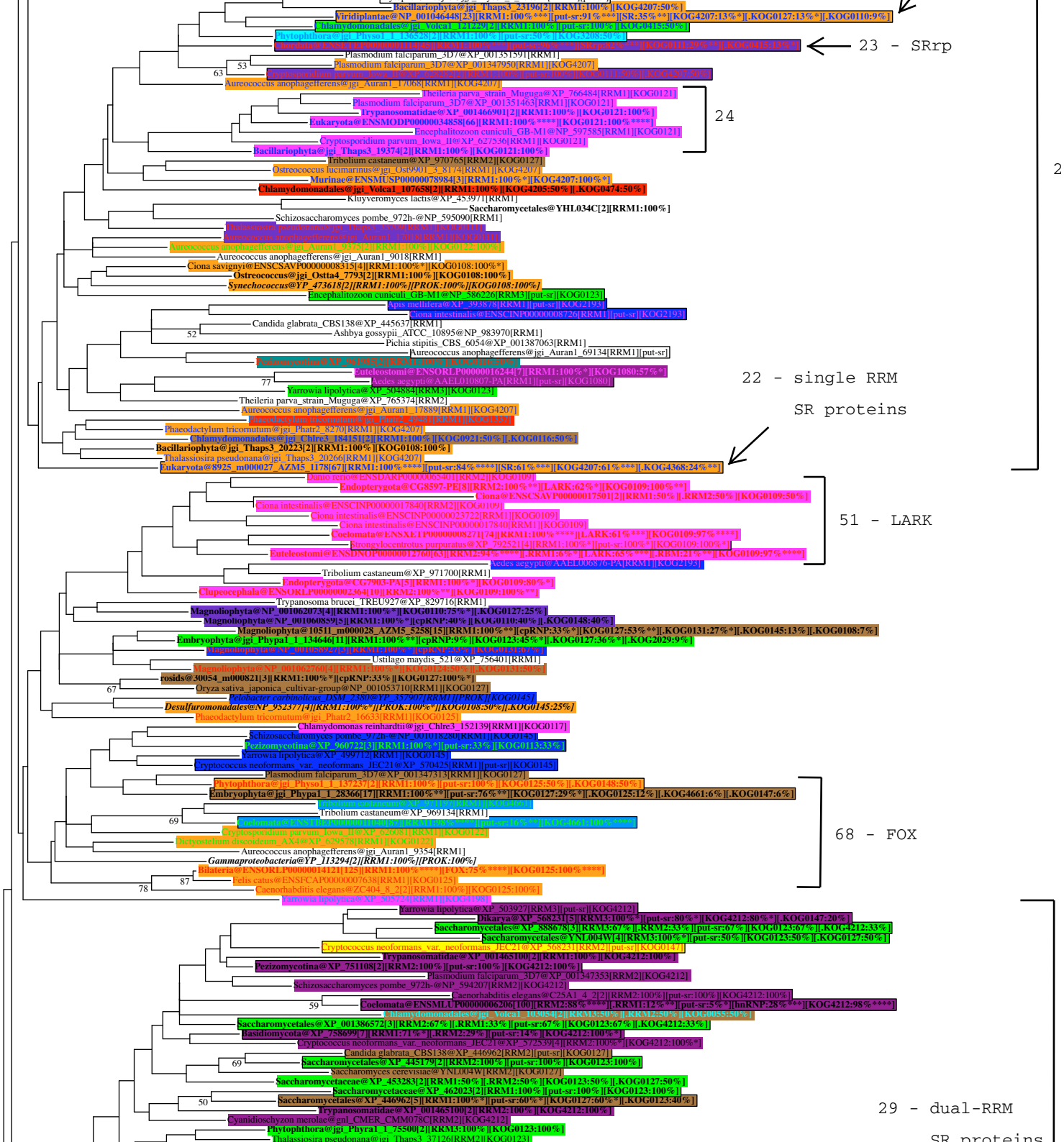

23 - SRrp
24
25
22 - single RRM
SR proteins
51 - LARK
68 - FOX
29 - dual-RRM
SR proteins
Part of 30
- hnRNP

Fig. S8

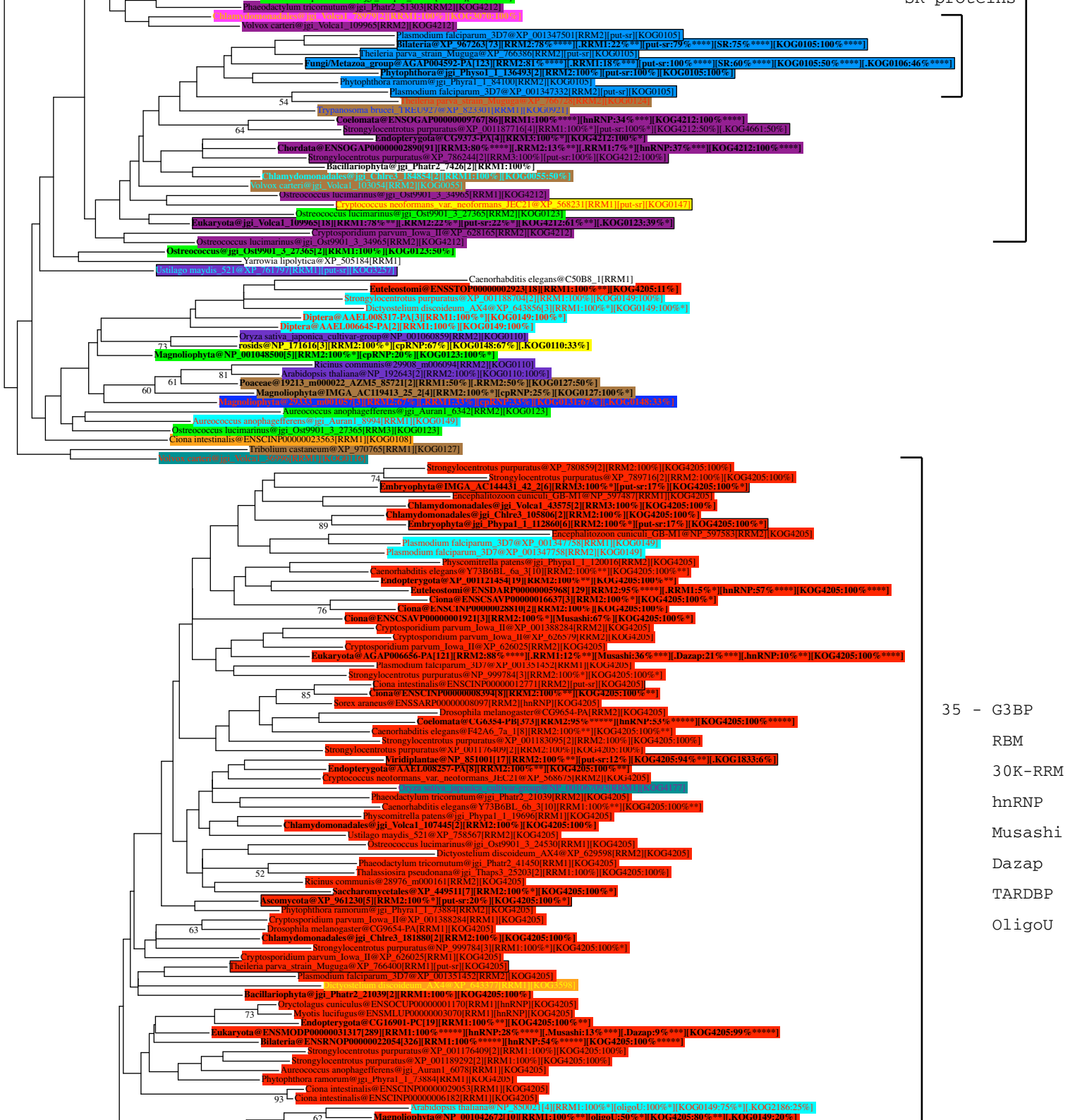

SR proteins
Phaeodactylum tricornutum@jgi_Phatr2_51303[RRM2][KOG4212]
Volvox carteri@jgi_Volca1_109965[RRM2][KOG4212]
Plasmodium falciparum_3D7@XP_001347501[RRM2][put-sr][KOG0105]
Bilateria@XP_967263[73][RRM2:78%****][.RRM1:22%**][put-sr:79%****][SR:75%****][KOG0105:100%****]
Theileria parva_strain_Muguga@XP_766386[RRM2][put-sr][KOG0105]
Fungi/Metazoa_group@AGAP004592-PA[123][RRM2:81%****][.RRM1:18%***][put-sr:100%****][SR:60%****][KOG0105:50%****][.KOG0106:46%****]
Phytophthora@jgi_Physo1_1_136493[2][RRM2:100%][put-sr:100%][KOG0105:100%]
Phytophthora ramorum@jgi_Phyra1_1_84100[RRM2][KOG0105]
Plasmodium falciparum_3D7@XP_001347332[RRM2][put-sr][KOG0105]
Theileria parva_strain_Muguga@XP_766728[RRM2][KOG0124]
Trypanosoma brucei_TREU927@XP_823301[RRM1][KOG0921]
Coelomata@ENSOGAP00000009767[86][RRM1:100%****][hnRNP:34%***][KOG4212:100%****]
Strongylocentrotus purpuratus@XP_001187716[4][RRM1:100%*][put-sr:100%*][KOG4212:50%][.KOG4661:50%]
Endopterygota@CG9373-PA[4][RRM3:100%**][KOG4212:100%**]
Chordata@ENSOGAP00000002890[91][RRM3:80%****][.RRM2:13%**][.RRM1:7%*][hnRNP:37%***][KOG4212:100%****]
Strongylocentrotus purpuratus@XP_786244[2][RRM3:100%][put-sr:100%][KOG4212:100%]
Bacillariophyta@jgi_Phatr2_7426[2][RRM1:100%]
Chlamydomonadales@jgi_Chlre3_174852[2][RRM3:100%][KOG0055:50%]
Volvox carteri@jgi_Volca1_103054[RRM2][KOG0055]
Ostreococcus lucimarinus@jgi_Ost9901_3_34965[RRM1][KOG4212]
Cryptococcus neoformans_var._neoformans_JEC21@XP_568231[RRM1][put-sr][KOG0147]
Ostreococcus lucimarinus@jgi_Ost9901_3_27365[RRM2][KOG0123]
Eukaryota@jgi_Volca1_109965[18][RRM1:78%**][.RRM2:22%*][put-sr:22%*][KOG4212:61%**][.KOG0123:39%*]
Cryptosporidium parvum_Iowa_II@XP_628165[RRM2][KOG4212]
Ostreococcus lucimarinus@jgi_Ost9901_3_34965[RRM2][KOG4212]
Ostreococcus@jgi_Ost9901_3_27365[2][RRM1:100%][KOG0123:50%]
Yarrowia lipolytica@XP_505184[RRM1]
Ustilago maydis_521@XP_761797[RRM1][put-sr][KOG3257]
Caenorhabditis elegans@C50B8_1[RRM1]
Euteleostomi@ENSSTOP00000002923[18][RRM1:100%**][KOG4205:11%]
Strongylocentrotus purpuratus@XP_001188704[2][RRM1:100%*][KOG0149:100%]
Dictyostelium discoideum_AX4@XP_643856[3][RRM1:100%*][KOG0149:100%*]
Diptera@AAEL008317-PA[3][RRM1:100%*][KOG0149:100%*]
Diptera@AAEL006645-PA[2][RRM1:100%][KOG0149:100%]
Oryza sativa_japonica_cultivar-group@NP_001060859[RRM2][KOG0110]
rosids@NP_171616[3][RRM2:100%*][cpRNP:67%][KOG0148:67%][.KOG0110:33%]
Magnoliophyta@NP_001048500[5][RRM2:100%*][cpRNP:20%][KOG0123:100%*]
Ricinus communis@29908_m000094[RRM2][KOG0110]
Arabidopsis thaliana@NP_192643[2][RRM2:100%][KOG0110:100%]
Poaceae@19213_m000022_AZM5_85721[2][RRM1:50%][.RRM2:50%][KOG0127:50%]
Magnoliophyta@IMGA_AC119413_25_2[4][RRM2:100%*][cpRNP:25%][KOG0127:100%*]
Aureococcus anophagefferens@jgi_Auran1_6342[RRM2][KOG0123]
Aureococcus anophagefferens@jgi_Auran1_8994[RRM1][KOG0149]
Ostreococcus lucimarinus@jgi_Ost9901_3_27365[RRM3][KOG0123]
Ciona intestinalis@ENSCINP00000023563[RRM1][KOG0108]
Tribolium castaneum@XP_970765[RRM1][KOG0127]
Strongylocentrotus purpuratus@XP_780859[2][RRM2:100%][KOG4205:100%]
Strongylocentrotus purpuratus@XP_789716[2][RRM2:100%][KOG4205:100%]
Embryophyta@IMGA_AC144431_42_2[6][RRM3:100%*][put-sr:17%][KOG4205:100%**]
Encephalitozoon cuniculi_GB-M1@NP_597487[RRM1][KOG4205]
Chlamydomonadales@jgi_Volca1_43575[2][RRM3:100%][KOG4205:100%]
Chlamydomonadales@jgi_Chlre3_105806[2][RRM2:100%][KOG4205:100%]
Embryophyta@jgi_Phypa1_1_112860[6][RRM2:100%*][put-sr:17%][KOG4205:100%**]
Encephalitozoon cuniculi_GB-M1@NP_597583[RRM2][KOG4205]
Plasmodium falciparum_3D7@XP_001347758[RRM1][KOG0149]
Plasmodium falciparum_3D7@XP_001347758[RRM2][KOG0149]
Physcomitrella patens@jgi_Phypa1_1_120016[RRM2][KOG4205]
Caenorhabditis elegans@Y73B6BL_6a_3[10][RRM2:100%**][KOG4205:100%**]
Endopterygota@XP_001121454[19][RRM2:100%**][KOG4205:100%**]
Euteleostomi@ENSDARP00000005968[129][RRM2:95%****][.RRM1:5%*][hnRNP:57%****][KOG4205:100%****]
Ciona@ENSCSAVP00000016637[3][RRM2:100%*][KOG4205:100%*]
Ciona@ENSCINP00000028810[2][RRM2:100%][KOG4205:100%]
Ciona@ENSCSAVP00000001921[3][RRM2:100%*][Musashi:67%][KOG4205:100%*]
Cryptosporidium parvum_Iowa_II@XP_001388284[RRM2][KOG4205]
Cryptosporidium parvum_Iowa_II@XP_626579[RRM2][KOG4205]
Cryptosporidium parvum_Iowa_II@XP_626025[RRM2][KOG4205]
Eukaryota@AGAP006656-PA[121][RRM2:88%****][.RRM1:12%**][Musashi:36%***][.Dazap:21%***][.hnRNP:10%**][KOG4205:100%****]
Plasmodium falciparum_3D7@XP_001351452[RRM1][KOG4205]
Strongylocentrotus purpuratus@NP_999784[3][RRM2:100%*][KOG4205:100%*]
Ciona intestinalis@ENSCINP00000012771[RRM2][put-sr][KOG4205]
Ciona@ENSCINP00000008394[8][RRM2:100%**][KOG4205:100%**]
Sorex araneus@ENSSARP00000008097[RRM2][hnRNP][KOG4205]
Drosophila melanogaster@CG9654-PA[RRM2][KOG4205]
Coelomata@CG6354-PB[373][RRM2:95%*****][hnRNP:53%*****][KOG4205:100%*****]
Caenorhabditis elegans@F42A6_7a_1[8][RRM2:100%**][KOG4205:100%**]
Strongylocentrotus purpuratus@XP_001183095[2][RRM2:100%][KOG4205:100%]
Strongylocentrotus purpuratus@XP_001176409[2][RRM2:100%][KOG4205:100%]
Viridiplantae@NP_851001[17][RRM2:100%**][put-sr:12%][KOG4205:94%**][.KOG1833:6%]
Endopterygota@AAEL008257-PA[8][RRM2:100%**][KOG4205:100%**]
Cryptococcus neoformans_var._neoformans_JEC21@XP_568675[RRM2][KOG4205]
Phaeodactylum tricornutum@jgi_Phatr2_21039[RRM2][KOG4205]
Caenorhabditis elegans@Y73B6BL_6b_3[10][RRM1:100%**][KOG4205:100%**]
Physcomitrella patens@jgi_Phypa1_1_19696[RRM1][KOG4205]
Chlamydomonadales@jgi_Volca1_107445[2][RRM2:100%][KOG4205:100%]
Ustilago maydis_521@XP_758567[RRM2][KOG4205]
Ostreococcus lucimarinus@jgi_Ost9901_3_24530[RRM1][KOG4205]
Dictyostelium discoideum_AX4@XP_629598[RRM2][KOG4205]
Phaeodactylum tricornutum@jgi_Phatr2_41450[RRM1][KOG4205]
Thalassiosira pseudonana@jgi_Thaps3_25203[2][RRM1:100%][KOG4205:100%]
Ricinus communis@28976_m000161[RRM2][KOG4205]
Saccharomycetales@XP_449511[7][RRM2:100%*][KOG4205:100%*]
Ascomycota@XP_961230[5][RRM2:100%*][put-sr:20%][KOG4205:100%*]
Phytophthora ramorum@jgi_Phyra1_1_73884[RRM2][KOG4205]
Cryptosporidium parvum_Iowa_II@XP_001388284[RRM1][KOG4205]
Drosophila melanogaster@CG9654-PA[RRM1][KOG4205]
Chlamydomonadales@jgi_Chlre3_181880[2][RRM2:100%][KOG4205:100%]
Strongylocentrotus purpuratus@NP_999784[3][RRM1:100%*][KOG4205:100%*]
Cryptosporidium parvum_Iowa_II@XP_626025[RRM1][KOG4205]
Theileria parva_strain_Muguga@XP_766400[RRM1][put-sr][KOG4205]
Plasmodium falciparum_3D7@XP_001351452[RRM2][KOG4205]
Dictyostelium discoideum_AX4@XP_643977[RRM1][KOG4598]
Bacillariophyta@jgi_Phatr2_21039[2][RRM1:100%][KOG4205:100%]
Oryctolagus cuniculus@ENSOCUP00000001170[RRM1][hnRNP][KOG4205]
Myotis lucifugus@ENSMLUP00000003070[RRM1][hnRNP][KOG4205]
Endopterygota@CG16901-PC[19][RRM1:100%**][KOG4205:100%**]
Eukaryota@ENSMODP00000031317[289][RRM1:100%*****][hnRNP:28%****][.Musashi:13%***][.Dazap:9%***][KOG4205:99%*****]
Bilateria@ENSRNOP00000022054[326][RRM1:100%*****][hnRNP:54%*****][KOG4205:100%*****]
Strongylocentrotus purpuratus@XP_001176409[2][RRM1:100%][KOG4205:100%]
Strongylocentrotus purpuratus@XP_001189292[2][RRM1:100%][KOG4205:100%]
Aureococcus anophagefferens@jgi_Auran1_6078[RRM1][KOG4205]
Phytophthora ramorum@jgi_Phyra1_1_73884[RRM1][KOG4205]
Ciona intestinalis@ENSCINP00000029053[RRM1][KOG4205]
Ciona intestinalis@ENSCINP00000006182[RRM1][KOG4205]
Arabidopsis thaliana@NP_850021[4][RRM1:100%*][oligoU:100%*][KOG0149:75%*][.KOG2186:25%]
Magnoliophyta@NP_001042672[10][RRM1:100%**][oligoU:50%*][KOG4205:80%**][.KOG0149:20%]
35 - G3BP
RBM
30K-RRM
hnRNP
Musashi
Dazap
TARDBP
OligoU

Fig. S8

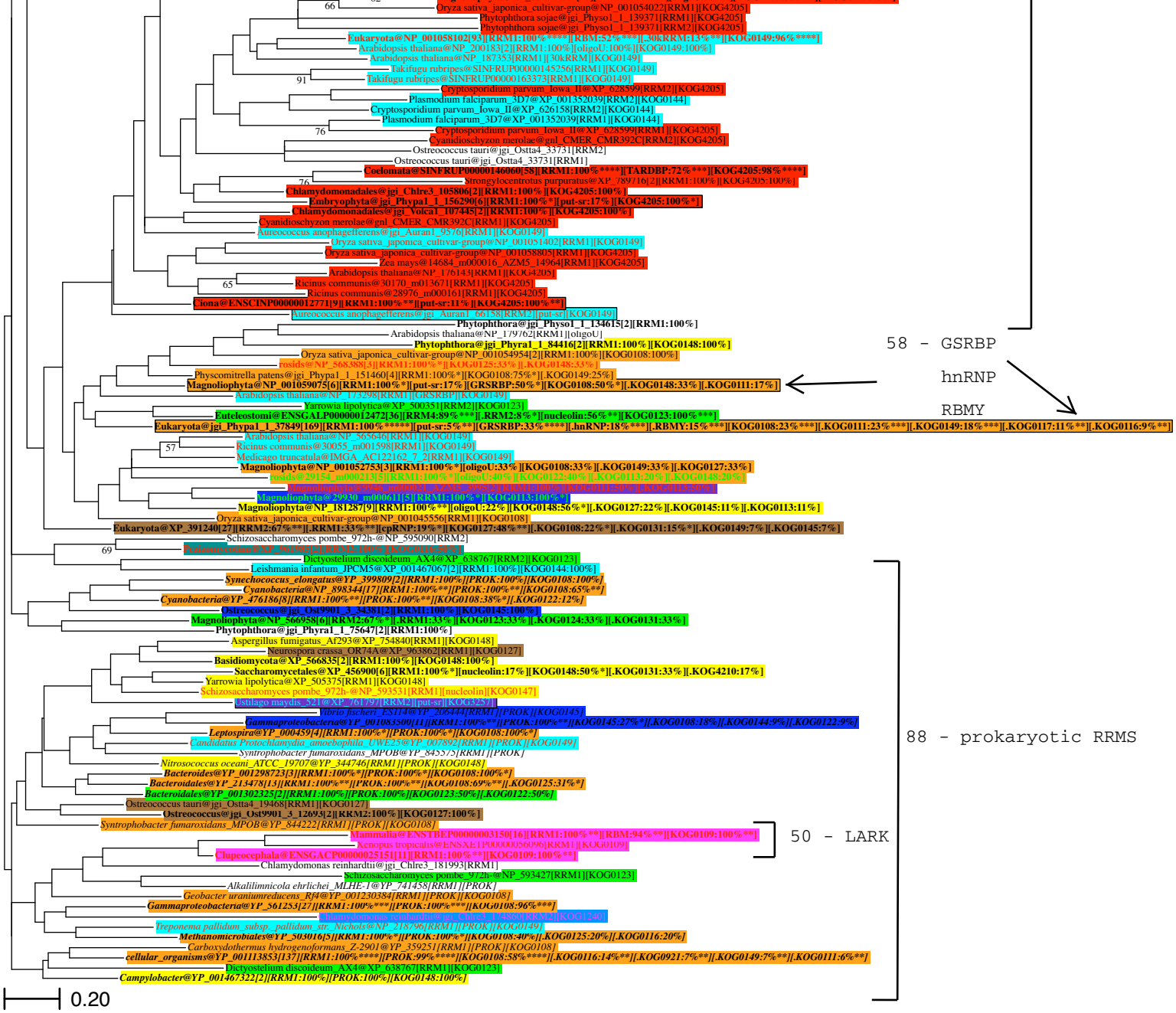


Fig. S9

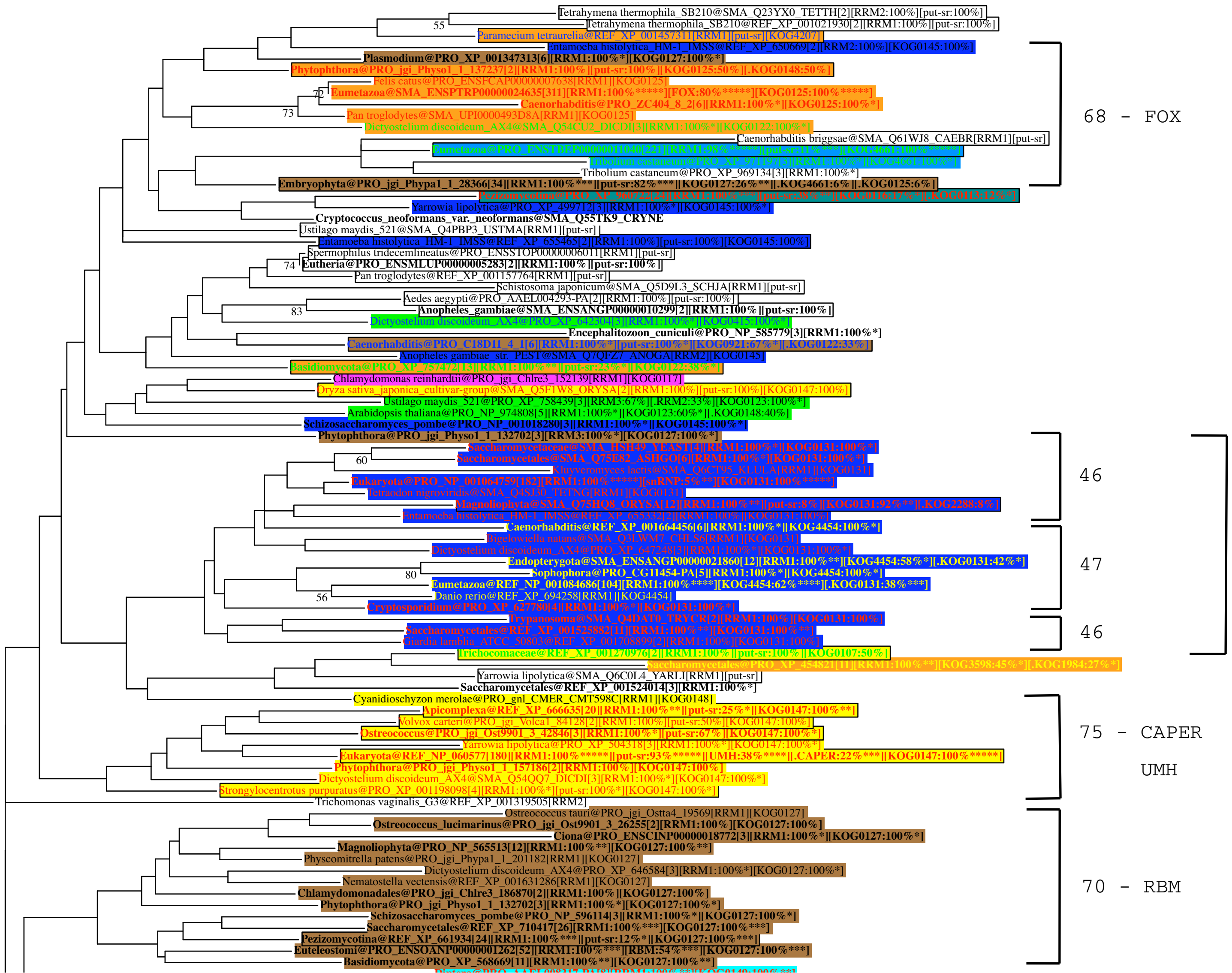
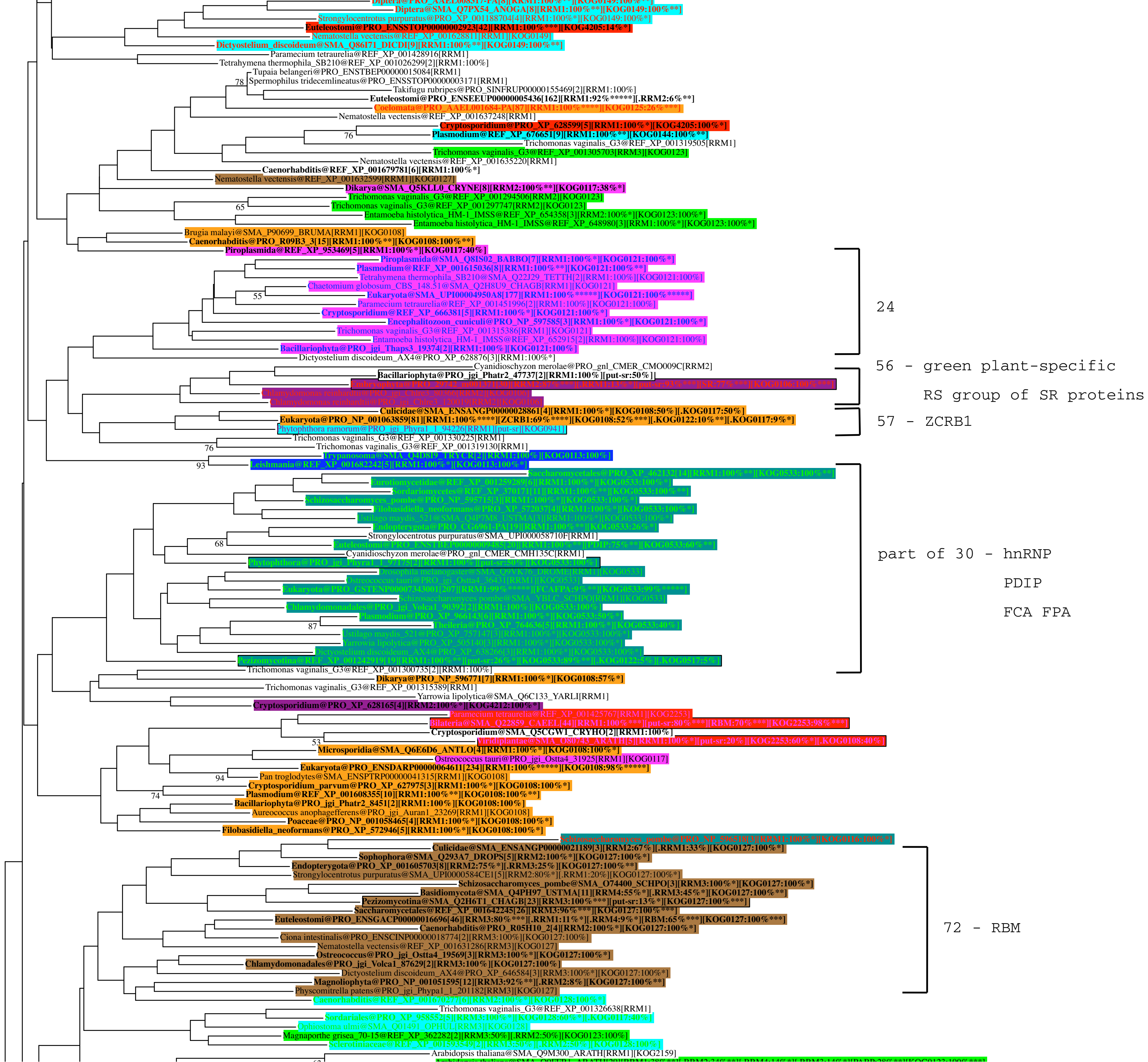


Fig. S9



24
56 - green plant-specific
RS group of SR proteins
57 - ZCRB1
part of 30 - hnRNP
PDIP
FCA FPA
72 - RBM

Fig. S9

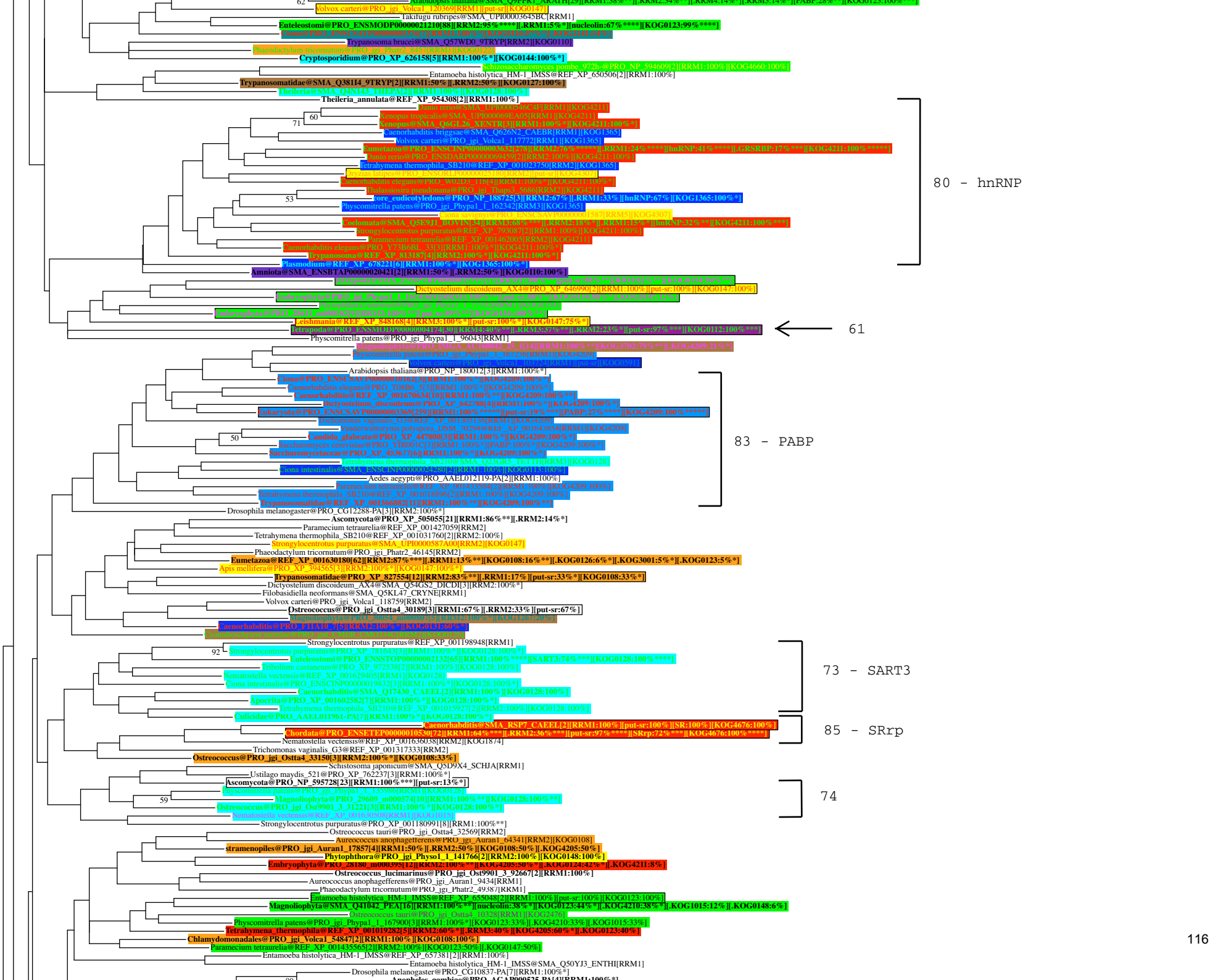


Fig. S9

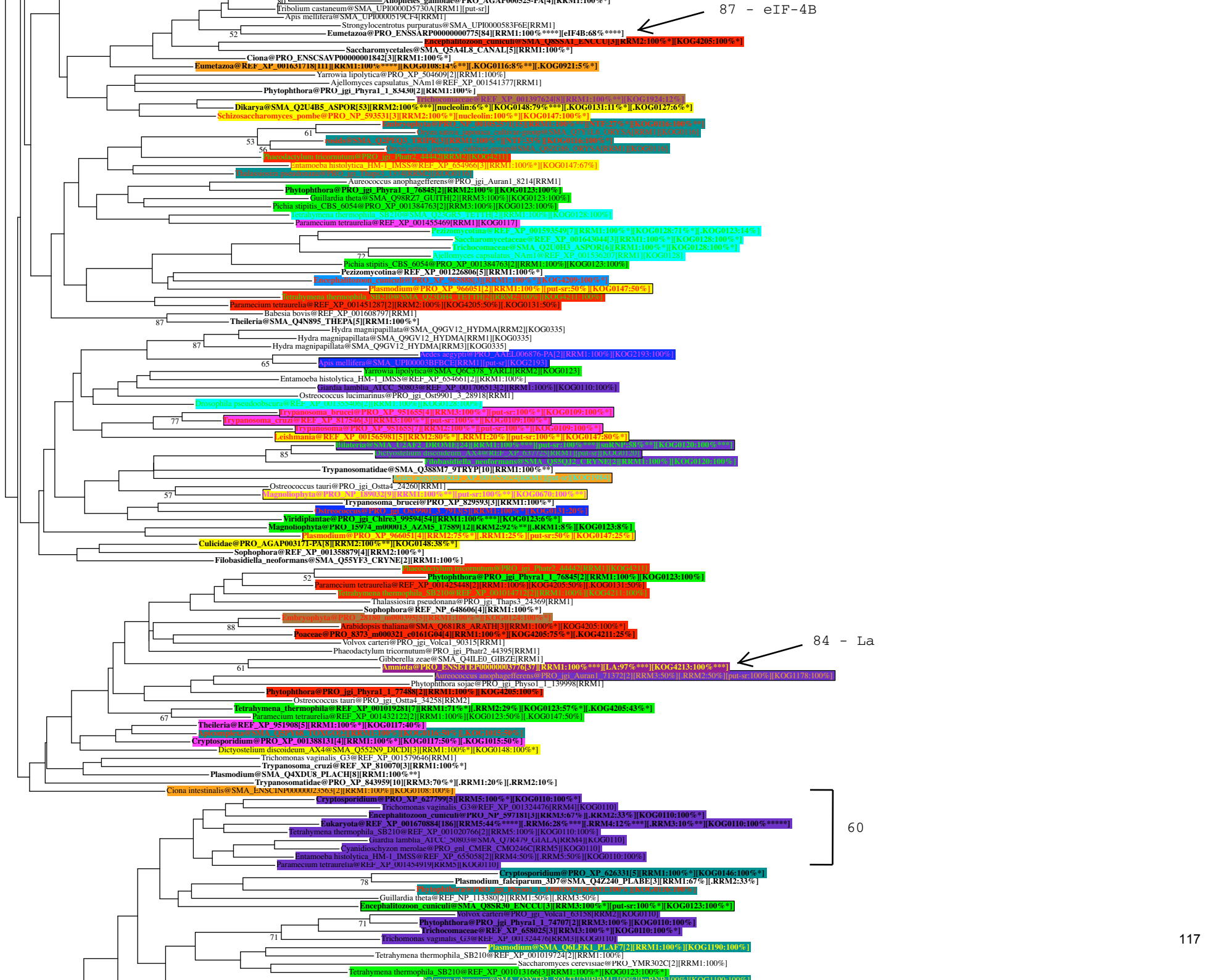


Fig. S9

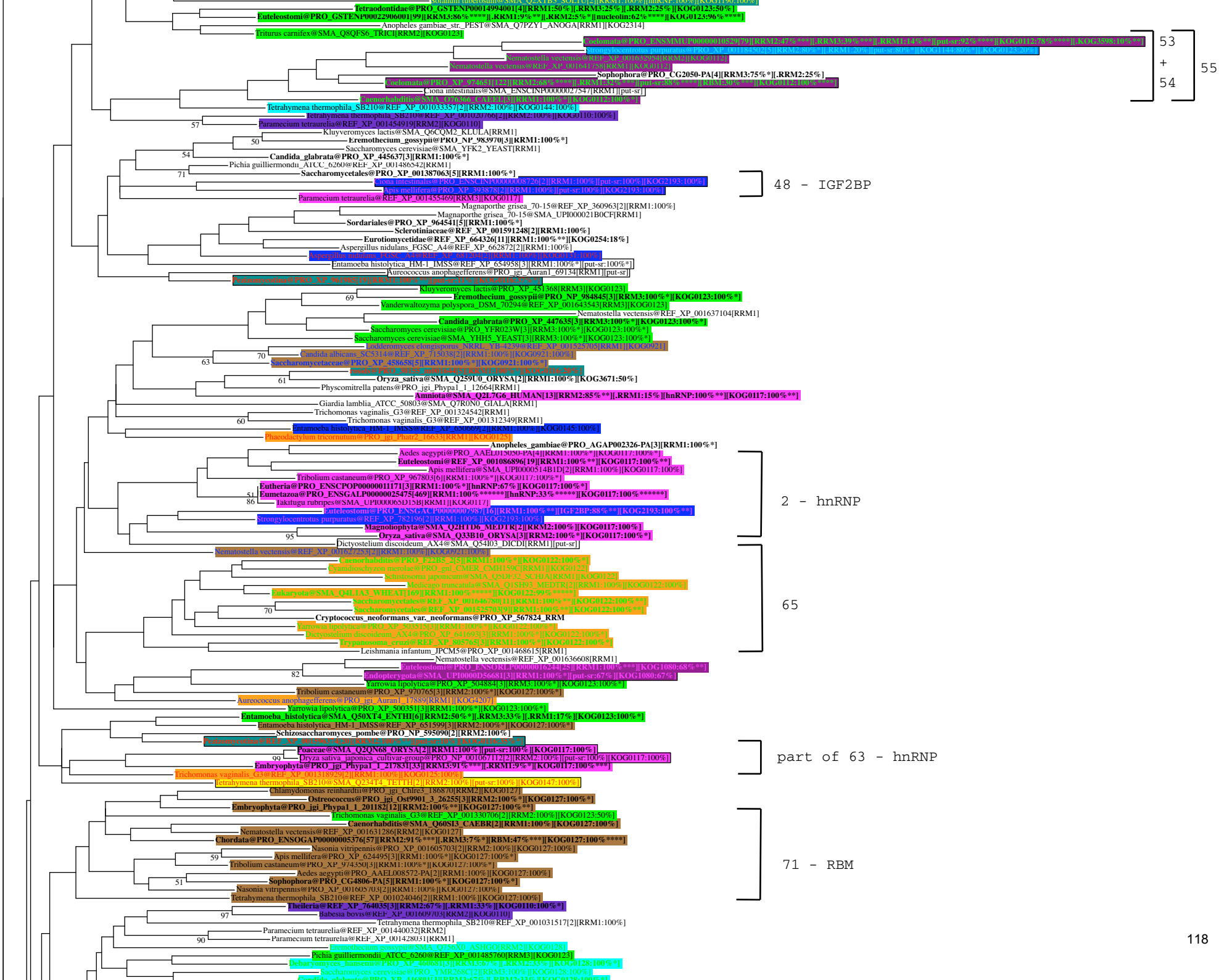

Tetraodontidae@PRO_GSTENP00014994001[4][RRM1:50%*][.RRM3:25%*][.RRM2:25%*][KOG0123:50%*]
Euteleostomi@PRO_GSTENP00022906001[99][RRM1:86%****][.RRM1:9%**][.RRM2:5%*][nucleolin:62%****][KOG0123:96%****]
Anopheles gambiae_str._PEST@SMA_Q7PZY1_ANOGA[RRM1][KOG2314]
Triturus carnifex@SMA_Q8QFS6_TRICI[RRM2][KOG0123]
Strongylocentrotus purpuratus@PRO_XP_001184502[5][RRM2:80%*][.RRM1:20%*][put-sr:80%*][KOG0144:80%*][KOG0123:20%*]
Nematostella vectensis@REF_XP_001641759[RRM1][KOG0124]
Sophophora@PRO_CG2050-PA[4][RRM3:75%*][.RRM2:25%*]
Ciona intestinalis@SMA_ENSCINP00000027547[RRM1][put-sr]
Caenorhabditis@SMA_Q76366_CAELE[3][RRM1:100%*][KOG0122:100%*]
Tetrahymena thermophila_SB210@REF_XP_001033357[2][RRM2:100%][KOG0144:100%]
Tetrahymena thermophila_SB210@REF_XP_001020766[2][RRM2:100%][KOG0110:100%]
Paramecium tetraurelia@REF_XP_001454919[RRM2][KOG0110]
Kluyveromyces lactis@SMA_Q6CQM2_KLULA[RRM1]
Eremothecium_gossypii@PRO_NP_983970[3][RRM1:100%*]
Saccharomyces cerevisiae@SMA_YFK2_YEAST[RRM1]
Candida_glabrata@PRO_XP_445637[3][RRM1:100%*]
Pichia guilliermondii_ATCC_6260@REF_XP_001486542[RRM1]
Saccharomycetales@PRO_XP_001387063[5][RRM1:100%*]
Ciona intestinalis@PRO_ENSCINP00000008726[2][RRM1:100%][put-sr:100%][KOG2193:100%]
Apis mellifera@PRO_XP_393878[2][RRM1:100%][put-sr:100%][KOG2193:100%]
Paramecium tetraurelia@REF_XP_001455469[RRM3][KOG0117]
Magnaporthe grisea_70-15@REF_XP_360963[2][RRM1:100%]
Magnaporthe grisea_70-15@SMA_UPI000021B0CF[RRM1]
Sordariales@PRO_XP_964541[5][RRM1:100%*]
Sclerotiniaceae@REF_XP_001591248[2][RRM1:100%]
Eurotiomycetidae@REF_XP_664326[11][RRM1:100%**][KOG0254:18%]
Aspergillus nidulans_FGSC_A4@REF_XP_662872[2][RRM1:100%]
Entamoeba histolytica_HM-1_IMSS@REF_XP_654958[3][RRM1:100%*][put-sr:100%*]
Aureococcus anophagefferens@PRO_jgi_Auran1_69134[RRM1][put-sr]
Kluyveromyces lactis@PRO_XP_451368[RRM3][KOG0123]
Eremothecium_gossypii@PRO_NP_984845[3][RRM3:100%*][KOG0123:100%*]
Vanderwaltozyma polyspora_DSM_70294@REF_XP_001643543[RRM3][KOG0123]
Nematostella vectensis@REF_XP_001637104[RRM1]
Candida_glabrata@PRO_XP_447635[3][RRM3:100%*][KOG0123:100%*]
Saccharomyces cerevisiae@PRO_YFR023W[3][RRM3:100%*][KOG0123:100%*]
Saccharomyces cerevisiae@SMA_YHH5_YEAST[3][RRM3:100%*][KOG0123:100%*]
Candida albicans_SC5314@REF_XP_715038[2][RRM1:100%][KOG0921:100%]
Saccharomycetaceae@PRO_XP_458658[5][RRM1:100%*][KOG0921:100%*]
Oryza_sativa@SMA_Q259U0_ORYSA[2][RRM1:100%][KOG3671:50%]
Physcomitrella patens@PRO_jgi_Phypa1_1_12664[RRM1]
Amniota@SMA_Q2L7G6_HUMAN[13][RRM2:85%**][.RRM1:15%][hnRNP:100%**][KOG0117:100%**]
Giardia lamblia_ATCC_50803@SMA_Q7R0N0_GIALA[RRM1]
Trichomonas vaginalis_G3@REF_XP_001324542[RRM1]
Trichomonas vaginalis_G3@REF_XP_001312349[RRM1]
Entamoeba histolytica_HM-1_IMSS@REF_XP_650069[2][RRM1:100%][KOG0145:100%]
Phaeodactylum tricornutum@PRO_jgi_Phatr2_16633[RRM1][KOG0125]
Anopheles_gambiae@PRO_AGAP002326-PA[3][RRM1:100%*]
Aedes aegypti@PRO_AAEL015050-PA[4][RRM1:100%*][KOG0117:100%*]
Euteleostomi@REF_XP_001086896[19][RRM1:100%**][KOG0117:100%**]
Apis mellifera@SMA_UPI0000514B1D[2][RRM1:100%][KOG0117:100%]
Tribolium castaneum@PRO_XP_967803[6][RRM1:100%*][KOG0117:100%*]
Eutheria@PRO_ENSCPOP00000011171[3][RRM1:100%*][hnRNP:67%][KOG0117:100%*]
Eumetazoa@PRO_ENSGALP00000025475[469][RRM1:100%*****][hnRNP:33%*****][KOG0117:100%******]
Takifugu rubripes@SMA_UPI000065D15B[RRM1][KOG0117]
Euteleostomi@PRO_ENSGACP00000007987[16][RRM1:100%**][IGF2BP:88%**][KOG2193:100%**]
Strongylocentrotus purpuratus@REF_XP_782196[2][RRM1:100%][KOG2193:100%]
Magnoliophyta@SMA_Q2H1D6_MEDTR[2][RRM2:100%][KOG0117:100%]
Oryza_sativa@SMA_Q33B10_ORYSA[3][RRM2:100%][KOG0117:100%*]
Dictyostelium discoideum_AX4@SMA_Q54I03_DICDI[RRM1][put-sr]
Nematostella vectensis@REF_XP_001627451[1][RRM1:100%][KOG0144:100%]
Caenorhabditis@PRO_P22B8_2[5][RRM1:100%*][KOG0122:100%*]
Cyanidioschyzon merolae@PRO_gnl_CMER_CMH159C[RRM1][KOG0122]
Schistosoma japonicum@SMA_Q5DF32_SCHJA[RRM1][KOG0122]
Medicago truncatula@SMA_Q1SH93_MEDTR[2][RRM1:100%][KOG0122:100%]
Eukaryota@SMA_Q4L1A3_WHEAT[169][RRM1:100%*****][KOG0122:99%*****]
Saccharomycetales@REF_XP_001646780[11][RRM1:100%**][KOG0122:100%**]
Saccharomycetales@REF_XP_001525703[9][RRM1:100%**][KOG0122:100%**]
Cryptococcus_neoformans_var._neoformans@PRO_XP_567824_RRM
Yarrowia lipolytica@PRO_XP_503515[3][RRM1:100%*][KOG0122:100%*]
Dictyostelium discoideum_AX4@PRO_XP_641693[3][RRM1:100%*][KOG0122:100%*]
Trypanosoma_cruzi@REF_XP_805765[3][RRM1:100%*][KOG0122:100%*]
Leishmania infantum_JPCM5@PRO_XP_001468615[RRM1]
Nematostella vectensis@REF_XP_001636608[RRM1]
Euteleostomi@PRO_ENSORLP00000016243[23][RRM1:100%***][KOG1080:68%**]
Endopterygota@SMA_UPI0000D56681[3][RRM1:100%*][put-sr:67%][KOG1080:67%]
Yarrowia lipolytica@PRO_XP_504884[3][RRM3:100%*][KOG0123:100%*]
Tribolium castaneum@PRO_XP_970765[3][RRM2:100%*][KOG0127:100%*]
Aureococcus anophagefferens@PRO_jgi_Auran1_17889[RRM1][KOG4207]
Yarrowia lipolytica@PRO_XP_500351[3][RRM1:100%*][KOG0123:100%*]
Entamoeba_histolytica@SMA_Q50XT4_ENTHI[6][RRM2:50%*][.RRM3:33%][.RRM1:17%][KOG0123:100%*]
Entamoeba histolytica_HM-1_IMSS@REF_XP_651599[3][RRM2:100%][KOG0127:100%*]
Schizosaccharomyces_pombe@PRO_NP_595090[2][RRM2:100%]
Poaceae@SMA_Q2QN68_ORYSA[2][RRM1:100%][put-sr:100%][KOG0117:100%]
Oryza sativa_japonica_cultivar-group@PRO_NP_001067112[2][RRM2:100%][put-sr:100%][KOG0117:100%]
Embryophyta@PRO_jgi_Phypa1_1_21783[33][RRM3:91%***][.RRM1:9%*][KOG0117:100%***]
Trichomonas vaginalis_G3@REF_XP_001318929[2][RRM1:100%][KOG0125:100%]
Tetrahymena thermophila_SB210@SMA_Q234T4_TETTH[2][RRM2:100%][put-sr:100%][KOG0147:100%]
Chlamydomonas reinhardtii@PRO_jgi_Chlre3_186870[RRM2][KOG0127]
Ostreococcus@PRO_jgi_Ost9901_3_26255[3][RRM2:100%*][KOG0127:100%*]
Embryophyta@PRO_jgi_Phypa1_1_201182[12][RRM2:100%**][KOG0127:100%**]
Trichomonas vaginalis_G3@REF_XP_001330706[2][RRM2:100%][KOG0123:50%]
Caenorhabditis@SMA_Q60SI3_CAEBR[2][RRM1:100%][KOG0127:100%]
Nematostella vectensis@REF_XP_001631286[RRM2][KOG0127]
Chordata@PRO_ENSOGAP00000005376[57][RRM2:91%***][.RRM3:7%*][RBM:47%***][KOG0127:100%****]
Nasonia vitripennis@PRO_XP_001605703[2][RRM2:100%][KOG0127:100%]
Apis mellifera@PRO_XP_624495[3][RRM1:100%*][KOG0127:100%*]
Tribolium castaneum@PRO_XP_974350[3][RRM1:100%*][KOG0127:100%*]
Aedes aegypti@PRO_AAEL008572-PA[2][RRM1:100%][KOG0127:100%]
Sophophora@PRO_CG4806-PA[5][RRM1:100%*][KOG0127:100%*]
Nasonia vitripennis@PRO_XP_001605703[2][RRM1:100%][KOG0127:100%]
Tetrahymena thermophila_SB210@REF_XP_001024046[2][RRM1:100%][KOG0127:100%]
Theileria@REF_XP_764035[3][RRM2:67%][.RRM1:33%][KOG0110:100%*]
Babesia bovis@REF_XP_001609703[RRM2][KOG0110]
Tetrahymena thermophila_SB210@REF_XP_001031517[2][RRM1:100%]
Paramecium tetraurelia@REF_XP_001440032[RRM2]
Paramecium tetraurelia@REF_XP_001428031[RRM1]
Eremothecium gossypii@SMA_Q756X0_ASHGO[RRM2][KOG0128]
Pichia guilliermondii_ATCC_6260@REF_XP_001485760[RRM3][KOG0123]
Debaryomyces_hansenii@PRO_XP_460681[3][RRM3:67%][RRM2:33%][KOG0128:100%*]
Saccharomyces cerevisiae@PRO_YMR268C[2][RRM3:100%][KOG0128:100%]
53 + 54
55
48 - IGF2BP
2 - hnRNP
65
part of 63 - hnRNP
71 - RBM

Fig. S9

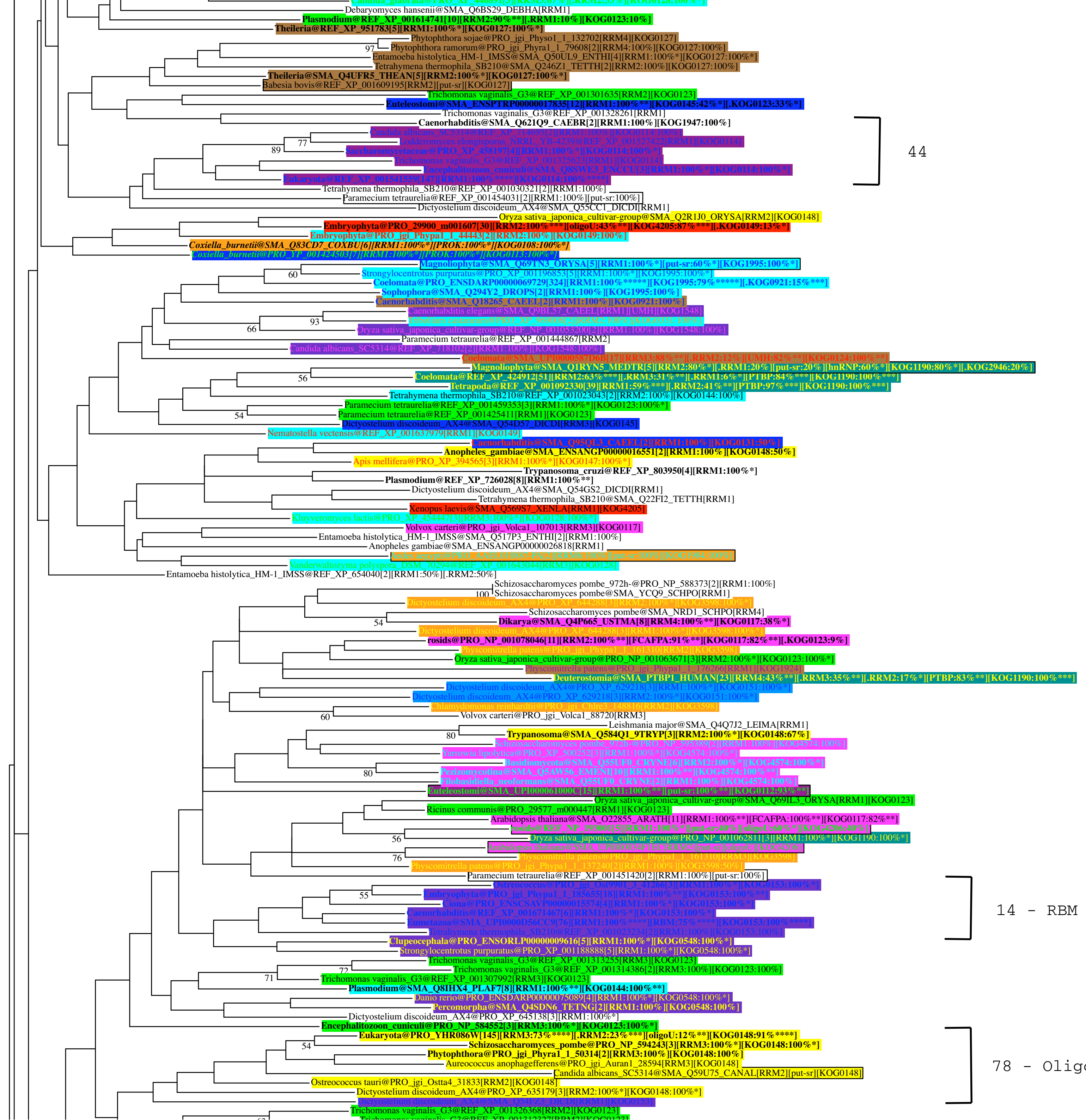


Fig. S9

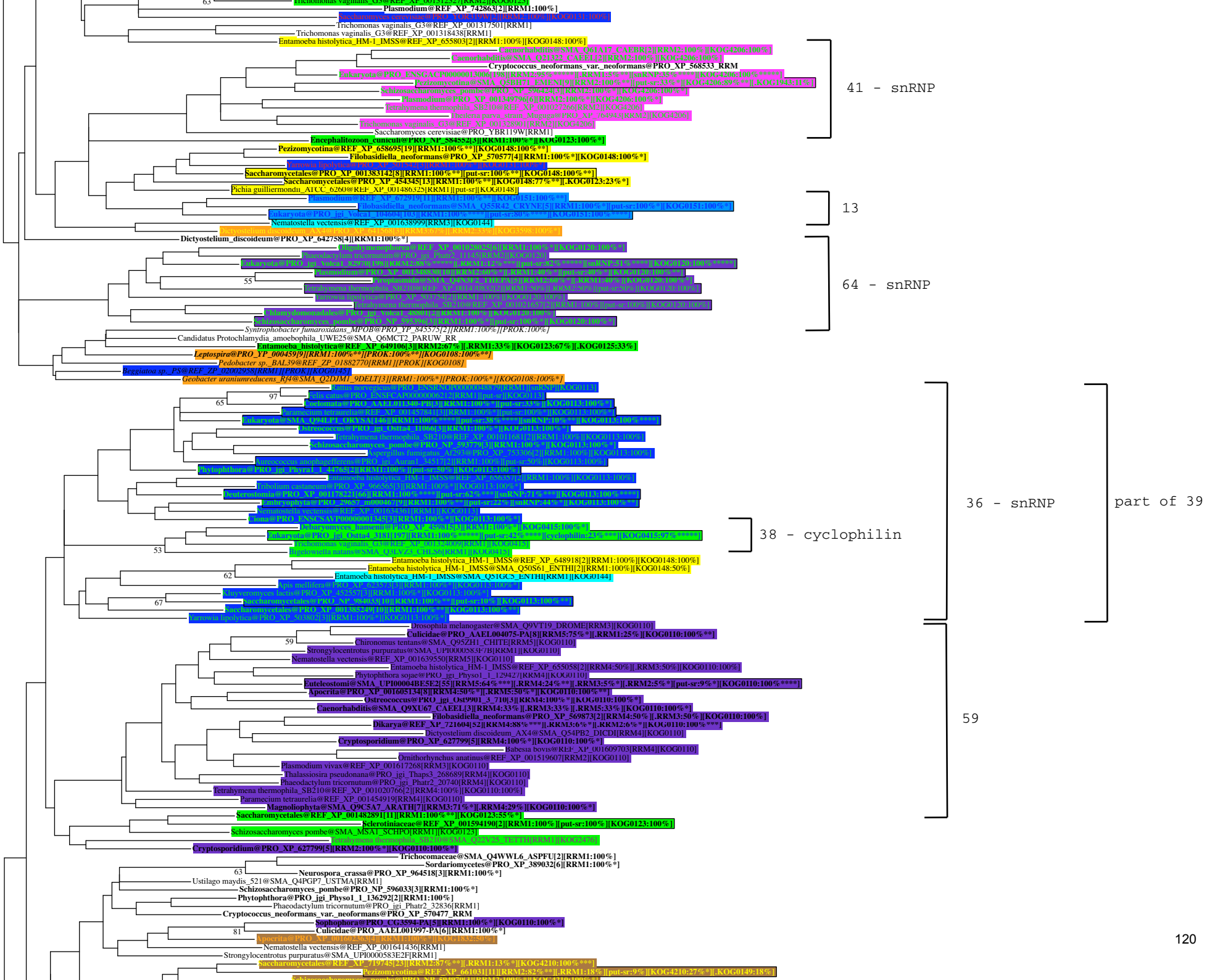

Trichomonas vaginalis_G3@REF_XP_001317501[RRM1]
Entamoeba histolytica_HM-1_IMSS@REF_XP_655803[2][RRM1:100%][KOG0148:100%]
Cryptococcus_neoformans_var._neoformans@PRO_XP_568533_RRM
Schizosaccharomyces_pombe@PRO_NP_596223[3][RRM2:100%*][KOG4206:100%*]
Plasmodium@REF_XP_001349796[6][RRM2:100%*][KOG4206:100%*]
Tetrahymena thermophila_SB210@REF_XP_001027206[RRM2][KOG4206]
Thalassiosira pseudonana@PRO_jgi_Thaps3_268689[RRM2][KOG4206]
Trichomonas vaginalis_G3@REF_XP_001328901[RRM2][KOG4206]
Saccharomyces cerevisiae@PRO_YBR119W[RRM1]
Encephalitozoon_cuniculi@PRO_NP_584552[3][RRM1:100%*][KOG0123:100%*]
Pezizomycotina@REF_XP_658695[19][RRM1:100%**][KOG0148:100%**]
Filobasidiella_neoformans@PRO_XP_570577[4][RRM1:100%*][KOG0148:100%*]
Saccharomycetales@PRO_XP_001388142[8][RRM1:100%**][put-sr:100%**][KOG0148:100%**]
Saccharomycetales@PRO_XP_454345[13][RRM1:100%**][KOG0148:77%**][.KOG0123:23%*]
Pichia guilliermondii_ATCC_6260@REF_XP_001486325[RRM1][put-sr][KOG0148]
Filobasidiella_neoformans@SMA_Q5SH42_CRYNE[5][RRM1:100%*][put-sr:100%*][KOG0151:100%*]
Eukaryota@PRO_jgi_Volca1_104604[103][RRM1:100%****][put-sr:80%****][KOG0151:100%****]
Nematostella vectensis@REF_XP_001638999[RRM3][KOG0144]
Dictyostelium_discoideum@PRO_XP_642758[4][RRM1:100%*]
Eukaryota@PRO_jgi_Volca1_82970[198][RRM2:98%*****][.RRM1:12%**][put-sr:82%*****][snRNP:51%****][KOG0120:100%*****]
Tetrahymena thermophila_SB210@REF_XP_001014683[3][RRM1:50%][RRM2:50%][put-sr:50%][KOG0120:100%]
Cryptosporidium@PRO_XP_627799[5][RRM2:100%*][KOG0110:100%*]
Trichomonas vaginalis_G3@REF_XP_001318438[RRM1]
Entamoeba histolytica_HM-1_IMSS@REF_XP_649106[3][RRM2:67%][.RRM1:33%][KOG0123:67%][.KOG0125:33%]
Leptospira@PRO_YP_000459[9][RRM1:100%**][PROK:100%**][KOG0108:100%**]
Pedobacter sp._BAL39@REF_ZP_01882770[RRM1][PROK][KOG0108]
Beggiatoa sp._PS@REF_ZP_02009858[RRM1][PROK][KOG0108]
Geobacter uraniumreducens_Rf4@SMA_Q2DM1_9DELT[3][RRM1:100%*][PROK:100%*][KOG0108:100%*]
Eukaryota@PRO_ENSCAFP00000006212[RRM1][put-sr][KOG0113]
Coelomata@PRO_AAEL001340-PH[3][RRM1:100%*][put-sr:33%*][KOG0113:100%*]
Caenorhabditis@SMA_Q9XU67_CAEEL[3][RRM4:33%][.RRM3:33%][.RRM5:33%][KOG0110:100%*]
Filobasidiella_neoformans@PRO_XP_569873[2][RRM4:50%][.RRM3:50%][KOG0110:100%]
Dikarya@REF_XP_721604[52][RRM4:88%***][.RRM3:6%*][.RRM2:6%*][KOG0110:100%***]
Dictyostelium discoideum_AX4@SMA_Q54PB2_DICDI[RRM4][KOG0110]
Cryptosporidium@PRO_XP_627799[5][RRM4:100%*][KOG0110:100%*]
Babesia bovis@REF_XP_001609703[RRM4][KOG0110]
Ornithorhynchus anatinus@REF_XP_001519607[RRM2][KOG0110]
Plasmodium vivax@REF_XP_001617268[RRM3][KOG0110]
Thalassiosira pseudonana@PRO_jgi_Thaps3_268689[RRM4][KOG0110]
Phaeodactylum tricornutum@PRO_jgi_Phatr2_20740[RRM4][KOG0110]
Tetrahymena thermophila_SB210@REF_XP_001020766[2][RRM4:100%][KOG0110:100%]
Paramecium tetraurelia@REF_XP_001454919[RRM4][KOG0110]
Magnoliophyta@SMA_Q9C5A7_ARATH[7][RRM3:71%*][.RRM4:29%][KOG0110:100%*]
Saccharomycetales@REF_XP_001482891[11][RRM1:100%**][KOG0123:55%*]
Sclerotiniaceae@REF_XP_001594190[2][RRM1:100%][put-sr:100%][KOG0123:100%*]
Schizosaccharomyces pombe@SMA_MSA1_SCHPO[RRM1][KOG0123]
Cryptosporidium@PRO_XP_627799[5][RRM2:100%*][KOG0110:100%*]
Trichocomaceae@SMA_Q4WWL6_ASPFU[2][RRM1:100%]
Sordariomycetes@PRO_XP_389032[6][RRM1:100%*]
Neurospora_crassa@PRO_XP_964518[3][RRM1:100%*]
Ustilago maydis_521@SMA_Q4PGP7_USTMA[RRM1]
Schizosaccharomyces_pombe@PRO_596033[3][RRM1:100%*]
Phytophthora@PRO_jgi_Physo1_1_136292[2][RRM1:100%]
Phaeodactylum tricornutum@PRO_jgi_Phatr2_32836[RRM1]
Cryptococcus_neoformans_var._neoformans@PRO_XP_570477_RRM
Sophophora@PRO_CG3594-PA[5][RRM1:100%*][KOG0110:100%*]
Culicidae@PRO_AAEL001997-PA[6][RRM1:100%*]
Nematostella vectensis@REF_XP_001641436[RRM1]
Strongylocentrotus purpuratus@SMA_UPI0000583E2F[RRM1]
Saccharomycetales@REF_XP_719745[23][RRM2:87%**][.RRM1:13%*][KOG4210:100%***]
Pezizomycotina@REF_XP_661031[11][RRM2:82%**][.RRM1:18%][put-sr:9%][KOG4210:27%*][.KOG0149:18%]
41 - snRNP
13
64 - snRNP
36 - snRNP
part of 39
38 - cyclophilin
59
120

Fig. S9

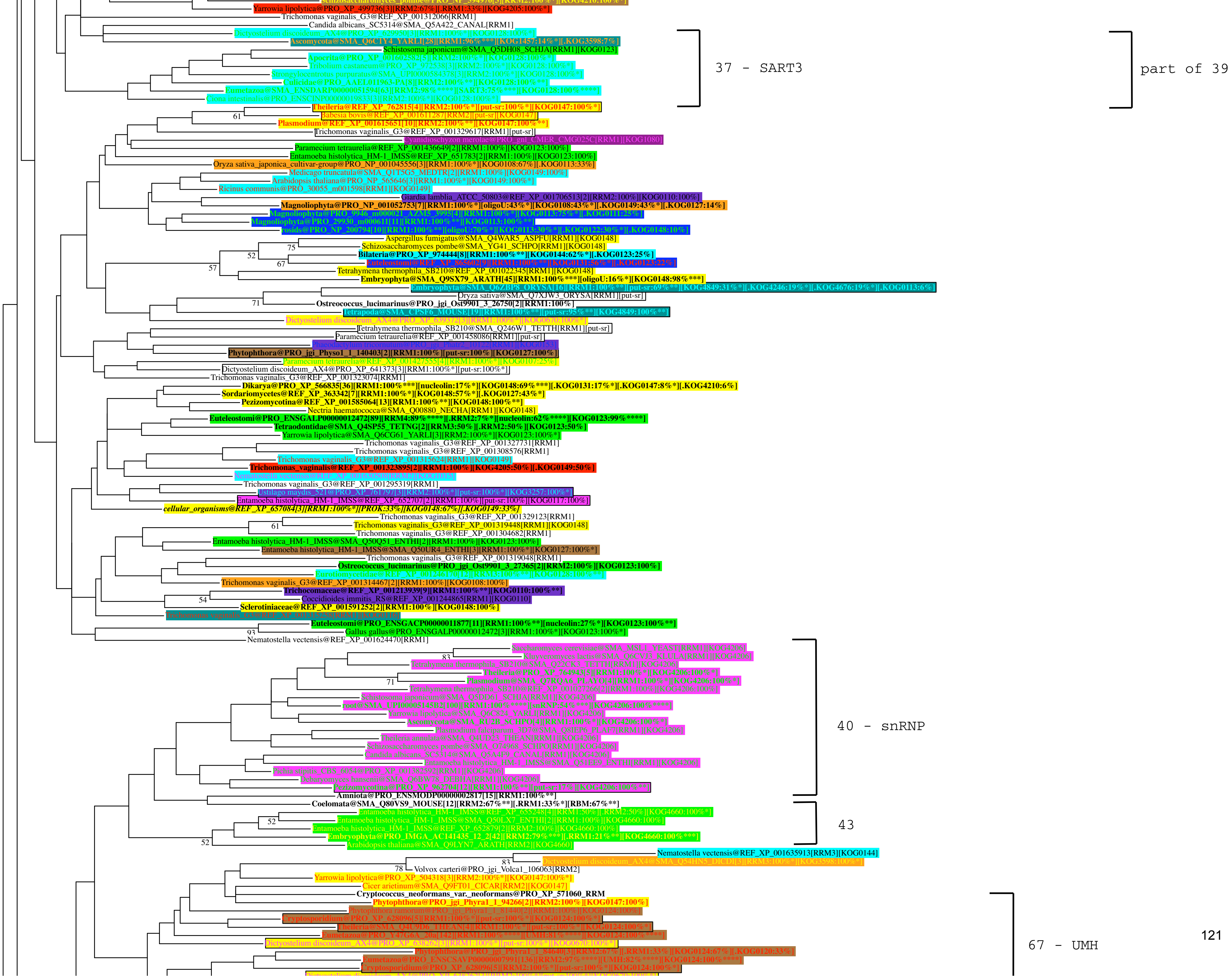


Fig. S9

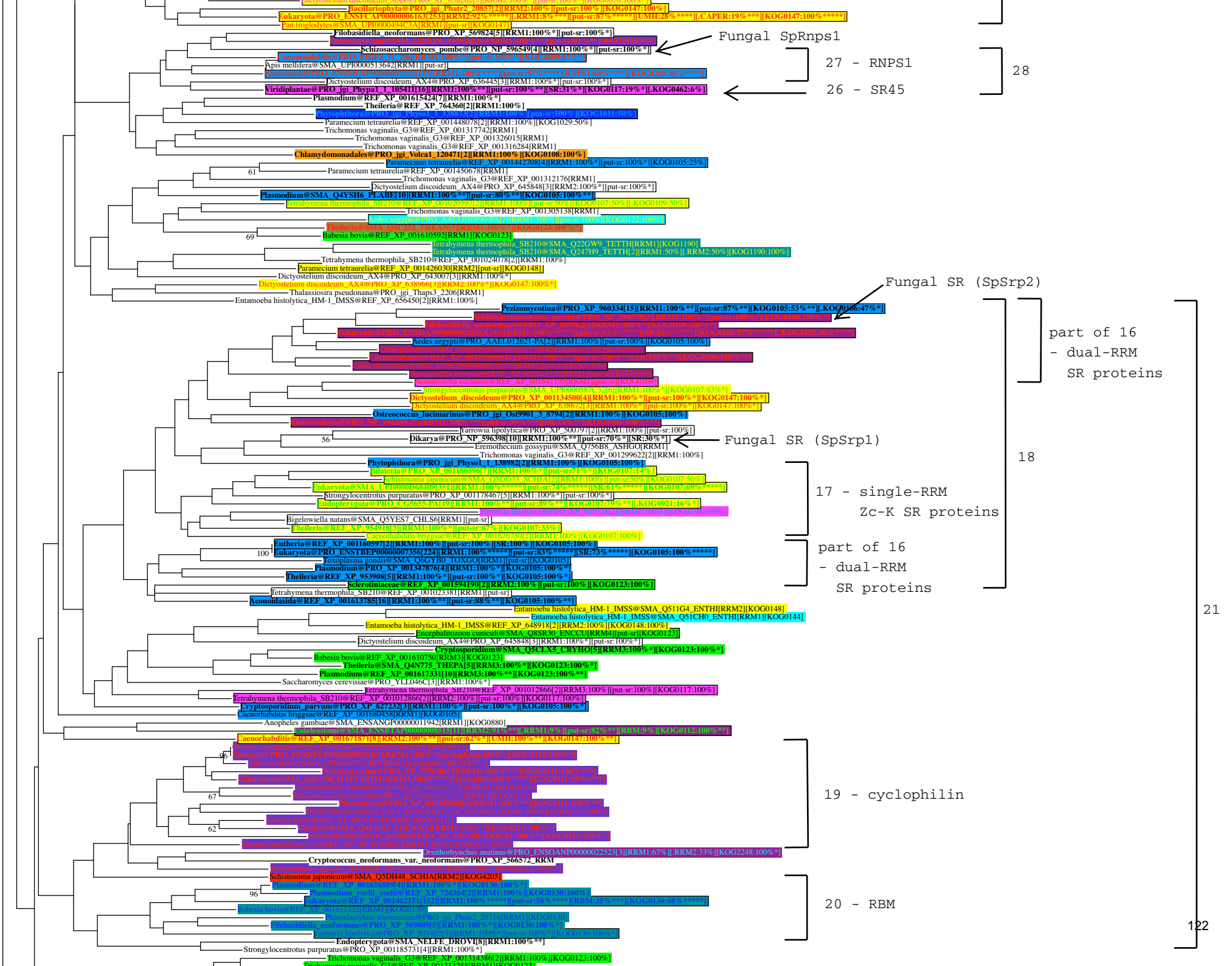

Fungal SpRnps1
27 - RNPS1
26 - SR45
28
Fungal SR (SpSrp2)
part of 16
- dual-RRM
SR proteins
Fungal SR (SpSrp1)
18
17 - single-RRM
Zc-K SR proteins
part of 16
- dual-RRM
SR proteins
21
19 - cyclophilin
20 - RBM
122

Fig. S9

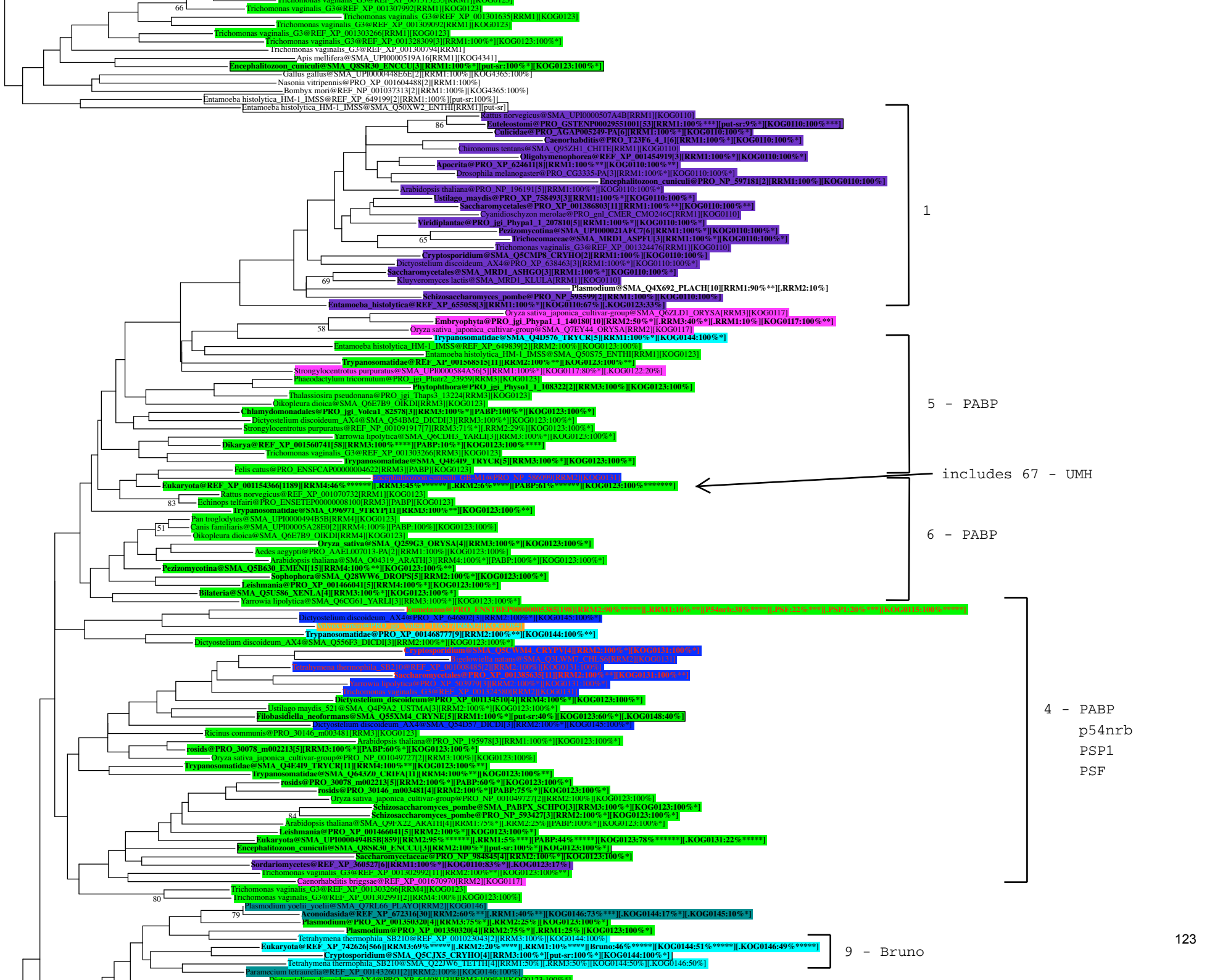
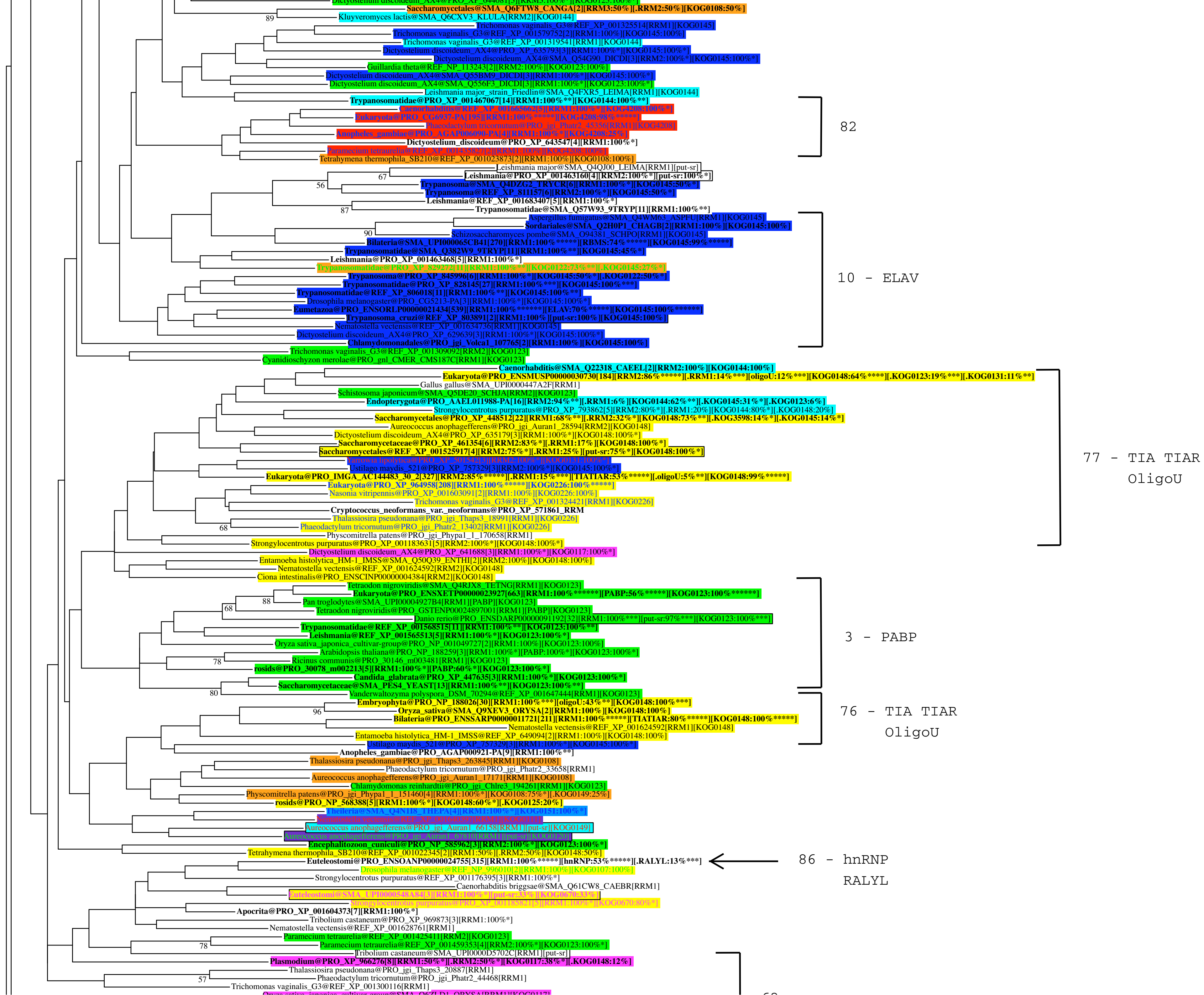
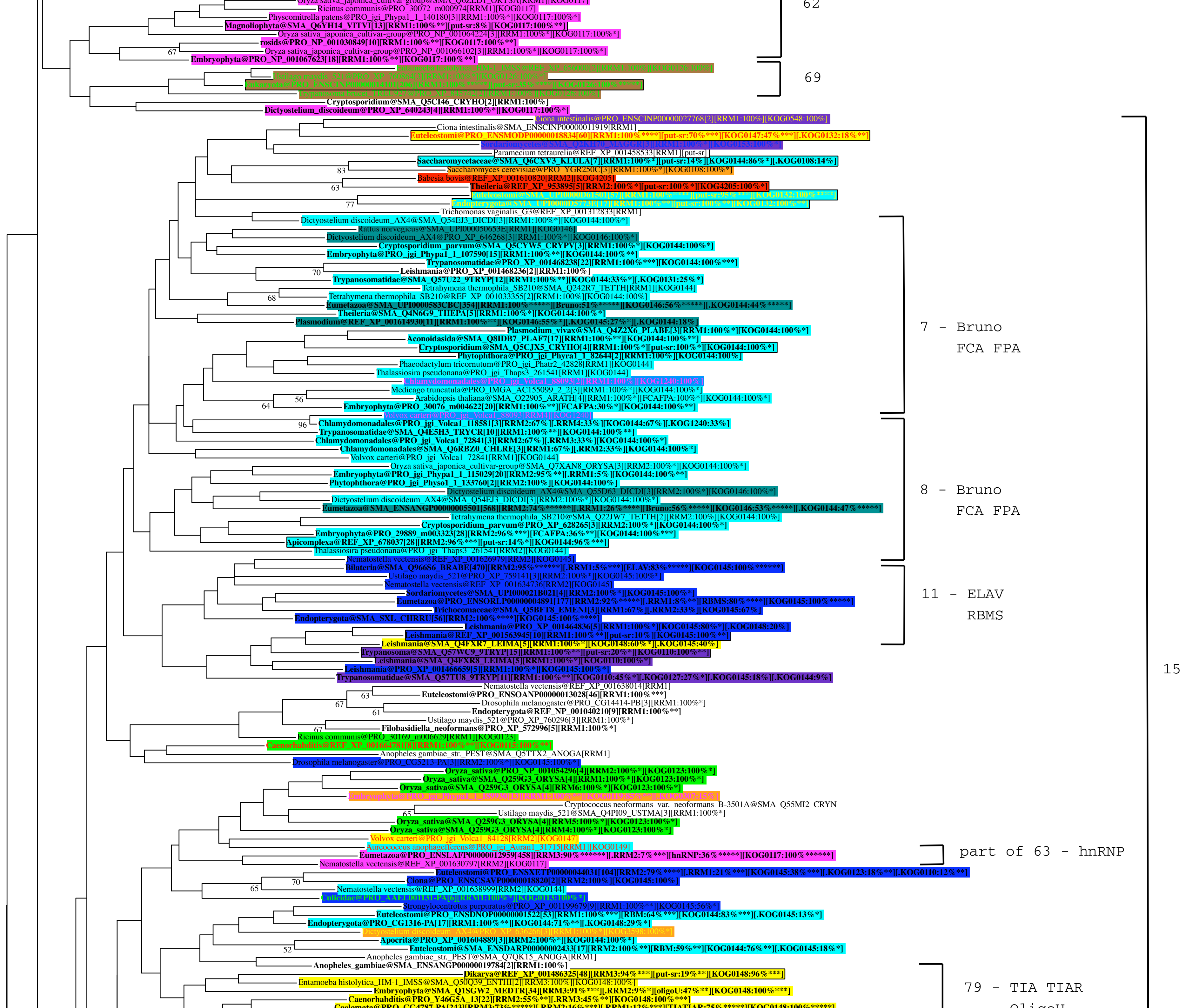


Fig. S9



82
10 - ELAV
77 - TIA TIAR
OligoU
3 - PABP
76 - TIA TIAR
OligoU
86 - hnRNP
RALYL

Fig. S9



62
69
7 - Bruno
FCA FPA
8 - Bruno
FCA FPA
11 - ELAV
RBMS
15
part of 63 - hnRNP
79 - TIA TIAR

Fig. S9

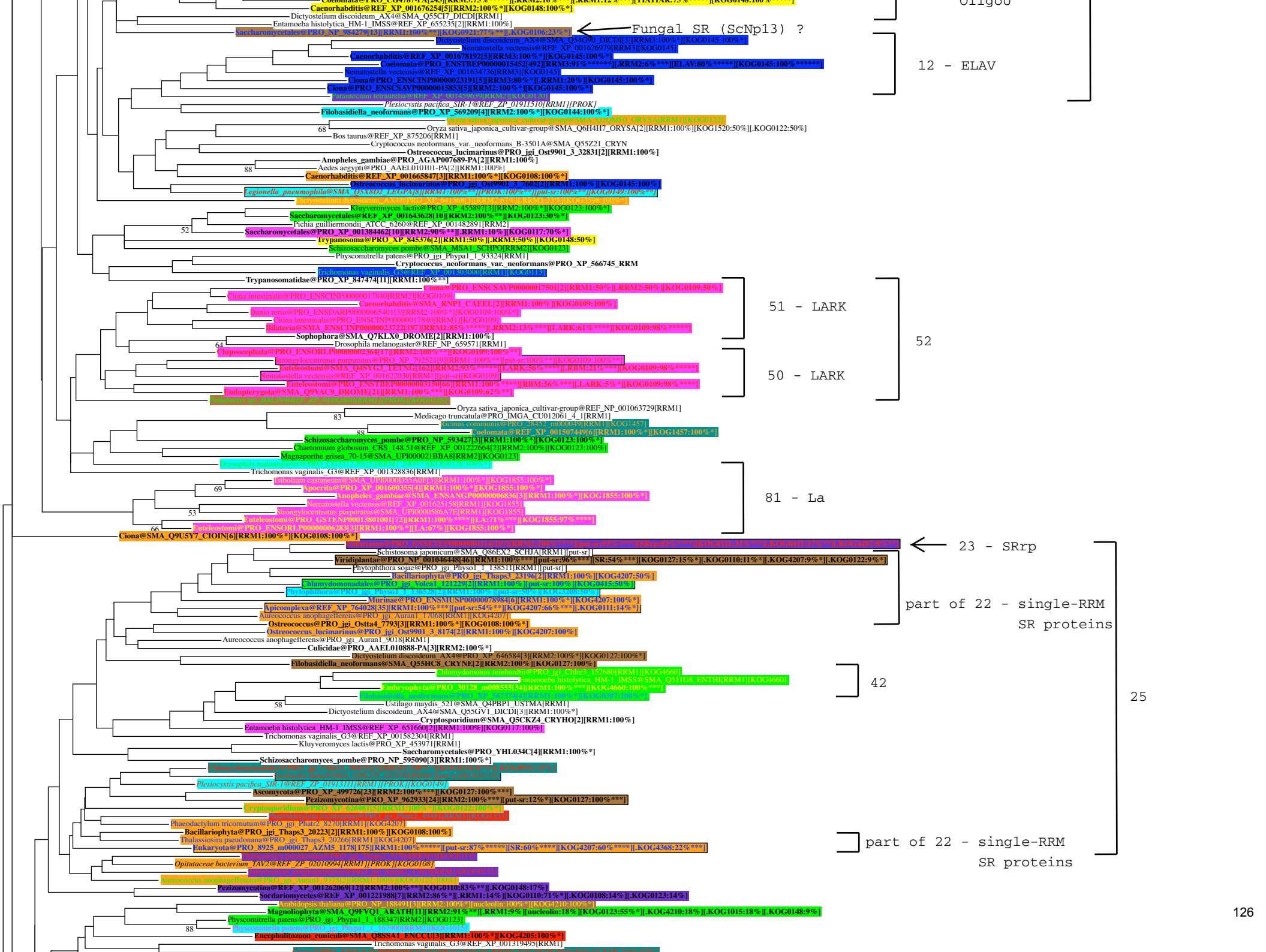

Fungal SR (ScNp13) ?
12 - ELAV
51 - LARK
50 - LARK
52
81 - La
23 - SRrp
part of 22 - single-RRM SR proteins
42
25
part of 22 - single-RRM SR proteins

Fig. S9

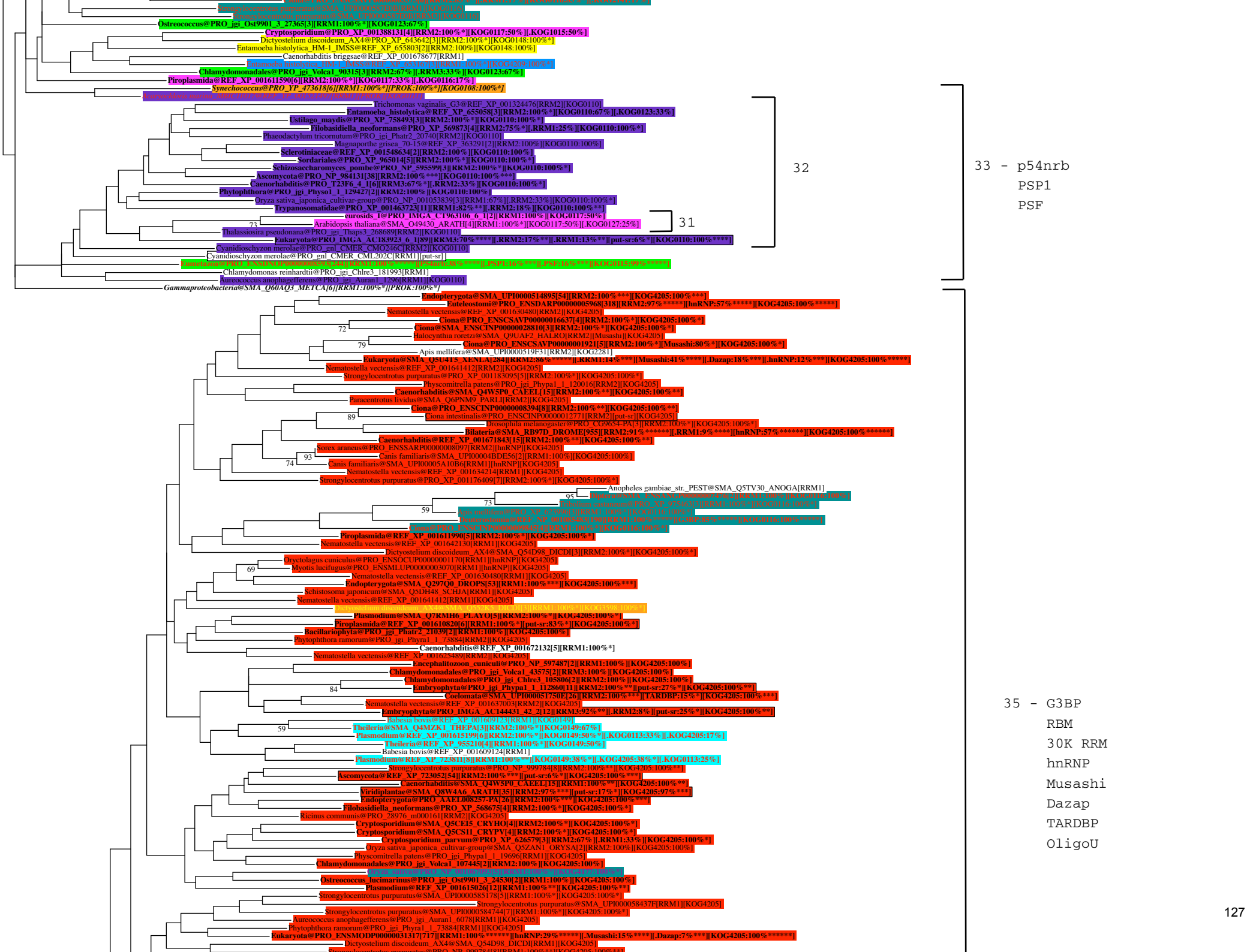

Ostreococcus@PRO_jgi_Ost9901_3_27365[3][RRM1:100%*][KOG0123:67%*]
Cryptosporidium@PRO_XP_001388131[4][RRM2:100%*][KOG0117:50%][.KOG1015:50%]
Dictyostelium discoideum_AX4@PRO_XP_643642[3][RRM2:100%*][KOG0148:100%*]
Entamoeba histolytica_HM-1_IMSS@REF_XP_655803[2][RRM2:100%][KOG0148:100%]
Caenorhabditis briggsae@REF_XP_001678677[RRM1]
Chlamydomonadales@PRO_jgi_Volca1_90315[3][RRM2:67%][.RRM3:33%][KOG0123:67%]
Piroplasmida@REF_XP_001611590[6][RRM2:100%*][KOG0117:33%][.KOG0116:17%]
Synechococcus@PRO_YP_473618[6][RRM1:100%*][PROK:100%*][KOG0108:100%*]
Trichomonas vaginalis_G3@REF_XP_001324476[RRM2][KOG0110]
Entamoeba_histolytica@REF_XP_655058[3][RRM2:100%*][KOG0110:67%][.KOG0123:33%]
Ustilago_maydis@PRO_XP_758493[3][RRM2:100%*][KOG0110:100%*]
Filobasidiella_neoformans@PRO_XP_569873[4][RRM2:75%*][.RRM1:25%][KOG0110:100%*]
Phaeodactylum tricornutum@PRO_jgi_Phatr2_20740[RRM2][KOG0110]
Magnaporthe grisea_70-15@REF_XP_363291[2][RRM2:100%][KOG0110:100%]
Sclerotiniaceae@REF_XP_001548634[2][RRM2:100%][KOG0110:100%]
Sordariales@PRO_XP_965014[5][RRM2:100%*][KOG0110:100%*]
Schizosaccharomyces_pombe@PRO_NP_595599[3][RRM2:100%*][KOG0110:100%*]
Ascomycota@PRO_NP_984131[38][RRM2:100%***][KOG0110:100%***]
Caenorhabditis@PRO_T23F6_4_1[6][RRM3:67%*][.RRM2:33%][KOG0110:100%*]
Phytophthora@PRO_jgi_Physo1_1_129472[2][RRM2:100%][KOG0110:100%]
Oryza sativa_japonica_cultivar-group@PRO_NP_001053839[3][RRM1:67%][.RRM2:33%][KOG0110:100%*]
Trypanosomatidae@PRO_XP_001463723[11][RRM1:82%**][.RRM2:18%][KOG0110:100%**]
eurosids_I@PRO_IMGA_C1963106_6_1[2][RRM1:100%*][KOG0117:50%]
Arabidopsis thaliana@SMA_O49430_ARATH[4][RRM1:100%*][KOG0117:50%][.KOG0127:25%]
Thalassiosira pseudonana@PRO_jgi_Thaps3_268689[RRM2][KOG0110]
Eukaryota@PRO_IMGA_AC183923_6_1[89][RRM3:70%****][.RRM2:17%**][.RRM1:13%**][put-sr:6%*][KOG0110:100%****]
Cyanidioschyzon merolae@PRO_gnl_CMER_CMO246C[RRM2][KOG0110]
Cyanidioschyzon merolae@PRO_gnl_CMER_CML202C[RRM1][put-sr]
Chlamydomonas reinhardtii@PRO_jgi_Chlre3_181993[RRM1]
Aureococcus anophagefferens@PRO_jgi_Auran1_1296[RRM1][KOG0110]
Gammaproteobacteria@SMA_Q60AQ3_METCA[6][RRM1:100%*][PROK:100%*]
32
31
33 - p54nrb
PSP1
PSF
Endopterygota@SMA_UPI00000514895[54][RRM2:100%***][KOG4205:100%***]
Euteleostomi@PRO_ENSDARP00000005968[318][RRM2:97%*****][hnRNP:57%*****][KOG4205:100%*****]
Nematostella vectensis@REF_XP_001630480[RRM2][KOG4205]
Ciona@PRO_ENSCSAVP00000016637[4][RRM2:100%*][KOG4205:100%*]
Ciona@SMA_ENSCINP00000028810[3][RRM2:100%*][KOG4205:100%*]
Halocynthia roretzi@SMA_Q9UAF2_HALRO[RRM2][Musashi][KOG4205]
Ciona@PRO_ENSCSAVP00000001921[5][RRM2:100%*][Musashi:80%*][KOG4205:100%*]
Apis mellifera@SMA_UPI0000519F31[RRM2][KOG2281]
Eukaryota@SMA_Q5U415_XENLA[284][RRM2:80%*****][.RRM1:14%***][Musashi:41%****][.Dazap:18%***][.hnRNP:12%***][KOG4205:100%*****]
Nematostella vectensis@REF_XP_001641412[RRM2][KOG4205]
Strongylocentrotus purpuratus@PRO_XP_001183095[5][RRM2:100%*][KOG4205:100%*]
Physcomitrella patens@PRO_jgi_Phypa1_1_120016[RRM2][KOG4205]
Caenorhabditis@SMA_Q4W5P0_CAEEL[15][RRM2:100%**][KOG4205:100%**]
Paracentrotus lividus@SMA_Q6PNM9_PARLI[RRM2][KOG4205]
Ciona@PRO_ENSCINP00000008394[8][RRM2:100%**][KOG4205:100%**]
Ciona intestinalis@PRO_ENSCINP00000012771[RRM2][put-sr][KOG4205]
Drosophila melanogaster@PRO_CG9654-PA[3][RRM2:100%*][KOG4205:100%*]
Bilateria@SMA_RB97D_DROME[955][RRM2:91%******][.RRM1:9%****][hnRNP:57%******][KOG4205:100%******]
Caenorhabditis@REF_XP_001671843[15][RRM2:100%**][KOG4205:100%**]
Sorex araneus@PRO_ENSSARP00000008097[RRM2][hnRNP][KOG4205]
Canis familiaris@SMA_UPI00004BDE56[2][RRM1:100%][KOG4205:100%]
Canis familiaris@SMA_UPI00005A10B6[RRM1][hnRNP][KOG4205]
Nematostella vectensis@REF_XP_001634214[RRM1][KOG4205]
Strongylocentrotus purpuratus@PRO_XP_001176409[7][RRM2:100%*][KOG4205:100%*]
Anopheles gambiae_str._PEST@SMA_Q5TV30_ANOGA[RRM1]
Piroplasmida@REF_XP_001611990[5][RRM2:100%*][KOG4205:100%*]
Nematostella vectensis@REF_XP_001642130[RRM1][KOG4205]
Dictyostelium discoideum_AX4@SMA_Q54D98_DICDI[3][RRM2:100%*][KOG4205:100%*]
Oryctolagus cuniculus@PRO_ENSOCUP00000001170[RRM1][hnRNP][KOG4205]
Myotis lucifugus@PRO_ENSMLUP00000003070[RRM1][hnRNP][KOG4205]
Nematostella vectensis@REF_XP_001630480[RRM1][KOG4205]
Endopterygota@SMA_Q297Q0_DROPS[53][RRM1:100%***][KOG4205:100%***]
Schistosoma japonicum@SMA_Q5DH48_SCHJA[RRM1][KOG4205]
Nematostella vectensis@REF_XP_001641412[RRM1][KOG4205]
Dictyostelium discoideum_AX4@SMA_Q552K5_DICDI[3][RRM1:100%*][KOG3598:100%*]
Plasmodium@SMA_Q7KMH6_PLAYO[5][RRM2:100%*][KOG4205:100%*]
Piroplasmida@REF_XP_001610820[6][RRM1:100%*][put-sr:83%*][KOG4205:100%*]
Bacillariophyta@PRO_jgi_Phatr2_21039[2][RRM1:100%][KOG4205:100%]
Phytophthora ramorum@PRO_jgi_Phyra1_1_73884[RRM2][KOG4205]
Caenorhabditis@REF_XP_001672132[5][RRM1:100%*]
Nematostella vectensis@REF_XP_001625489[RRM2][KOG4205]
Encephalitozoon_cuniculi@PRO_NP_597487[2][RRM1:100%][KOG4205:100%]
Chlamydomonadales@PRO_jgi_Volca1_43575[2][RRM3:100%][KOG4205:100%]
Chlamydomonadales@PRO_jgi_Chlre3_105806[2][RRM2:100%][KOG4205:100%]
Embryophyta@PRO_jgi_Phypa1_1_112860[11][RRM2:100%**][put-sr:27%*][KOG4205:100%**]
Coelomata@SMA_UPI0000517501[26][RRM2:100%***][TARDBP:15%*][KOG4205:100%***]
Nematostella vectensis@REF_XP_001637003[RRM2][KOG4205]
Embryophyta@PRO_IMGA_AC144431_42_2[12][RRM3:92%**][.RRM2:8%][put-sr:25%*][KOG4205:100%**]
Babesia bovis@REF_XP_001609128[RRM1][KOG0149]
Theileria@SMA_Q4MZK1_THEPA[3][RRM2:100%*][KOG0149:67%]
Plasmodium@REF_XP_001615199[6][RRM2:100%*][KOG0149:50%*][.KOG0113:33%][.KOG4205:17%]
Theileria@REF_XP_955210[4][RRM1:100%*][KOG0149:50%]
Babesia bovis@REF_XP_001609124[RRM1]
Plasmodium@REF_XP_723811[8][RRM1:100%**][KOG0149:38%*][.KOG4205:38%*][.KOG0113:25%]
Strongylocentrotus purpuratus@PRO_NP_999784[8][RRM2:100%**][KOG4205:100%**]
Ascomycota@REF_XP_723052[54][RRM2:100%***][put-sr:6%*][KOG4205:100%***]
Caenorhabditis@SMA_Q4W5P0_CAEEL[15][RRM1:100%**][KOG4205:100%**]
Viridiplantae@SMA_Q8W4A6_ARATH[35][RRM2:97%***][put-sr:17%*][KOG4205:97%***]
Endopterygota@PRO_AAEL008257-PA[26][RRM2:100%***][KOG4205:100%***]
Filobasidiella_neoformans@PRO_XP_568675[4][RRM2:100%*][KOG4205:100%*]
Ricinus communis@PRO_28976_m000161[RRM2][KOG4205]
Cryptosporidium@SMA_Q5CEI5_CRYHO[4][RRM2:100%*][KOG4205:100%*]
Cryptosporidium@SMA_Q5CS11_CRYPV[4][RRM2:100%*][KOG4205:100%*]
Cryptosporidium_parvum@PRO_XP_626579[3][RRM2:67%][.RRM1:33%][KOG4205:100%*]
Oryza sativa_japonica_cultivar-group@SMA_Q5ZAN1_ORYSA[2][RRM2:100%][KOG4205:100%]
Physcomitrella patens@PRO_jgi_Phypa1_1_19696[RRM1][KOG4205]
Chlamydomonadales@PRO_jgi_Volca1_107445[2][RRM2:100%][KOG4205:100%]
Ostreococcus_lucimarinus@PRO_jgi_Ost9901_3_24530[2][RRM1:100%][KOG4205:100%]
Plasmodium@REF_XP_001615026[12][RRM1:100%**][KOG4205:100%**]
Strongylocentrotus purpuratus@SMA_UPI0000585178[5][RRM1:100%*][KOG4205:100%*]
Strongylocentrotus purpuratus@SMA_UPI000058437F[RRM1][KOG4205]
Strongylocentrotus purpuratus@SMA_UPI0000584744[7][RRM1:100%*][KOG4205:100%*]
Aureococcus anophagefferens@PRO_jgi_Auran1_6078[RRM1][KOG4205]
Phytophthora ramorum@PRO_jgi_Phyra1_1_73884[RRM1][KOG4205]
Eukaryota@PRO_ENSMODP00000031317[717][RRM1:100%******][hnRNP:29%*****][.Musashi:15%****][.Dazap:7%***][KOG4205:100%******]
Dictyostelium discoideum_AX4@SMA_Q54D98_DICDI[RRM1][KOG4205]
35 - G3BP
RBM
30K RRM
hnRNP
Musashi
Dazap
TARDBP
OligoU

Fig. S9

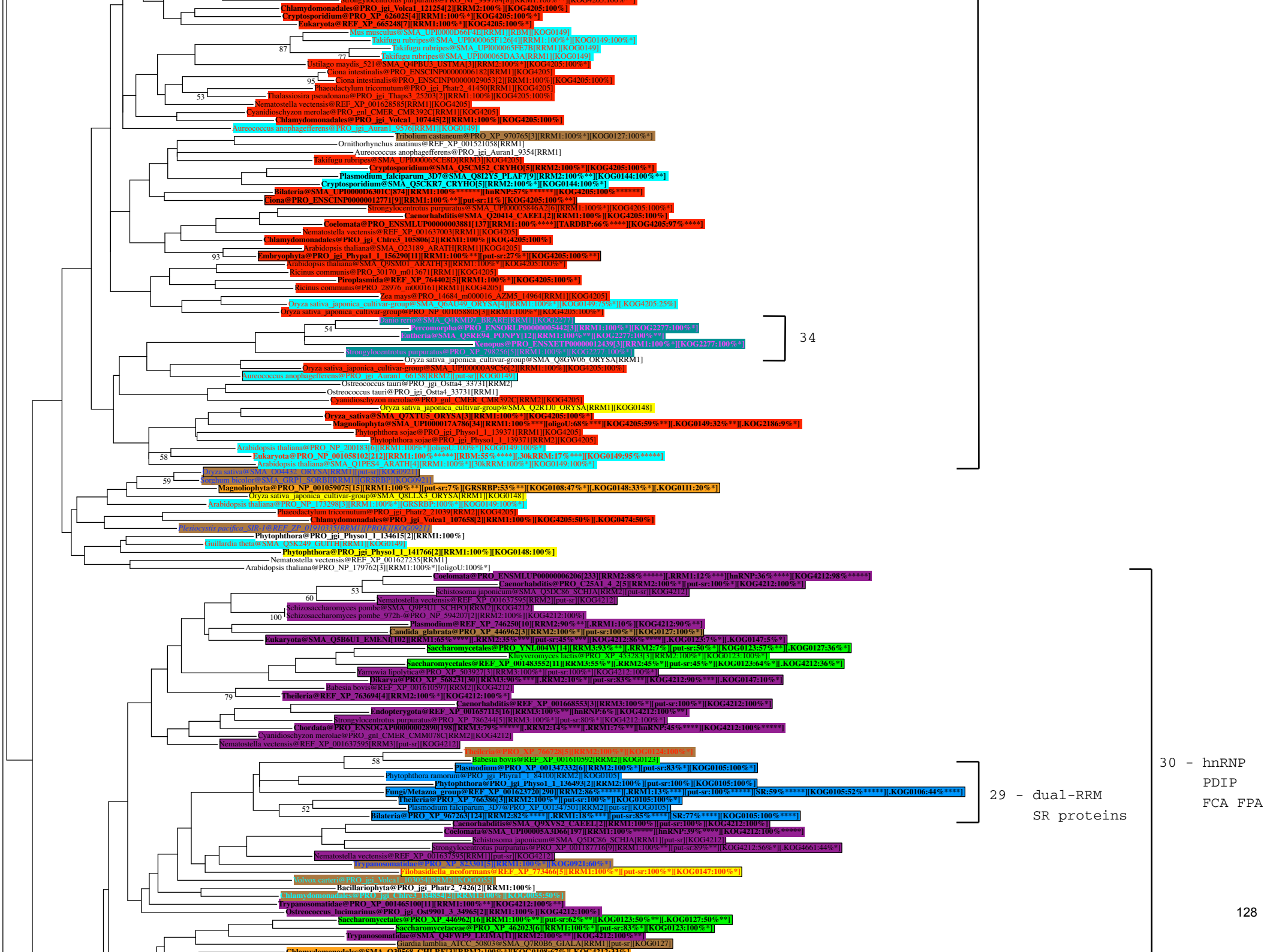
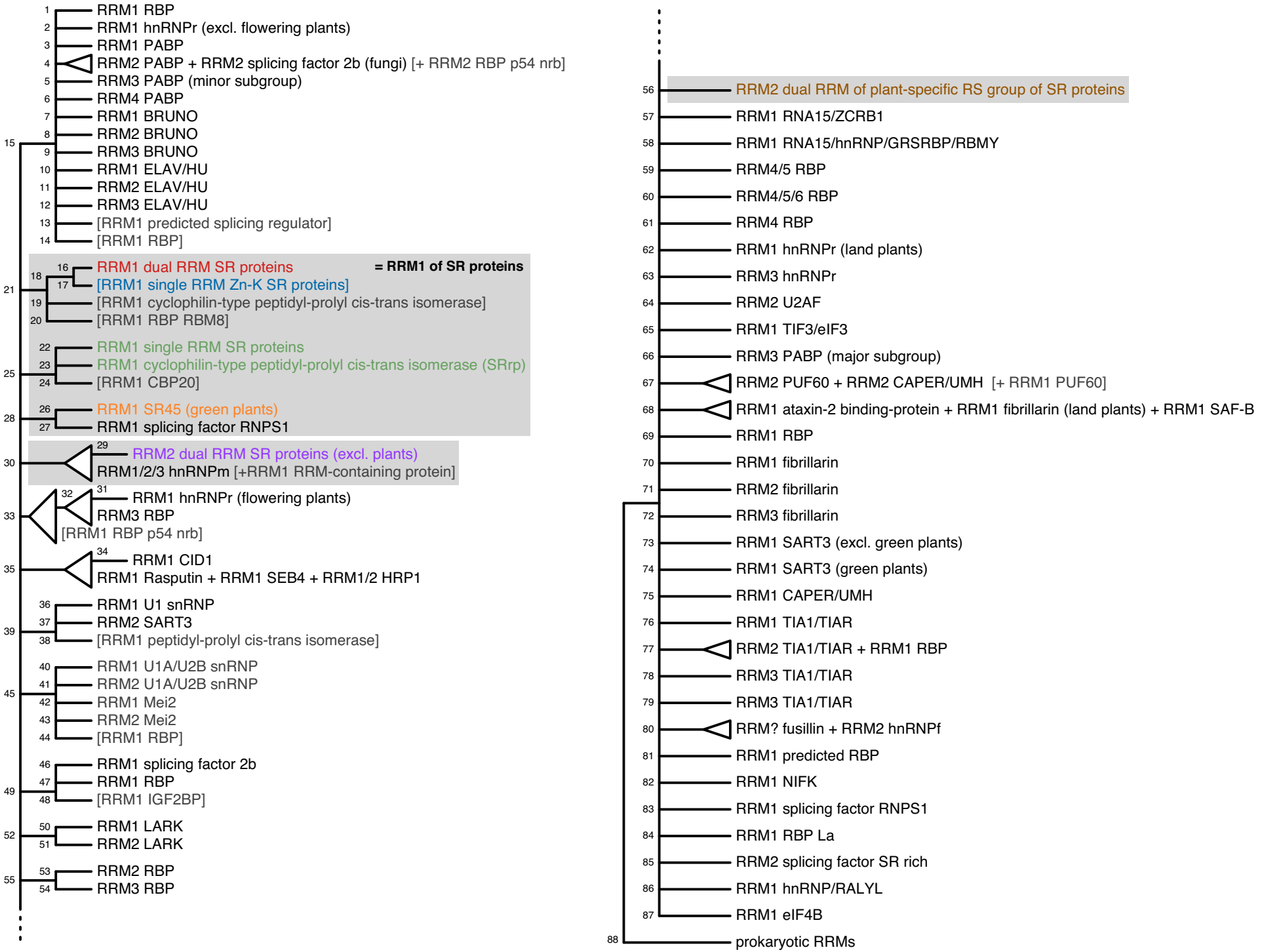


Fig. S9

Fig. S10


1 RRM1 RBP
2 RRM1 hnRNPr (excl. flowering plants)
3 RRM1 PABP
4 RRM2 PABP + RRM2 splicing factor 2b (fungi) [+ RRM2 RBP p54 nrb]
5 RRM3 PABP (minor subgroup)
6 RRM4 PABP
7 RRM1 BRUNO
8 RRM2 BRUNO
9 RRM3 BRUNO
10 RRM1 ELAV/HU
11 RRM2 ELAV/HU
12 RRM3 ELAV/HU
13 [RRM1 predicted splicing regulator]
14 [RRM1 RBP]
15
16 RRM1 dual RRM SR proteins
= RRM1 of SR proteins
17 [RRM1 single RRM Zn-K SR proteins]
18
19 [RRM1 cyclophilin-type peptidyl-prolyl cis-trans isomerase]
20 [RRM1 RBP RBM8]
21
22 RRM1 single RRM SR proteins
23 RRM1 cyclophilin-type peptidyl-prolyl cis-trans isomerase (SRrp)
24 [RRM1 CBP20]
25
26 RRM1 SR45 (green plants)
27 RRM1 splicing factor RNPS1
28
29 RRM2 dual RRM SR proteins (excl. plants)
RRM1/2/3 hnRNPm [+RRM1 RRM-containing protein]
30
31 RRM1 hnRNPr (flowering plants)
32
RRM3 RBP
[RRM1 RBP p54 nrb]
33
34 RRM1 CID1
RRM1 Rasputin + RRM1 SEB4 + RRM1/2 HRP1
35
36 RRM1 U1 snRNP
37 RRM2 SART3
38 [RRM1 peptidyl-prolyl cis-trans isomerase]
39
40 RRM1 U1A/U2B snRNP
41 RRM2 U1A/U2B snRNP
42 RRM1 Mei2
43 RRM2 Mei2
44 [RRM1 RBP]
45
46 RRM1 splicing factor 2b
47 RRM1 RBP
48 [RRM1 IGF2BP]
49
50 RRM1 LARK
51 RRM2 LARK
52
53 RRM2 RBP
54 RRM3 RBP
55
56 RRM2 dual RRM of plant-specific RS group of SR proteins
57 RRM1 RNA15/ZCRB1
58 RRM1 RNA15/hnRNP/GRSRBP/RBMY
59 RRM4/5 RBP
60 RRM4/5/6 RBP
61 RRM4 RBP
62 RRM1 hnRNPr (land plants)
63 RRM3 hnRNPr
64 RRM2 U2AF
65 RRM1 TIF3/eIF3
66 RRM3 PABP (major subgroup)
67 RRM2 PUF60 + RRM2 CAPER/UMH [+ RRM1 PUF60]
68 RRM1 ataxin-2 binding-protein + RRM1 fibrillarin (land plants) + RRM1 SAF-B
69 RRM1 RBP
70 RRM1 fibrillarin
71 RRM2 fibrillarin
72 RRM3 fibrillarin
73 RRM1 SART3 (excl. green plants)
74 RRM1 SART3 (green plants)
75 RRM1 CAPER/UMH
76 RRM1 TIA1/TIAR
77 RRM2 TIA1/TIAR + RRM1 RBP
78 RRM3 TIA1/TIAR
79 RRM3 TIA1/TIAR
80 RRM? fusillin + RRM2 hnRNPf
81 RRM1 predicted RBP
82 RRM1 NIFK
83 RRM1 splicing factor RNPS1
84 RRM1 RBP La
85 RRM2 splicing factor SR rich
86 RRM1 hnRNP/RALYL
87 RRM1 eIF4B
88 prokaryotic RRMs

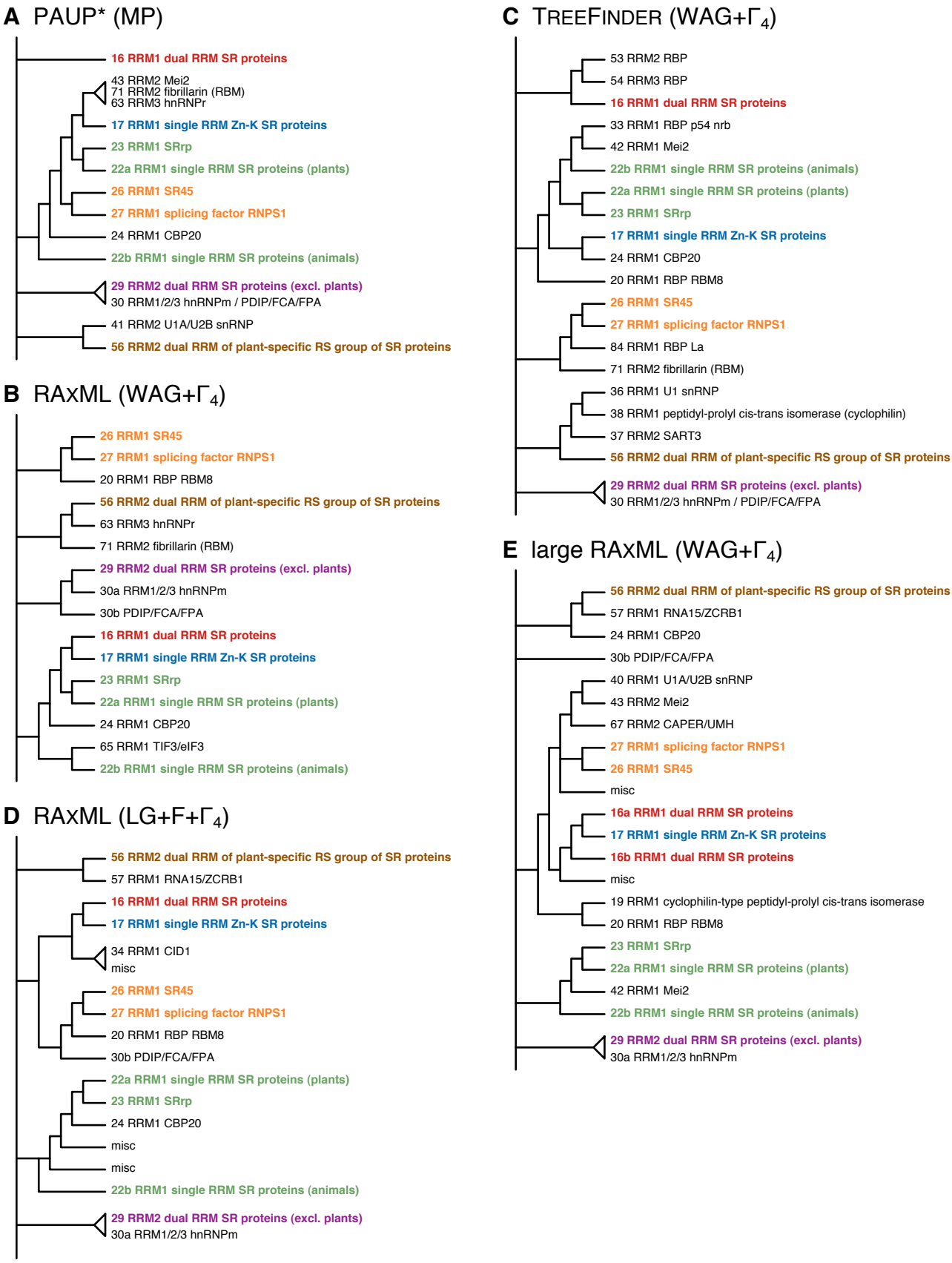

A PAUP* (MP)
16 RRM1 dual RRM SR proteins
43 RRM2 Mei2
71 RRM2 fibrillarin (RBM)
63 RRM3 hnRNPr
17 RRM1 single RRM Zn-K SR proteins
23 RRM1 SRrp
22a RRM1 single RRM SR proteins (plants)
26 RRM1 SR45
27 RRM1 splicing factor RNPS1
24 RRM1 CBP20
22b RRM1 single RRM SR proteins (animals)
29 RRM2 dual RRM SR proteins (excl. plants)
30 RRM1/2/3 hnRNPm / PDIP/FCA/FPA
41 RRM2 U1A/U2B snRNP
56 RRM2 dual RRM of plant-specific RS group of SR proteins
B RAxML (WAG+Γ4)
26 RRM1 SR45
27 RRM1 splicing factor RNPS1
20 RRM1 RBP RBM8
56 RRM2 dual RRM of plant-specific RS group of SR proteins
63 RRM3 hnRNPr
71 RRM2 fibrillarin (RBM)
29 RRM2 dual RRM SR proteins (excl. plants)
30a RRM1/2/3 hnRNPm
30b PDIP/FCA/FPA
16 RRM1 dual RRM SR proteins
17 RRM1 single RRM Zn-K SR proteins
23 RRM1 SRrp
22a RRM1 single RRM SR proteins (plants)
24 RRM1 CBP20
65 RRM1 TIF3/eIF3
22b RRM1 single RRM SR proteins (animals)
D RAxML (LG+F+Γ4)
56 RRM2 dual RRM of plant-specific RS group of SR proteins
57 RRM1 RNA15/ZCRB1
16 RRM1 dual RRM SR proteins
17 RRM1 single RRM Zn-K SR proteins
34 RRM1 CID1
misc
26 RRM1 SR45
27 RRM1 splicing factor RNPS1
20 RRM1 RBP RBM8
30b PDIP/FCA/FPA
22a RRM1 single RRM SR proteins (plants)
23 RRM1 SRrp
24 RRM1 CBP20
misc
misc
22b RRM1 single RRM SR proteins (animals)
29 RRM2 dual RRM SR proteins (excl. plants)
30a RRM1/2/3 hnRNPm
C TREEFINDER (WAG+Γ4)
53 RRM2 RBP
54 RRM3 RBP
16 RRM1 dual RRM SR proteins
33 RRM1 RBP p54 nrb
42 RRM1 Mei2
22b RRM1 single RRM SR proteins (animals)
22a RRM1 single RRM SR proteins (plants)
23 RRM1 SRrp
17 RRM1 single RRM Zn-K SR proteins
24 RRM1 CBP20
20 RRM1 RBP RBM8
26 RRM1 SR45
27 RRM1 splicing factor RNPS1
84 RRM1 RBP La
71 RRM2 fibrillarin (RBM)
36 RRM1 U1 snRNP
38 RRM1 peptidyl-prolyl cis-trans isomerase (cyclophilin)
37 RRM2 SART3
56 RRM2 dual RRM of plant-specific RS group of SR proteins
29 RRM2 dual RRM SR proteins (excl. plants)
30 RRM1/2/3 hnRNPm / PDIP/FCA/FPA
E large RAxML (WAG+Γ4)
56 RRM2 dual RRM of plant-specific RS group of SR proteins
57 RRM1 RNA15/ZCRB1
24 RRM1 CBP20
30b PDIP/FCA/FPA
40 RRM1 U1A/U2B snRNP
43 RRM2 Mei2
67 RRM2 CAPER/UMH
27 RRM1 splicing factor RNPS1
26 RRM1 SR45
misc
16a RRM1 dual RRM SR proteins
17 RRM1 single RRM Zn-K SR proteins
16b RRM1 dual RRM SR proteins
misc
19 RRM1 cyclophilin-type peptidyl-prolyl cis-trans isomerase
20 RRM1 RBP RBM8
23 RRM1 SRrp
22a RRM1 single RRM SR proteins (plants)
42 RRM1 Mei2
22b RRM1 single RRM SR proteins (animals)
29 RRM2 dual RRM SR proteins (excl. plants)
30a RRM1/2/3 hnRNPm

Fig. S12

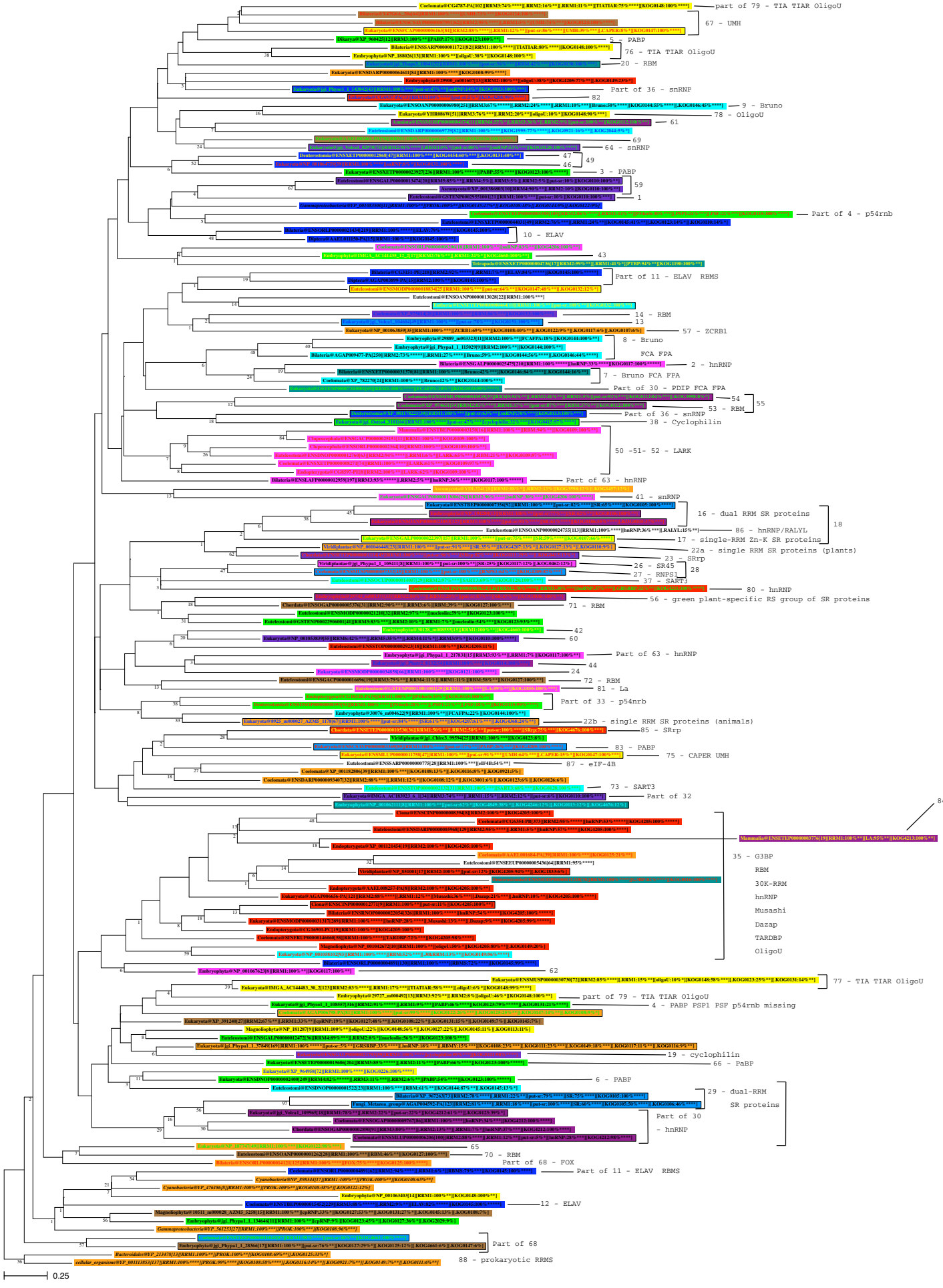

part of 79 - TIA TIAR OligoU
67 - UMH
5 - PABP
76 - TIA TIAR OligoU
20 - RBM
Part of 36 - snRNP
82
9 - Bruno
78 - OligoU
61
69
64 - snRNP
47
49
46
3 - PABP
59
1
Part of 4 - p54rnb
10 - ELAV
43
Part of 11 - ELAV RBMS
14 - RBM
13
57 - ZCRB1
8 - Bruno
FCA FPA
2 - hnRNP
7 - Bruno FCA FPA
Part of 30 - PDIP FCA FPA
54
55
53 - RBM
Part of 36 - snRNP
38 - Cyclophilin
50 -51- 52 - LARK
Part of 63 - hnRNP
41 - snRNP
16 - dual RRM SR proteins
18
86 - hnRNP/RALYL
17 - single-RRM Zn-K SR proteins
22a - single RRM SR proteins (plants)
23 - SRrp
26 - SR45
28
27 - RNPS1
37 - SART3
80 - hnRNP
56 - green plant-specific RS group of SR proteins
71 - RBM
42
60
Part of 63 - hnRNP
44
24
72 - RBM
81 - La
Part of 33 - p54nrb
22b - single RRM SR proteins (animals)
85 - SRrp
83 - PABP
75 - CAPER UMH
87 - eIF-4B
73 - SART3
Part of 32
84 - La
35 - G3BP
RBM
30K-RRM
hnRNP
Musashi
Dazap
TARDBP
OligoU
62
77 - TIA TIAR OligoU
part of 79 - TIA TIAR OligoU
4 - PABP PSP1 PSF p54rnb missing
19 - cyclophilin
66 - PaBP
6 - PABP
29 - dual-RRM
SR proteins
Part of 30
- hnRNP
65
70 - RBM
Part of 68 - FOX
Part of 11 - ELAV RBMS
12 - ELAV
Part of 68
88 - prokaryotic RRMS
0.25

Fig. S13

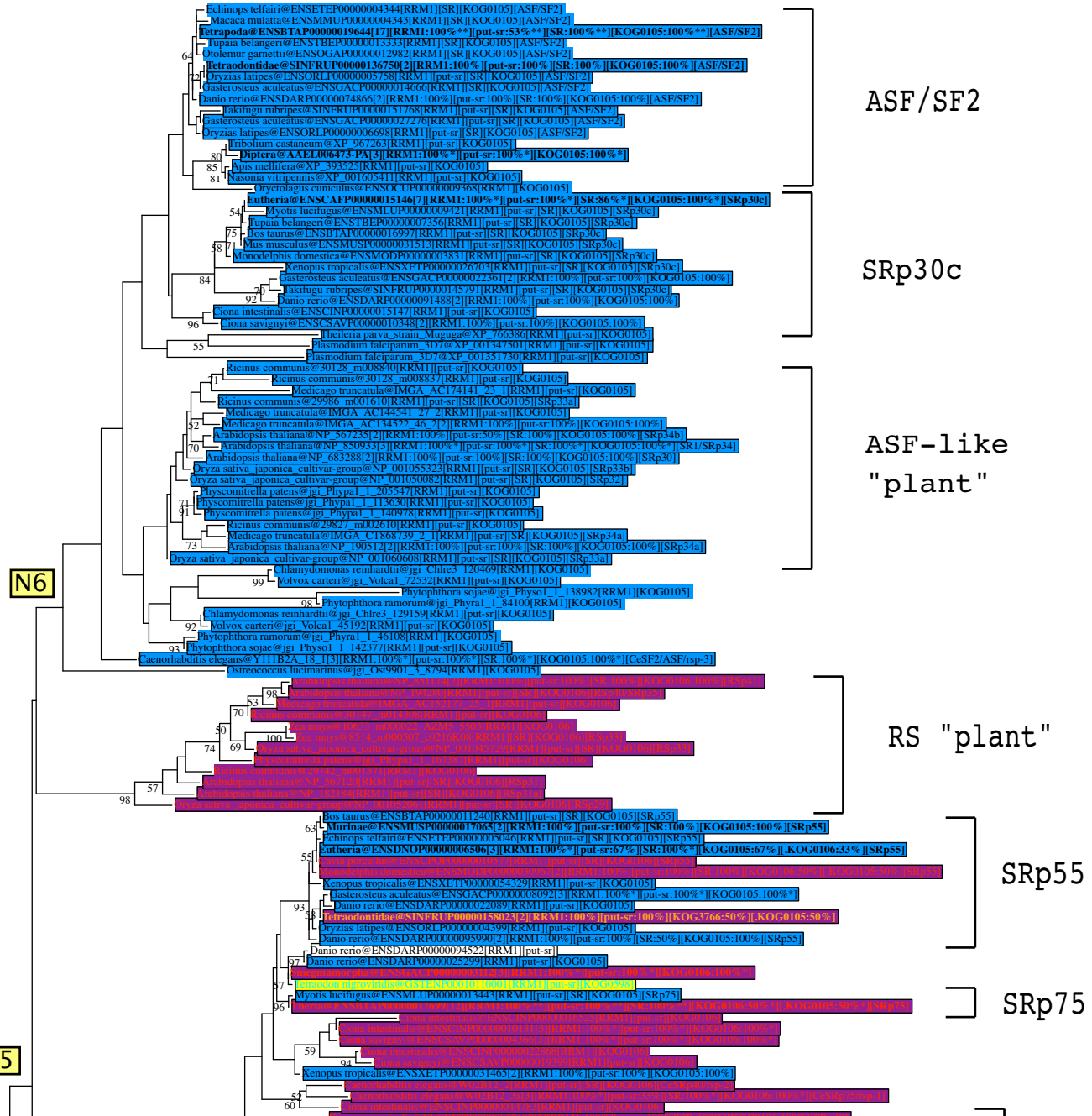


Fig. S13

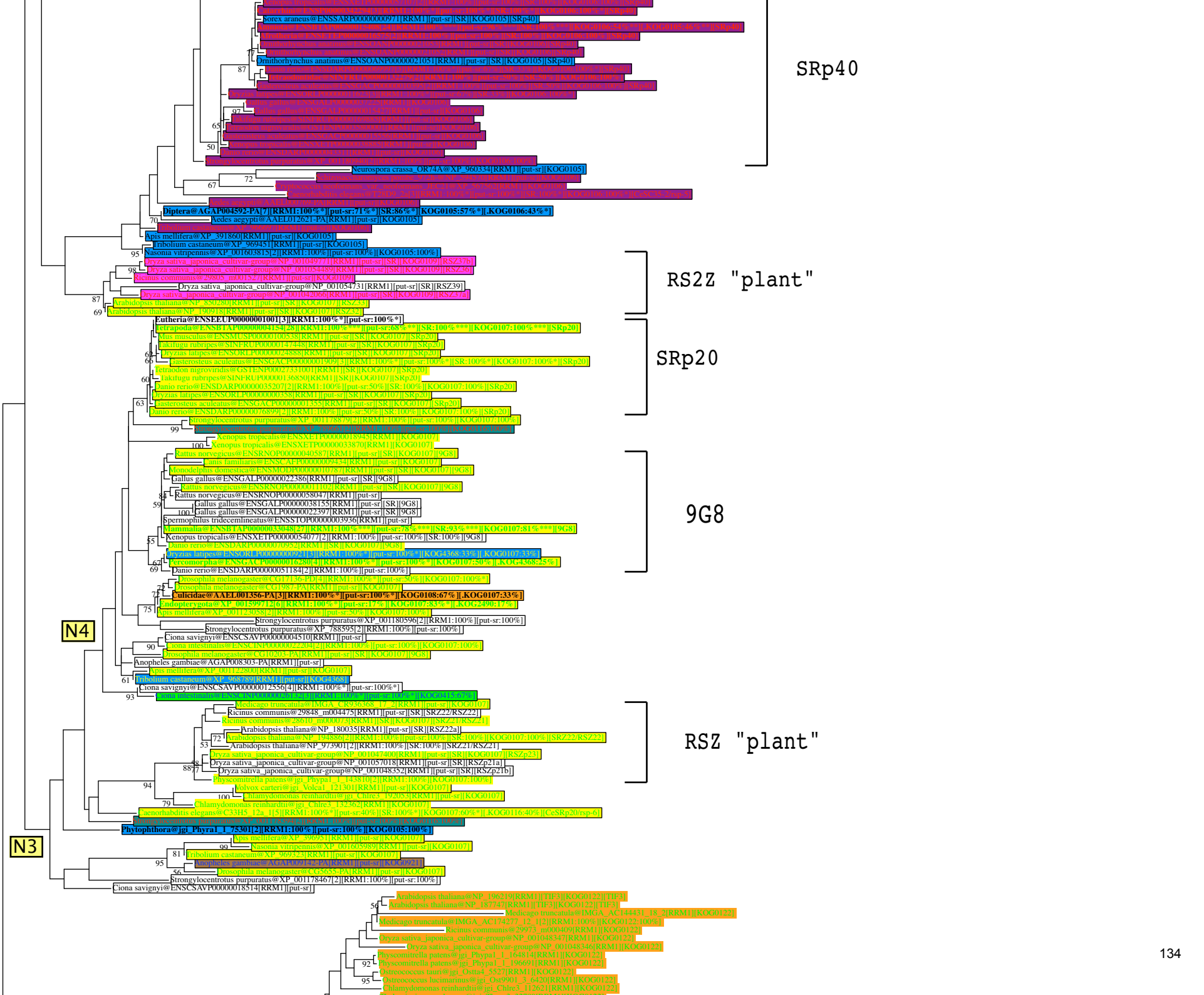

SRp40
RS2Z "plant"
SRp20
9G8
RSZ "plant"
N4
N3
Sorex araneus@ENSSARP00000000971[RRM1][put-sr][SR][KOG0105][SRp40]
Ornithorhynchus anatinus@ENSOANP00000021051[RRM1][put-sr][SR][KOG0105][SRp40]
Neurospora crassa_OR74A@XP_960334[RRM1][put-sr][KOG0105]
Diptera@AGAP004592-PA[7][RRM1:100%*][put-sr:71%*][SR:86%*][KOG0105:57%*][.KOG0106:43%*]
Aedes aegypti@AAEL012621-PA[RRM1][put-sr][KOG0105]
Apis mellifera@XP_391860[RRM1][put-sr][KOG0105]
Tribolium castaneum@XP_969451[RRM1][put-sr][KOG0105]
Nasonia vitripennis@XP_001603815[2][RRM1:100%][put-sr:100%][KOG0105:100%]
Oryza sativa_japonica_cultivar-group@NP_001049771[RRM1][put-sr][SR][KOG0109][RSZ37b]
Oryza sativa_japonica_cultivar-group@NP_001054489[RRM1][put-sr][SR][KOG0109][RSZ36]
Ricinus communis@29805_m001527[RRM1][put-sr][KOG0109]
Oryza sativa_japonica_cultivar-group@NP_001054731[RRM1][put-sr][SR][RSZ39]
Oryza sativa_japonica_cultivar-group@NP_001049266[RRM1][put-sr][SR][KOG0109][RSZ37a]
Arabidopsis thaliana@NP_850280[RRM1][put-sr][SR][KOG0107][RSZ33]
Arabidopsis thaliana@NP_190918[RRM1][put-sr][SR][KOG0107][RSZ32]
Eutheria@ENSEEUP00000001001[3][RRM1:100%*][put-sr:100%*]
Tetrapoda@ENSTRUP00000004154[28][RRM1:100%***][put-sr:68%**][SR:100%***][KOG0107:100%***][SRp20]
Mus musculus@ENSMUSP00000100538[RRM1][put-sr][SR][KOG0107][SRp20]
Takifugu rubripes@SINFRUP00000147448[RRM1][put-sr][SR][KOG0107][SRp20]
Oryzias latipes@ENSORLP00000024888[RRM1][put-sr][SR][KOG0107][SRp20]
Gasterosteus aculeatus@ENSGACP00000019095[3][RRM1:100%*][put-sr:100%*][SR:100%*][KOG0107:100%*][SRp20]
Tetraodon nigroviridis@GSTENP00027331001[RRM1][SR][KOG0107][SRp20]
Takifugu rubripes@SINFRUP00000136850[RRM1][SR][KOG0107][SRp20]
Danio rerio@ENSDARP00000035207[2][RRM1:100%][put-sr:50%][SR:100%][KOG0107:100%][SRp20]
Oryzias latipes@ENSORLP00000000355[RRM1][put-sr][SR][KOG0107][SRp20]
Gasterosteus aculeatus@ENSGACP00000001355[RRM1][put-sr][SR][KOG0107][SRp20]
Danio rerio@ENSDARP00000076899[2][RRM1:100%][put-sr:50%][SR:100%][KOG0107:100%][SRp20]
Strongylocentrotus purpuratus@XP_001178879[2][RRM1:100%][put-sr:100%][KOG0107:100%]
Xenopus tropicalis@ENSXETP00000018945[RRM1][KOG0107]
Xenopus tropicalis@ENSXETP00000033870[RRM1][KOG0107]
Rattus norvegicus@ENSRNOP00000040587[RRM1][put-sr][SR][KOG0107][9G8]
Canis familiaris@ENSCAFP00000009434[RRM1][put-sr][KOG0107]
Monodelphis domestica@ENSMODP00000010787[RRM1][put-sr][SR][KOG0107][9G8]
Gallus gallus@ENSGALP00000022386[RRM1][put-sr][SR][9G8]
Rattus norvegicus@ENSRNOP00000011102[RRM1][put-sr][SR][KOG0107][9G8]
Rattus norvegicus@ENSRNOP00000058047[RRM1][put-sr]
Gallus gallus@ENSGALP00000038155[RRM1][put-sr][SR][9G8]
Gallus gallus@ENSGALP00000022397[RRM1][put-sr][SR][9G8]
Spermophilus tridecemlineatus@ENSSTOP00000003936[RRM1][put-sr]
Mammalia@ENSBTAP00000033048[27][RRM1:100%***][put-sr:78%***][SR:93%***][KOG0107:81%***][9G8]
Xenopus tropicalis@ENSXETP00000054077[2][RRM1:100%][put-sr:100%][SR:100%][9G8]
Danio rerio@ENSDARP00000070952[RRM1][SR][KOG0107][9G8]
Oryzias latipes@ENSORLP00000000921[3][RRM1:100%*][put-sr:100%*][KOG4368:33%][.KOG0107:33%]
Percomorpha@ENSGACP00000016280[4][RRM1:100%*][put-sr:100%*][KOG0107:50%][.KOG4368:25%]
Danio rerio@ENSDARP00000051184[2][RRM1:100%][put-sr:100%]
Drosophila melanogaster@CG17136-PD[4][RRM1:100%*][put-sr:50%][KOG0107:100%*]
Drosophila melanogaster@CG11087-PA[RRM1][put-sr][KOG0107]
Culicidae@AAEL001356-PA[5][RRM1:100%*][put-sr:100%*][KOG0108:67%][.KOG0107:33%]
Endopterygota@XP_001599712[6][RRM1:100%*][put-sr:17%][.KOG0107:83%*][.KOG2490:17%]
Apis mellifera@XP_001123058[2][RRM1:100%][put-sr:50%][KOG0107:100%]
Strongylocentrotus purpuratus@XP_001180596[2][RRM1:100%][put-sr:100%]
Strongylocentrotus purpuratus@XP_788595[2][RRM1:100%][put-sr:100%]
Ciona savignyi@ENSCSAVP00000004510[RRM1][put-sr]
Ciona intestinalis@ENSCINP00000022204[2][RRM1:100%][put-sr:100%][KOG0107:100%]
Drosophila melanogaster@CG10203-PA[RRM1][put-sr][SR][KOG0107][9G8]
Anopheles gambiae@AGAP008303-PA[RRM1][put-sr]
Apis mellifera@XP_001122800[RRM1][put-sr][KOG0107]
Tribolium castaneum@XP_968789[RRM1][put-sr][KOG4368]
Ciona savignyi@ENSCSAVP00000012556[4][RRM1:100%*][put-sr:100%*]
Ciona intestinalis@ENSCINP00000026132[3][RRM1:100%*][put-sr:100%*][KOG0115:67%]
Medicago truncatula@IMGA_CR936368_17_2[RRM1][put-sr][KOG0107]
Ricinus communis@29848_m004475[RRM1][put-sr][SR][SRZ22/RSZ22]
Ricinus communis@28610_m000075[RRM1][SR][KOG0107][SRZ21/RSZ21]
Arabidopsis thaliana@NP_180035[RRM1][put-sr][SR][RSZ22a]
Arabidopsis thaliana@NP_194886[2][RRM1:100%][put-sr:100%][SR:100%][KOG0107:100%][SRZ22/RSZ22]
Arabidopsis thaliana@NP_973901[2][RRM1:100%][SR:100%][SRZ21/RSZ21]
Oryza sativa_japonica_cultivar-group@NP_001047400[RRM1][put-sr][SR][KOG0107][RSZp23]
Oryza sativa_japonica_cultivar-group@NP_001057018[RRM1][put-sr][SR][RSZp21a]
Oryza sativa_japonica_cultivar-group@NP_001048352[RRM1][put-sr][SR][RSZp21b]
Physcomitrella patens@jgi_Phypa1_1_143810[2][RRM1:100%][KOG0107:100%]
Volvox carteri@jgi_Volca1_121301[RRM1][put-sr][KOG0107]
Chlamydomonas reinhardtii@jgi_Chlre3_192055[RRM1][put-sr][KOG0107]
Chlamydomonas reinhardtii@jgi_Chlre3_132262[RRM1][KOG0107]
Caenorhabditis elegans@C33H5_12a_1[5][RRM1:100%*][put-sr:40%][SR:100%*][KOG0107:60%*][.KOG0116:40%][CeSRp20/rsp-6]
Phytophthora@jgi_Phyra1_1_75301[2][RRM1:100%][put-sr:100%][KOG0105:100%]
Apis mellifera@XP_396951[RRM1][put-sr][KOG0107]
Nasonia vitripennis@XP_001605989[RRM1][put-sr][KOG0107]
Tribolium castaneum@XP_969323[RRM1][put-sr][KOG0107]
Anopheles gambiae@AGAP009142-PA[RRM1][put-sr][KOG0921]
Drosophila melanogaster@CG5655-PA[RRM1][put-sr][KOG0107]
Strongylocentrotus purpuratus@XP_001178467[2][RRM1:100%][put-sr:100%]
Ciona savignyi@ENSCSAVP00000018514[RRM1][put-sr]
Arabidopsis thaliana@NP_196219[RRM1][TIF3][KOG0122][TIF3]
Arabidopsis thaliana@NP_187747[RRM1][TIF3][KOG0122][TIF3]
Medicago truncatula@IMGA_AC144431_18_2[RRM1][KOG0122]
Medicago truncatula@IMGA_AC174277_12_1[2][RRM1:100%][KOG0122:100%]
Ricinus communis@29973_m000409[RRM1][KOG0122]
Oryza sativa_japonica_cultivar-group@NP_001048347[RRM1][KOG0122]
Oryza sativa_japonica_cultivar-group@NP_001048346[RRM1][KOG0122]
Physcomitrella patens@jgi_Phypa1_1_164814[RRM1][KOG0122]
Physcomitrella patens@jgi_Phypa1_1_196691[RRM1][KOG0122]
Ostreococcus tauri@jgi_Ostta4_5527[RRM1][KOG0122]
Ostreococcus lucimarinus@jgi_Ost9901_3_6420[RRM1][KOG0122]
Chlamydomonas reinhardtii@jgi_Chlre3_112621[RRM1][KOG0122]

Fig. S13

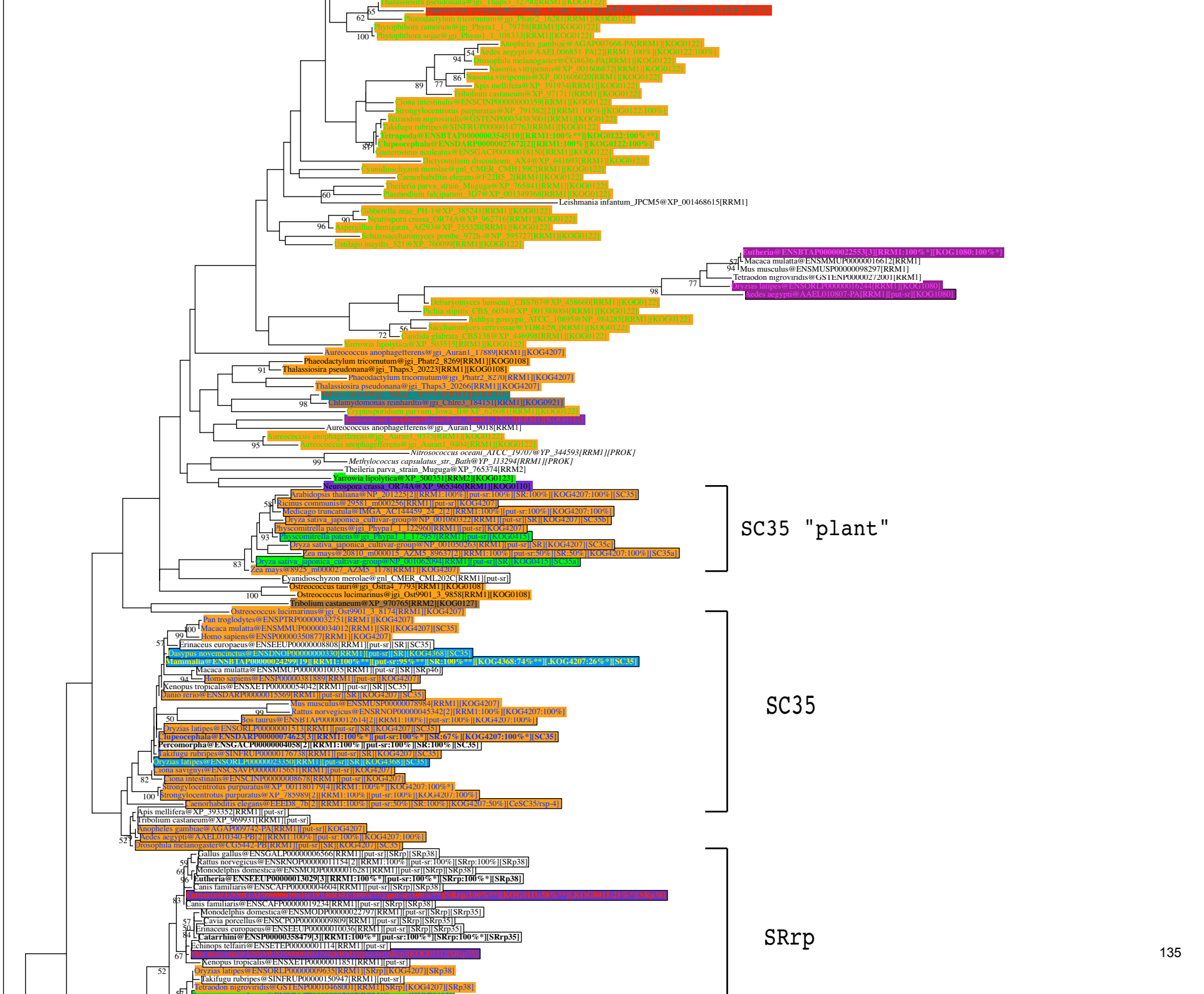

Eutheria@ENSBTAP00000022553[3][RRM1:100%*][KOG1080:100%*]
Macaca mulatta@ENSMMUP00000016612[RRM1]
Mus musculus@ENSMUSP00000098297[RRM1]
Tetraodon nigroviridis@GSTENP00000272001[RRM1]
Oryzias latipes@ENSORLP00000016244[RRM1][KOG1080]
Aedes aegypti@AAEL010807-PA[RRM1][put-sr][KOG1080]
Debaryomyces hansenii_CBS767@XP_458660[RRM1][KOG0122]
Pichia stipitis_CBS_6054@XP_001388004[RRM1][KOG0122]
Ashbya gossypii_ATCC_10895@NP_984285[RRM1][KOG0122]
Saccharomyces cerevisiae@YDR429C[RRM1][KOG0122]
Candida glabrata_CBS138@XP_446998[RRM1][KOG0122]
Yarrowia lipolytica@XP_503515[RRM1][KOG0122]
Aureococcus anophagefferens@jgi_Auran1_17889[RRM1][KOG4207]
Phaeodactylum tricornutum@jgi_Phatr2_8269[RRM1][KOG0108]
Thalassiosira pseudonana@jgi_Thaps3_20223[RRM1][KOG0108]
Phaeodactylum tricornutum@jgi_Phatr2_8270[RRM1][KOG4207]
Thalassiosira pseudonana@jgi_Thaps3_20266[RRM1][KOG4207]
Chlamydomonas reinhardtii@jgi_Chlre3_184151[RRM1][KOG0921]
Cryptosporidium parvum_Iowa_II@XP_626081[RRM1][KOG0122]
Aureococcus anophagefferens@jgi_Auran1_9018[RRM1]
Aureococcus anophagefferens@jgi_Auran1_9379[RRM1][KOG0122]
Aureococcus anophagefferens@jgi_Auran1_9404[RRM1][KOG0122]
Nitrosococcus oceani_ATCC_19707@YP_344593[RRM1][PROK]
Methylococcus capsulatus_str._Bath@YP_113294[RRM1][PROK]
Theileria parva_strain_Muguga@XP_765374[RRM2]
Yarrowia lipolytica@XP_500351[RRM2][KOG0123]
Neurospora crassa_OR74A@XP_965346[RRM1][KOG0110]
Arabidopsis thaliana@NP_201225[2][RRM1:100%][put-sr:100%][SR:100%][KOG4207:100%][SC35]
Ricinus communis@29581_m000256[RRM1][put-sr][KOG4207]
Medicago truncatula@IMGA_AC144459_24_2[2][RRM1:100%][put-sr:100%][KOG4207:100%]
Oryza sativa_japonica_cultivar-group@NP_001060322[RRM1][put-sr][SR][KOG4207][SC35b]
Physcomitrella patens@jgi_Phypa1_1_122960[RRM1][put-sr][KOG4207]
Physcomitrella patens@jgi_Phypa1_1_172957[RRM1][put-sr][KOG4415]
Oryza sativa_japonica_cultivar-group@NP_001050263[RRM1][put-sr][SR][KOG4207][SC35c]
Zea mays@20810_m000015_AZM5_89637[2][RRM1:100%][put-sr:50%][SR:50%][KOG4207:100%][SC35a]
Oryza sativa_japonica_cultivar-group@NP_001060744[RRM1][put-sr][SR][KOG4415][SC35d]
Zea mays@8925_m000027_AZM5_1[1][RRM1][KOG4207]
Cyanidioschyzon merolae@gnl_CMER_CML202C[RRM1][put-sr]
Ostreococcus tauri@jgi_Ostta4_7793[RRM1][KOG0108]
Ostreococcus lucimarinus@jgi_Ost9901_3_9858[RRM1][KOG0108]
Tribolium castaneum@XP_970765[RRM2][KOG0127]
Ostreococcus lucimarinus@jgi_Ost9901_3_8174[RRM1][KOG4207]
Pan troglodytes@ENSPTRP00000032751[RRM1][KOG4207]
Macaca mulatta@ENSMMUP00000034012[RRM1][SR][KOG4207][SC35]
Homo sapiens@ENSP00000350877[RRM1][KOG4207]
Erinaceus europaeus@ENSEEUP00000008808[RRM1][put-sr][SR][SC35]
Dasypus novemcinctus@ENSDNOP00000000330[RRM1][put-sr][SR][KOG4368][SC35]
Mammalia@ENSBTAP00000024299[19][RRM1:100%**][put-sr:95%**][SR:100%**][KOG4368:74%**][KOG4207:26%*][SC35]
Macaca mulatta@ENSMMUP00000010035[RRM1][put-sr][SR][SRp46]
Homo sapiens@ENSP00000381889[RRM1][put-sr][KOG4207]
Xenopus tropicalis@ENSXETP00000054042[RRM1][put-sr][SR][SC35]
Danio rerio@ENSDARP00000015569[RRM1][put-sr][SR][KOG4207][SC35]
Mus musculus@ENSMUSP00000078984[RRM1][KOG4207]
Rattus norvegicus@ENSRNOP00000045342[2][RRM1:100%][KOG4207:100%]
Bos taurus@ENSBTAP00000012614[2][RRM1:100%][put-sr:100%][KOG4207:100%]
Oryzias latipes@ENSORLP00000001513[RRM1][put-sr][SR][KOG4207][SC35]
Lipoecephala@ENSDARP00000074623[3][RRM1:100%*][put-sr:100%*][SR:67%][KOG4207:100%*][SC35]
Percomorpha@ENSGACP00000004058[2][RRM1:100%][put-sr:100%][SR:100%][SC35]
Takifugu rubripes@SINFRUP00000176738[RRM1][put-sr][SR][KOG4207][SC35]
Oryzias latipes@ENSORLP00000123550[RRM1][put-sr][SR][KOG4368][SC35]
Gasterosteus aculeatus@ENSGACP00000015581[RRM1][put-sr][KOG4207]
Ciona intestinalis@ENSCINP00000008678[RRM1][put-sr][KOG4207]
Strongylocentrotus purpuratus@XP_001180179[4][RRM1:100%*][KOG4207:100%*]
Strongylocentrotus purpuratus@XP_785989[2][RRM1:100%][put-sr:100%][KOG4207:100%]
Caenorhabditis elegans@F11D8_7b[2][RRM1:100%][put-sr:50%][SR:100%][KOG4207:50%][CeSC35/rsp-4]
Apis mellifera@XP_393352[RRM1][put-sr]
Tribolium castaneum@XP_969931[RRM1][put-sr]
Anopheles gambiae@AGAP009742-PA[RRM1][put-sr][KOG4207]
Aedes aegypti@AAEL010340-PB[2][RRM1:100%][put-sr:100%][KOG4207:100%]
Drosophila melanogaster@CG5442-PB[RRM1][put-sr][SR][KOG4207][SC35]
Gallus gallus@ENSGALP00000006566[RRM1][put-sr][SRp][SRp38]
Rattus norvegicus@ENSRNOP00000011154[2][RRM1:100%][put-sr:100%][SRrp:100%][SRp38]
Monodelphis domestica@ENSMODP00000016281[RRM1][put-sr][SRp][SRp38]
Eutheria@ENSEEUP00000013029[3][RRM1:100%*][put-sr:100%*][SRrp:100%*][SRp38]
Canis familiaris@ENSCAFP00000004604[RRM1][put-sr][SRp][SRp38]
Canis familiaris@ENSCAFP00000019234[RRM1][put-sr][SRp][SRp38]
Monodelphis domestica@ENSMODP00000022797[RRM1][put-sr][SRp][SRp35]
Cavia porcellus@ENSCPOP00000009809[RRM1][put-sr][SRp][SRp35]
Erinaceus europaeus@ENSEEUP00000010036[RRM1][put-sr][SRp][SRp35]
Catarrhini@ENSP00000358479[3][RRM1:100%*][put-sr:100%*][SRrp:100%*][SRrp35]
Echinops telfairi@ENSETEP00000001114[RRM1][put-sr]
Xenopus tropicalis@ENSXETP00000011851[RRM1][put-sr]
Oryzias latipes@ENSORLP00000009635[RRM1][SRp][KOG4207][SRp38]
Takifugu rubripes@SINFRUP00000150947[RRM1][put-sr]
Tetraodon nigroviridis@GSTENP00001046801[RRM1][SRp][KOG4207][SRp38]
SC35 "plant"
SC35
SRrp

Fig. S13

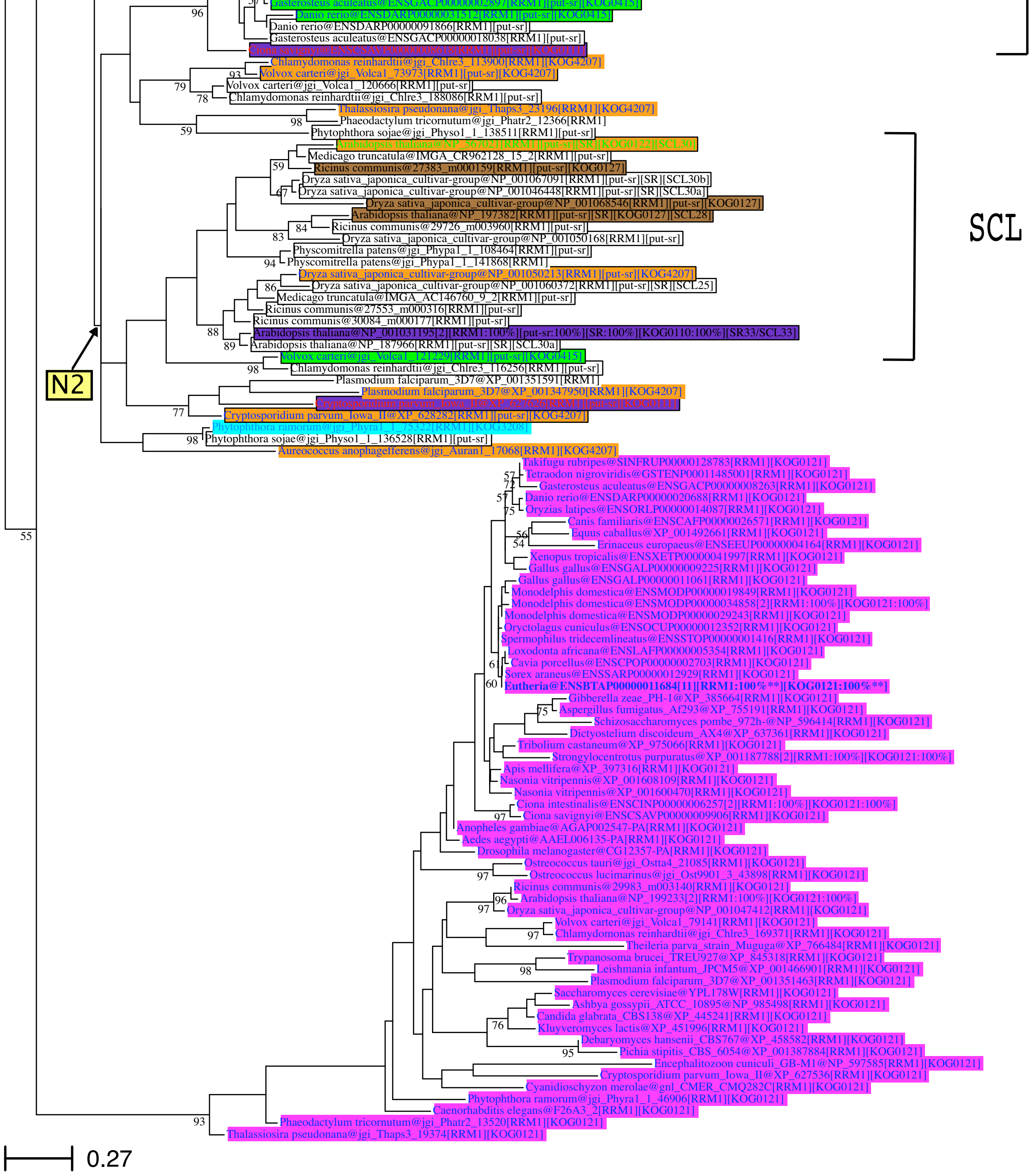


Fig. S14

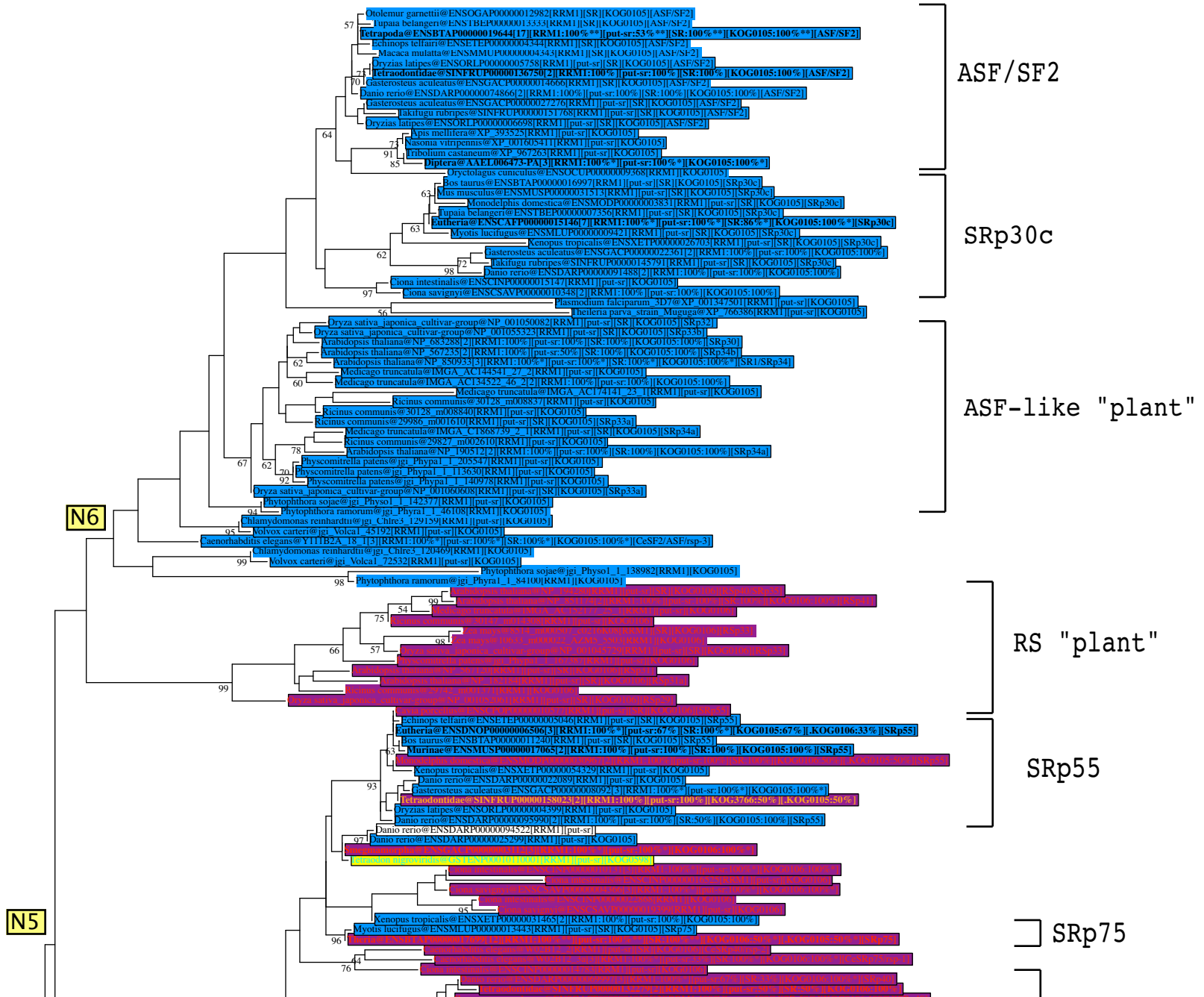

Otolemur garnettii@ENSOGAP00000012982[RRM1][SR][KOG0105][ASF/SF2]
Tupaia belangeri@ENSTBEP00000013333[RRM1][SR][KOG0105][ASF/SF2]
Tetrapoda@ENSBTAP00000019644[17][RRM1:100%*][put-sr:53%**][SR:100%**][KOG0105:100%**][ASF/SF2]
Echinops telfairi@ENSETEP00000043444[RRM1][SR][KOG0105][ASF/SF2]
Macaca mulatta@ENSMMUP00000004343[RRM1][SR][KOG0105][ASF/SF2]
Oryzias latipes@ENSORLP00000005758[RRM1][put-sr][SR][KOG0105][ASF/SF2]
Tetraodontidae@SINFRUP00000136750[2][RRM1:100%*][put-sr:100%*][SR:100%*][KOG0105:100%*][ASF/SF2]
Gasterosteus aculeatus@ENSGACP00000014666[RRM1][SR][KOG0105][ASF/SF2]
Danio rerio@ENSDARP00000074866[2][RRM1:100%][put-sr:100%][SR:100%][KOG0105:100%][ASF/SF2]
Gasterosteus aculeatus@ENSGACP00000027276[RRM1][put-sr][SR][KOG0105][ASF/SF2]
Takifugu rubripes@SINFRUP00000151768[RRM1][put-sr][SR][KOG0105][ASF/SF2]
Oryzias latipes@ENSORLP00000006698[RRM1][put-sr][SR][KOG0105][ASF/SF2]
Apis mellifera@XP_393525[RRM1][put-sr][KOG0105]
Nasonia vitripennis@XP_001605411[RRM1][put-sr][KOG0105]
Tribolium castaneum@XP_967263[RRM1][put-sr][KOG0105]
Diptera@AAEL006473-PA[3][RRM1:100%*][put-sr:100%*][KOG0105:100%*]
Oryctolagus cuniculus@ENSOCUP00000009368[RRM1][KOG0105]
Bos taurus@ENSBTAP00000016997[RRM1][put-sr][SR][KOG0105][SRp30c]
Mus musculus@ENSMUSP00000031513[RRM1][put-sr][SR][KOG0105][SRp30c]
Monodelphis domestica@ENSMODP00000003831[RRM1][put-sr][SR][KOG0105][SRp30c]
Tupaia belangeri@ENSTBEP00000007356[RRM1][put-sr][SR][KOG0105][SRp30c]
Eutheria@ENSCAFP00000015146[7][RRM1:100%*][put-sr:100%*][SR:86%*][KOG0105:100%*][SRp30c]
Myotis lucifugus@ENSMLUP00000009421[RRM1][put-sr][SR][KOG0105][SRp30c]
Xenopus tropicalis@ENSXETP00000026703[RRM1][put-sr][SR][KOG0105][SRp30c]
Gasterosteus aculeatus@ENSGACP00000022361[2][RRM1:100%][put-sr:100%][KOG0105:100%]
Takifugu rubripes@SINFRUP00000145791[RRM1][put-sr][SR][KOG0105][SRp30c]
Danio rerio@ENSDARP00000091488[2][RRM1:100%][put-sr:100%][KOG0105:100%]
Ciona intestinalis@ENSCINP00000015147[RRM1][put-sr][KOG0105]
Ciona savignyi@ENSCSAVP00000010348[2][RRM1:100%][put-sr:100%][KOG0105:100%]
Plasmodium falciparum_3D7@XP_001347501[RRM1][put-sr][KOG0105]
Theileria parva_strain_Muguga@XP_766386[RRM1][put-sr][KOG0105]
Oryza sativa_japonica_cultivar-group@NP_001050082[RRM1][put-sr][SR][KOG0105][SRp32]
Oryza sativa_japonica_cultivar-group@NP_001055323[RRM1][put-sr][SR][KOG0105][SRp33b]
Arabidopsis thaliana@NP_683288[2][RRM1:100%][put-sr:100%][SR:100%][KOG0105:100%][SRp30]
Arabidopsis thaliana@NP_567235[2][RRM1:100%][put-sr:50%][SR:100%][KOG0105:100%][SRp34b]
Arabidopsis thaliana@NP_850933[3][RRM1:100%*][put-sr:100%*][SR:100%*][KOG0105:100%*][SR1/SRp34]
Medicago truncatula@IMGA_AC144541_27_2[RRM1][put-sr][KOG0105]
Medicago truncatula@IMGA_AC134522_46_2[2][RRM1:100%][put-sr:100%][KOG0105:100%]
Medicago truncatula@IMGA_AC174141_23_1[RRM1][put-sr][KOG0105]
Ricinus communis@30128_m008837[RRM1][put-sr][KOG0105]
Ricinus communis@30128_m008840[RRM1][put-sr][KOG0105]
Ricinus communis@29986_m001610[RRM1][put-sr][SR][KOG0105][SRp33a]
Medicago truncatula@IMGA_CT868739_2_1[RRM1][put-sr][SR][KOG0105][SRp34a]
Ricinus communis@29827_m002610[RRM1][put-sr][KOG0105]
Arabidopsis thaliana@NP_190512[2][RRM1:100%][put-sr:100%][SR:100%][KOG0105:100%][SRp34a]
Physcomitrella patens@jgi_Phypa1_1_205547[RRM1][put-sr][KOG0105]
Physcomitrella patens@jgi_Phypa1_1_113630[RRM1][put-sr][KOG0105]
Physcomitrella patens@jgi_Phypa1_1_140978[RRM1][put-sr][KOG0105]
Oryza sativa_japonica_cultivar-group@NP_001060608[RRM1][put-sr][SR][KOG0105][SRp33a]
Phytophthora sojae@jgi_Physo1_1_142377[RRM1][put-sr][KOG0105]
Phytophthora ramorum@jgi_Phyra1_1_46105[RRM1][KOG0105]
Chlamydomonas reinhardtii@jgi_Chlre3_129159[RRM1][put-sr][KOG0105]
Volvox carteri@jgi_Volca1_45192[RRM1][put-sr][KOG0105]
Caenorhabditis elegans@Y11D2A_18_1[3][RRM1:100%*][put-sr:100%*][SR:100%*][KOG0105:100%*][CeSF2/ASF/rsp-3]
Chlamydomonas reinhardtii@jgi_Chlre3_120469[RRM1][KOG0105]
Volvox carteri@jgi_Volca1_72532[RRM1][put-sr][KOG0105]
Phytophthora sojae@jgi_Physo1_1_138982[RRM1][KOG0105]
Phytophthora ramorum@jgi_Phyra1_1_84100[RRM1][KOG0105]
Echinops telfairi@ENSETEP00000050446[RRM1][put-sr][SR][KOG0105][SRp55]
Eutheria@ENSDNOP00000006506[3][RRM1:100%*][put-sr:67%][SR:100%*][KOG0105:67%][.KOG0106:33%][SRp55]
Bos taurus@ENSBTAP00000011240[RRM1][put-sr][SR][KOG0105][SRp55]
Murinae@ENSMUSP00000017065[2][RRM1:100%][put-sr:100%][SR:100%][KOG0105:100%][SRp55]
Xenopus tropicalis@ENSXETP00000054329[RRM1][put-sr][KOG0105]
Danio rerio@ENSDARP00000022089[RRM1][put-sr][KOG0105]
Gasterosteus aculeatus@ENSGACP00000008092[3][RRM1:100%*][put-sr:100%*][KOG0105:100%*]
Tetraodontidae@SINFRUP00000158023[2][RRM1:100%][put-sr:100%][KOG3766:50%][.KOG0105:50%]
Oryzias latipes@ENSORLP00000004399[RRM1][put-sr][KOG0105]
Danio rerio@ENSDARP00000095990[2][RRM1:100%][put-sr:100%][SR:50%][KOG0105:100%][SRp55]
Danio rerio@ENSDARP00000094522[RRM1][put-sr]
Danio rerio@ENSDARP00000025299[RRM1][put-sr][KOG0105]
Xenopus tropicalis@ENSXETP00000031465[2][RRM1:100%][put-sr:100%][KOG0105:100%]
Myotis lucifugus@ENSMLUP00000013443[RRM1][put-sr][SR][KOG0105][SRp75]
ASF/SF2
SRp30c
ASF-like "plant"
RS "plant"
SRp55
SRp75
N6
N5

Fig. S14

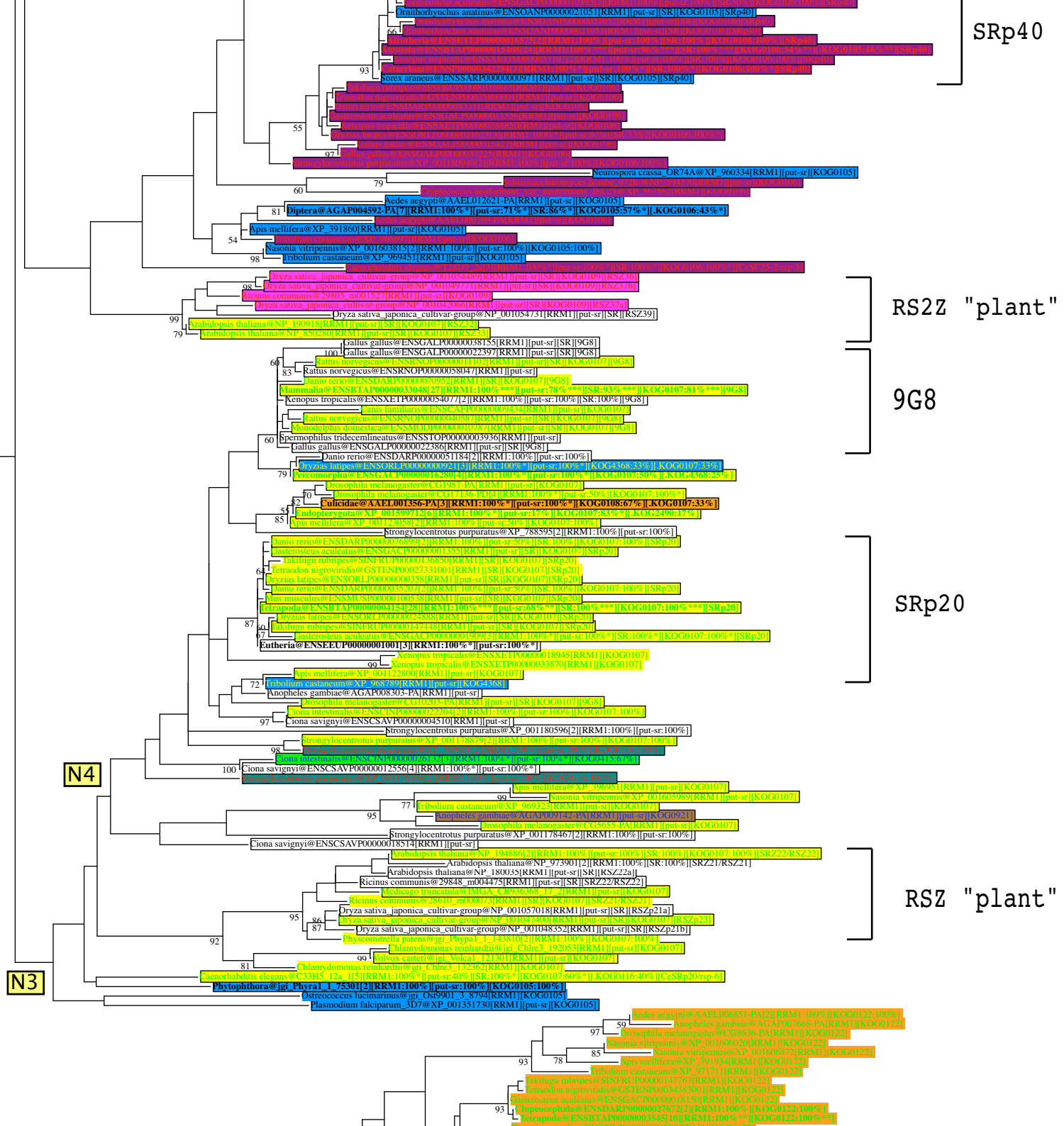

SRp40
RS2Z "plant"
9G8
SRp20
RSZ "plant"
N4
N3
Ornithorhynchus anatinus@ENSOANP00000021051[RRM1][put-sr][SR][KOG0105][SRp40]
Sorex araneus@ENSSARP00000000971[RRM1][put-sr][SR][KOG0105][SRp40]
Neurospora crassa_OR74A@XP_960334[RRM1][put-sr][KOG0105]
Aedes aegypti@AAEL012621-PA[RRM1][put-sr][KOG0105]
Diptera@AGAP004592-PA[7][RRM1:100%*][put-sr:71%*][SR:86%*][KOG0105:57%*][.KOG0106:43%*]
Apis mellifera@XP_391860[RRM1][put-sr][KOG0105]
Nasonia vitripennis@XP_001603815[2][RRM1:100%*][put-sr:100%*][KOG0105:100%]
Tribolium castaneum@XP_969451[RRM1][put-sr][KOG0105]
Oryza sativa_japonica_cultivar-group@NP_001054489[RRM1][put-sr][SR][KOG0109][RSZ36]
Oryza sativa_japonica_cultivar-group@NP_001049777[RRM1][put-sr][SR][KOG0109][RSZ34/35]
Ricinus communis@29805_m001527[RRM1][put-sr][KOG0109]
Oryza sativa_japonica_cultivar-group@NP_001042066[RRM1][put-sr][SR][KOG0109][RSZ37a]
Oryza sativa_japonica_cultivar-group@NP_001054731[RRM1][put-sr][SR][RSZ39]
Arabidopsis thaliana@NP_190918[RRM1][put-sr][SR][KOG0107][RSZ32]
Arabidopsis thaliana@NP_850280[RRM1][put-sr][SR][KOG0107][RSZ33]
Gallus gallus@ENSGALP00000038155[RRM1][put-sr][SR][9G8]
Gallus gallus@ENSGALP00000022397[RRM1][put-sr][SR][9G8]
Rattus norvegicus@ENSRNOP00000011102[RRM1][put-sr][SR][KOG0107][9G8]
Rattus norvegicus@ENSRNOP00000058047[RRM1][put-sr]
Danio rerio@ENSDARP00000070952[RRM1][SR][KOG0107][9G8]
Mammalia@ENSBTAP00000033048[27][RRM1:100%***][put-sr:78%***][SR:93%***][KOG0107:81%***][9G8]
Xenopus tropicalis@ENSXETP00000054077[2][RRM1:100%][put-sr:100%][SR:100%][9G8]
Canis familiaris@ENSCAFP00000009434[RRM1][put-sr][KOG0107]
Rattus norvegicus@ENSRNOP00000040587[RRM1][put-sr][SR][KOG0107][9G8]
Monodelphis domestica@ENSMODP00000010787[RRM1][put-sr][SR][KOG0107][9G8]
Spermophilus tridecemlineatus@ENSSTOP00000003936[RRM1][put-sr]
Gallus gallus@ENSGALP00000022386[RRM1][put-sr][SR][9G8]
Danio rerio@ENSDARP00000051184[2][RRM1:100%][put-sr:100%]
Oryzias latipes@ENSORLP00000000921[5][RRM1:100%*][put-sr:100%*][KOG4368:33%*][.KOG0107:33%*]
Percomorpha@ENSGACP00000016280[4][RRM1:100%*][put-sr:100%*][KOG0107:50%*][.KOG4368:25%*]
Drosophila melanogaster@CG1987-PA[RRM1][put-sr][KOG0107]
Drosophila melanogaster@CG17136-PD[4][RRM1:100%*][put-sr:50%][KOG0107:100%*]
Culicidae@AAEL001356-PA[3][RRM1:100%*][put-sr:100%*][KOG0108:67%*][.KOG0107:33%*]
Endopterygota@XP_001597126[6][RRM1:100%*][put-sr:17%][KOG0107:83%*][.KOG2490:17%]
Apis mellifera@XP_001123058[2][RRM1:100%][put-sr:50%][KOG0107:100%]
Strongylocentrotus purpuratus@XP_788595[2][RRM1:100%][put-sr:100%]
Danio rerio@ENSDARP00000076899[2][RRM1:100%][put-sr:50%][SR:100%][KOG0107:100%][SRp20]
Gasterosteus aculeatus@ENSGACP00000001355[RRM1][put-sr][SR][KOG0107][SRp20]
Takifugu rubripes@SINFRUP00000136850[RRM1][SR][KOG0107][SRp20]
Tetraodon nigroviridis@GSTENP00027331001[RRM1][SR][KOG0107][SRp20]
Oryzias latipes@ENSORLP00000000358[RRM1][put-sr][SR][KOG0107][SRp20]
Danio rerio@ENSDARP00000035207[2][RRM1:100%][put-sr:50%][SR:100%][KOG0107:100%][SRp20]
Mus musculus@ENSMUSP00000100538[RRM1][put-sr][SR][KOG0107][SRp20]
Tetrapoda@ENSBTAP00000004154[28][RRM1:100%***][put-sr:68%**][SR:100%***][KOG0107:100%***][SRp20]
Oryzias latipes@ENSORLP00000024888[RRM1][put-sr][SR][KOG0107][SRp20]
Takifugu rubripes@SINFRUP00000147448[RRM1][put-sr][SR][KOG0107][SRp20]
Gasterosteus aculeatus@ENSGACP00000010093[RRM1:100%*][put-sr:100%*][SR:100%*][KOG0107:100%*][SRp20]
Eutheria@ENSEEUP00000001001[3][RRM1:100%*][put-sr:100%*]
Xenopus tropicalis@ENSXALP00000018945[RRM1][KOG0107]
Xenopus tropicalis@ENSXETP00000033870[RRM1][KOG0107]
Apis mellifera@XP_001122800[RRM1][put-sr][KOG0107]
Tribolium castaneum@XP_968759[RRM1][put-sr][KOG4368]
Anopheles gambiae@AGAP008303-PA[RRM1][put-sr]
Drosophila melanogaster@CG10203-PA[RRM1][put-sr][SR][KOG0107][9G8]
Ciona intestinalis@ENSCINP00000022004[2][RRM1:100%][put-sr:100%][KOG0107:100%]
Ciona savignyi@ENSCSAVP00000004510[RRM1][put-sr]
Strongylocentrotus purpuratus@XP_001180596[2][RRM1:100%][put-sr:100%]
Strongylocentrotus purpuratus@XP_001178879[2][RRM1:100%][put-sr:100%][KOG0107:100%]
Ciona intestinalis@ENSCINP00000012172[3][RRM1:100%*][put-sr:100%*][KOG0115:67%]
Ciona savignyi@ENSCSAVP00000012556[4][RRM1:100%*][put-sr:100%*]
Apis mellifera@XP_396951[RRM1][put-sr][KOG0107]
Nasonia vitripennis@XP_001605989[RRM1][put-sr][KOG0107]
Tribolium castaneum@XP_969323[RRM1][put-sr][KOG0107]
Anopheles gambiae@AGAP009142-PA[RRM1][put-sr][KOG0921]
Drosophila melanogaster@CG5655-PA[RRM1][put-sr][KOG0107]
Strongylocentrotus purpuratus@XP_001178467[2][RRM1:100%][put-sr:100%]
Ciona savignyi@ENSCSAVP00000018514[RRM1][put-sr]
Arabidopsis thaliana@NP_194886[2][RRM1:100%][put-sr:100%][SR:100%][KOG0107:100%][SRZ22/RSZ22]
Arabidopsis thaliana@NP_973901[2][RRM1:100%][SR:100%][SRZ21/RSZ21]
Arabidopsis thaliana@NP_180035[RRM1][put-sr][SR][RSZ22a]
Ricinus communis@29848_m004475[RRM1][put-sr][SR][SRZ22/RSZ22]
Medicago truncatula@IMGA_CR936368_T1_2[RRM1][put-sr][KOG0107]
Ricinus communis@28610_m000073[RRM1][SR][KOG0107][SRZ21/RSZ21]
Oryza sativa_japonica_cultivar-group@NP_001057018[RRM1][put-sr][SR][RSZp21a]
Oryza sativa_japonica_cultivar-group@NP_001047400[RRM1][put-sr][SR][KOG0107][RSZp23]
Oryza sativa_japonica_cultivar-group@NP_001048352[RRM1][put-sr][SR][RSZp21b]
Physcomitrella patens@jgi_Phypa1_1_143810[2][RRM1:100%][KOG0107:100%]
Chlamydomonas reinhardtii@jgi_Chlre3_192053[RRM1][put-sr][KOG0107]
Volvox carteri@jgi_Volca1_121301[RRM1][put-sr][KOG0107]
Chlamydomonas reinhardtii@jgi_Chlre3_132362[RRM1][KOG0107]
Caenorhabditis elegans@C33H5_12a_1[5][RRM1:100%*][put-sr:40%][SR:100%*][KOG0107:60%*][.KOG0116:40%][CeSRp20/rsp-6]
Phytophthora@jgi_Phyra1_1_75301[2][RRM1:100%][put-sr:100%][KOG0105:100%]
Ostreococcus lucimarinus@jgi_Ost9901_3_8794[RRM1][KOG0105]
Plasmodium falciparum_3D7@XP_001351730[RRM1][put-sr][KOG0105]
Aedes aegypti@AAEL006851-PA[2][RRM1:100%][KOG0122:100%]
Anopheles gambiae@AGAP007668-PA[RRM1][KOG0122]
Drosophila melanogaster@CG8636-PA[RRM1][KOG0122]
Nasonia vitripennis@XP_001606020[RRM1][KOG0122]
Nasonia vitripennis@XP_001606872[RRM1][KOG0122]
Apis mellifera@XP_391934[RRM1][KOG0122]
Tribolium castaneum@XP_971711[RRM1][KOG0122]
Takifugu rubripes@SINFRUP00001476[3][RRM1][KOG0122]
Tetraodon nigroviridis@GSTENP00034383001[RRM1][KOG0122]
Gasterosteus aculeatus@ENSGACP00000018150[RRM1][KOG0122]
Clupeocephala@ENSDARP00000027672[2][RRM1:100%][KOG0122:100%]
Tetrapoda@ENSBTAP00000003545[10][RRM1:100%**][KOG0122:100%**]

Fig. S14

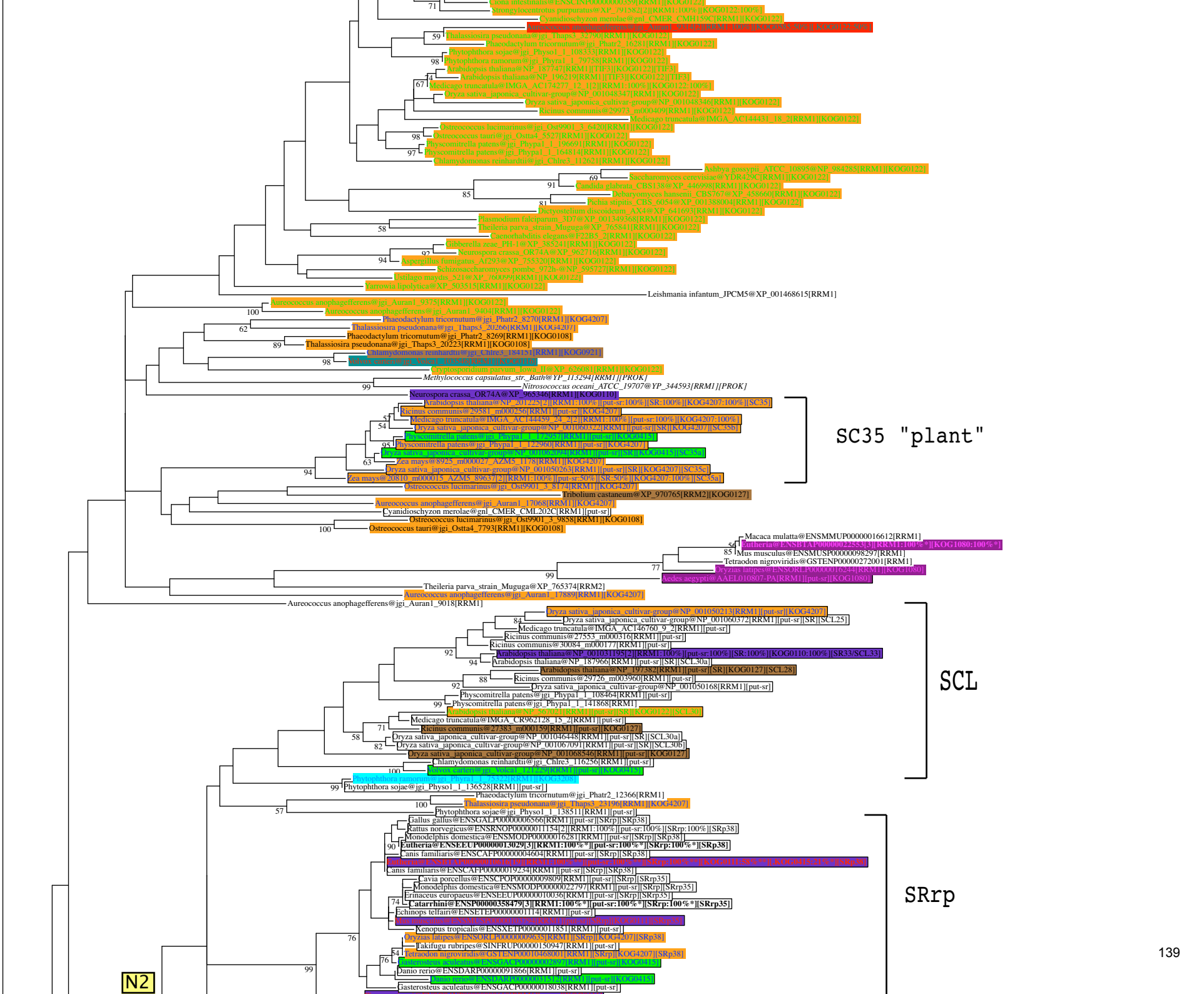

Leishmania infantum_JPCM5@XP_001468615[RRM1]
Methylococcus capsulatus_str._Bath@YP_113294[RRM1][PROK]
Nitrosococcus oceani_ATCC_19707@YP_344593[RRM1][PROK]
SC35 "plant"
SCL
SRrp
N2

Fig. S14

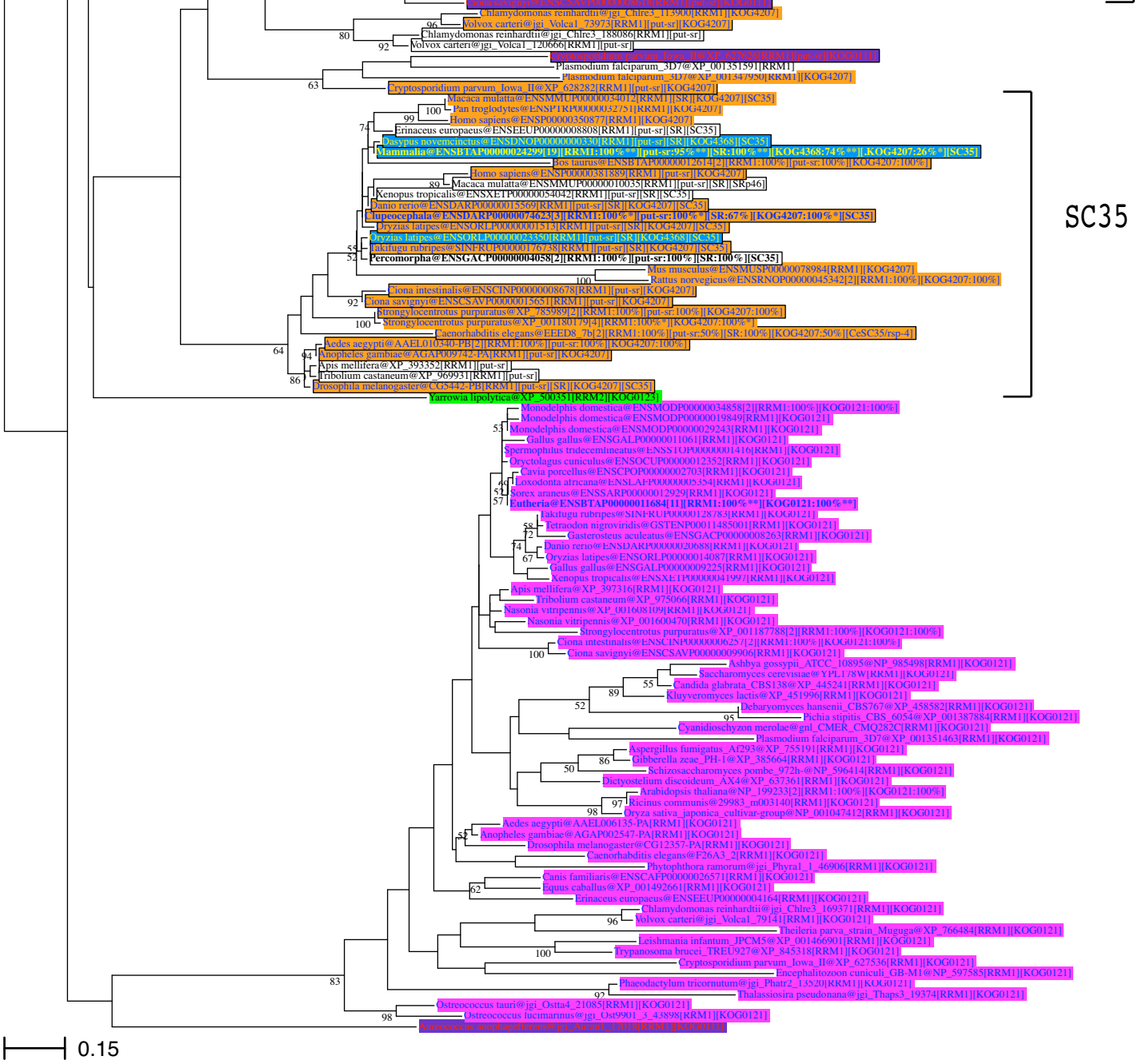
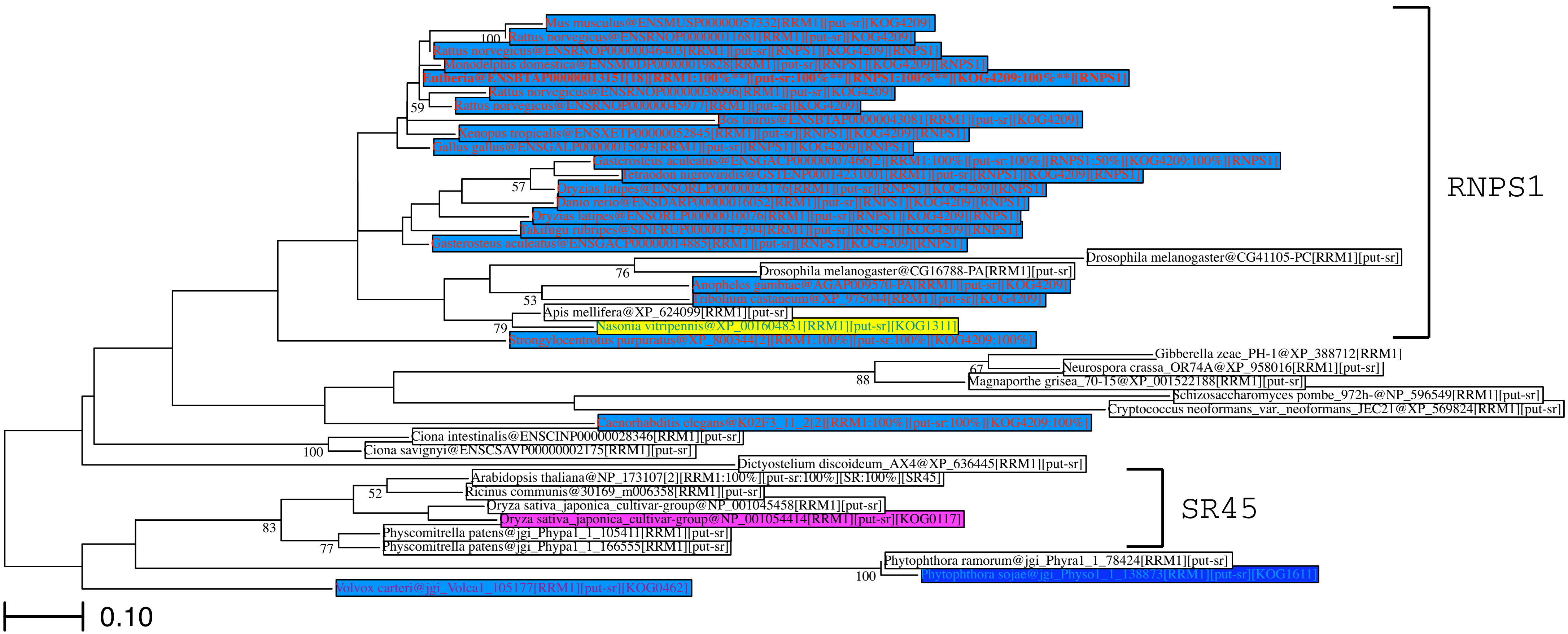

Chlamydomonas reinhardtii@jgi_Chlre3_113900[RRM1][KOG4207]
Volvox carteri@jgi_Volca1_73973[RRM1][put-sr][KOG4207]
Chlamydomonas reinhardtii@jgi_Chlre3_188086[RRM1][put-sr]
Volvox carteri@jgi_Volca1_120666[RRM1][put-sr]
Plasmodium falciparum_3D7@XP_001351591[RRM1]
Plasmodium falciparum_3D7@XP_001347930[RRM1][KOG4207]
Cryptosporidium parvum_Iowa_II@XP_628282[RRM1][put-sr][KOG4207]
Macaca mulatta@ENSMMUP00000034012[RRM1][SR][KOG4207][SC35]
Pan troglodytes@ENSPTRP00000032751[RRM1][KOG4207]
Homo sapiens@ENSP00000350877[RRM1][KOG4207]
Erinaceus europaeus@ENSEEUP00000008808[RRM1][put-sr][SR][SC35]
Dasypus novemcinctus@ENSDNOP00000000330[RRM1][put-sr][SR][KOG4368][SC35]
Mammalia@ENSBTAP00000024299[19][RRM1:100%**][put-sr:95%**][SR:100%**][KOG4368:74%**][KOG4207:26%*][SC35]
Bos taurus@ENSBTAP00000012614[2][RRM1:100%][put-sr:100%][KOG4207:100%]
Homo sapiens@ENSP00000381889[RRM1][put-sr][KOG4207]
Macaca mulatta@ENSMMUP00000010035[RRM1][put-sr][SR][SRp46]
Xenopus tropicalis@ENSXETP00000054042[RRM1][put-sr][SR][SC35]
Danio rerio@ENSDARP00000015569[RRM1][put-sr][SR][KOG4207][SC35]
Clupeocephala@ENSDARP00000074623[3][RRM1:100%**][put-sr:100%**][SR:67%][KOG4207:100%*][SC35]
Oryzias latipes@ENSORLP00000001513[RRM1][put-sr][SR][KOG4207][SC35]
Oryzias latipes@ENSORLP00000023350[RRM1][put-sr][SR][KOG4368][SC35]
Takifugu rubripes@SINFRUP00000176238[RRM1][put-sr][SR][KOG4207][SC35]
Percomorpha@ENSGACP00000004058[2][RRM1:100%][put-sr:100%][SR:100%][SC35]
Mus musculus@ENSMUSP00000078984[RRM1][KOG4207]
Rattus norvegicus@ENSRNOP00000045342[2][RRM1:100%][KOG4207:100%]
Ciona intestinalis@ENSCINP00000008678[RRM1][put-sr][KOG4207]
Ciona savignyi@ENSCSAVP00000015651[RRM1][put-sr][KOG4207]
Strongylocentrotus purpuratus@XP_785989[2][RRM1:100%][put-sr:100%][KOG4207:100%]
Strongylocentrotus purpuratus@XP_001180179[4][RRM1:100%*][KOG4207:100%*]
Caenorhabditis elegans@EEED8_7b[2][RRM1:100%][put-sr:50%][SR:100%][KOG4207:50%][CeSC35/rsp-4]
Aedes aegypti@AAEL010340-PB[2][RRM1:100%][put-sr:100%][KOG4207:100%]
Anopheles gambiae@AGAP009742-PA[RRM1][put-sr][KOG4207]
Apis mellifera@XP_393352[RRM1][put-sr]
Tribolium castaneum@XP_969931[RRM1][put-sr]
Drosophila melanogaster@CG5442-PB[RRM1][put-sr][SR][KOG4207][SC35]
Yarrowia lipolytica@XP_500351[RRM2][KOG0123]
SC35
Monodelphis domestica@ENSMODP00000034858[2][RRM1:100%][KOG0121:100%]
Monodelphis domestica@ENSMODP00000019849[RRM1][KOG0121]
Monodelphis domestica@ENSMODP00000029243[RRM1][KOG0121]
Gallus gallus@ENSGALP00000011061[RRM1][KOG0121]
Spermophilus tridecemlineatus@ENSSTOP00000001416[RRM1][KOG0121]
Oryctolagus cuniculus@ENSOCUP00000012352[RRM1][KOG0121]
Cavia porcellus@ENSCPOP00000002703[RRM1][KOG0121]
Loxodonta africana@ENSLAFP00000005354[RRM1][KOG0121]
Sorex araneus@ENSSARP00000012929[RRM1][KOG0121]
Eutheria@ENSBTAP00000011684[11][RRM1:100%**][KOG0121:100%**]
Takifugu rubripes@SINFRUP00000128783[RRM1][KOG0121]
Tetraodon nigroviridis@GSTENP00011485001[RRM1][KOG0121]
Gasterosteus aculeatus@ENSGACP00000008263[RRM1][KOG0121]
Danio rerio@ENSDARP00000020688[RRM1][KOG0121]
Oryzias latipes@ENSORLP00000014087[RRM1][KOG0121]
Gallus gallus@ENSGALP00000009225[RRM1][KOG0121]
Xenopus tropicalis@ENSXETP00000041997[RRM1][KOG0121]
Apis mellifera@XP_397316[RRM1][KOG0121]
Tribolium castaneum@XP_975066[RRM1][KOG0121]
Nasonia vitripennis@XP_001608109[RRM1][KOG0121]
Nasonia vitripennis@XP_001600470[RRM1][KOG0121]
Strongylocentrotus purpuratus@XP_001187788[2][RRM1:100%][KOG0121:100%]
Ciona intestinalis@ENSCINP00000006257[2][RRM1:100%][KOG0121:100%]
Ciona savignyi@ENSCSAVP00000009906[RRM1][KOG0121]
Ashbya gossypii_ATCC_10895@NP_985498[RRM1][KOG0121]
Saccharomyces cerevisiae@YPL178W[RRM1][KOG0121]
Candida glabrata_CBS138@XP_445241[RRM1][KOG0121]
Kluyveromyces lactis@XP_451996[RRM1][KOG0121]
Debaryomyces hansenii_CBS767@XP_458582[RRM1][KOG0121]
Pichia stipitis_CBS_6054@XP_001387884[RRM1][KOG0121]
Cyanidioschyzon merolae@gnl_CMER_CMQ282C[RRM1][KOG0121]
Plasmodium falciparum_3D7@XP_001351463[RRM1][KOG0121]
Aspergillus fumigatus_Af293@XP_755191[RRM1][KOG0121]
Gibberella zeae_PH-1@XP_385664[RRM1][KOG0121]
Schizosaccharomyces pombe_972h-@NP_596414[RRM1][KOG0121]
Dictyostelium discoideum_AX4@XP_637361[RRM1][KOG0121]
Arabidopsis thaliana@NP_199233[2][RRM1:100%][KOG0121:100%]
Ricinus communis@29983_m003140[RRM1][KOG0121]
Oryza sativa_japonica_cultivar-group@NP_001047412[RRM1][KOG0121]
Aedes aegypti@AAEL006135-PA[RRM1][KOG0121]
Anopheles gambiae@AGAP002547-PA[RRM1][KOG0121]
Drosophila melanogaster@CG12357-PA[RRM1][KOG0121]
Caenorhabditis elegans@F26A3_2[RRM1][KOG0121]
Phytophthora ramorum@jgi_Phyra1_1_46906[RRM1][KOG0121]
Canis familiaris@ENSCAFP00000026571[RRM1][KOG0121]
Equus caballus@XP_001492661[RRM1][KOG0121]
Erinaceus europaeus@ENSEEUP00000004164[RRM1][KOG0121]
Chlamydomonas reinhardtii@jgi_Chlre3_169371[RRM1][KOG0121]
Volvox carteri@jgi_Volca1_79141[RRM1][KOG0121]
Theileria parva_strain_Muguga@XP_766484[RRM1][KOG0121]
Leishmania infantum_JPCM5@XP_001466901[RRM1][KOG0121]
Trypanosoma brucei_TREU927@XP_845318[RRM1][KOG0121]
Cryptosporidium parvum_Iowa_II@XP_627536[RRM1][KOG0121]
Encephalitozoon cuniculi_GB-M1@NP_597585[RRM1][KOG0121]
Phaeodactylum tricornutum@jgi_Phatr2_11520[RRM1][KOG0121]
Thalassiosira pseudonana@jgi_Thaps3_19374[RRM1][KOG0121]
Ostreococcus tauri@jgi_Ostta4_21085[RRM1][KOG0121]
Ostreococcus lucimarinus@jgi_Ost9901_3_43898[RRM1][KOG0121]
0.15

Fig. S15



Mus musculus@ENSMUSP00000057332[RRM1][put-sr][KOG4209]
100
Rattus norvegicus@ENSRNOP00000011681[RRM1][put-sr][KOG4209]
Rattus norvegicus@ENSRNOP00000046403[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Monodelphis domestica@ENSMODP00000019828[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Eutheria@ENSBTAP00000013151[18][RRM1:100%**][put-sr:100%**][RNPS1:100%**][KOG4209:100%**][RNPS1]
Rattus norvegicus@ENSRNOP00000038996[RRM1][put-sr][KOG4209]
59
Rattus norvegicus@ENSRNOP00000045977[RRM1][put-sr][KOG4209]
Bos taurus@ENSBTAP00000043081[RRM1][put-sr][KOG4209]
Xenopus tropicalis@ENSXETP00000052845[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Gallus gallus@ENSGALP00000013093[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Gasterosteus aculeatus@ENSGACP00000007466[2][RRM1:100%][put-sr:100%][RNPS1:50%][KOG4209:100%][RNPS1]
Tetraodon nigroviridis@GSTENP00014231001[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
57
Oryzias latipes@ENSORLP00000023176[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Danio rerio@ENSDARP00000016052[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Oryzias latipes@ENSORLP00000010076[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Takifugu rubripes@SINFRUP00000147394[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Gasterosteus aculeatus@ENSGACP00000014885[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Drosophila melanogaster@CG41105-PC[RRM1][put-sr]
76
Drosophila melanogaster@CG16788-PA[RRM1][put-sr]
Anopheles gambiae@AGAP009570-PA[RRM1][put-sr][KOG4209]
53
Tribolium castaneum@XP_975044[RRM1][put-sr][KOG4209]
Apis mellifera@XP_624099[RRM1][put-sr]
79
Nasonia vitripennis@XP_001604831[RRM1][put-sr][KOG1311]
Strongylocentrotus purpuratus@XP_800344[2][RRM1:100%][put-sr:100%][KOG4209:100%]
RNPS1
Gibberella zeae_PH-1@XP_388712[RRM1]
67
Neurospora crassa_OR74A@XP_958016[RRM1][put-sr]
88
Magnaporthe grisea_70-15@XP_001522188[RRM1][put-sr]
Schizosaccharomyces pombe_972h-@NP_596549[RRM1][put-sr]
Cryptococcus neoformans_var._neoformans_JEC21@XP_569824[RRM1][put-sr]
Caenorhabditis elegans@K02F3_11_2[2][RRM1:100%][put-sr:100%][KOG4209:100%]
Ciona intestinalis@ENSCINP00000028346[RRM1][put-sr]
100
Ciona savignyi@ENSCSAVP00000002175[RRM1][put-sr]
Dictyostelium discoideum_AX4@XP_636445[RRM1][put-sr]
Arabidopsis thaliana@NP_173107[2][RRM1:100%][put-sr:100%][SR:100%][SR45]
52
Ricinus communis@30169_m006358[RRM1][put-sr]
Oryza sativa_japonica_cultivar-group@NP_001045458[RRM1][put-sr]
Oryza sativa_japonica_cultivar-group@NP_001054414[RRM1][put-sr][KOG0117]
83
Physcomitrella patens@jgi_Phypa1_1_105411[RRM1][put-sr]
77
Physcomitrella patens@jgi_Phypa1_1_166555[RRM1][put-sr]
SR45
Phytophthora ramorum@jgi_Phyra1_1_78424[RRM1][put-sr]
100
Phytophthora sojae@jgi_Physo1_1_138873[RRM1][put-sr][KOG1611]
Volvox carteri@jgi_Volca1_105177[RRM1][put-sr][KOG0462]
0.10

Fig. S16

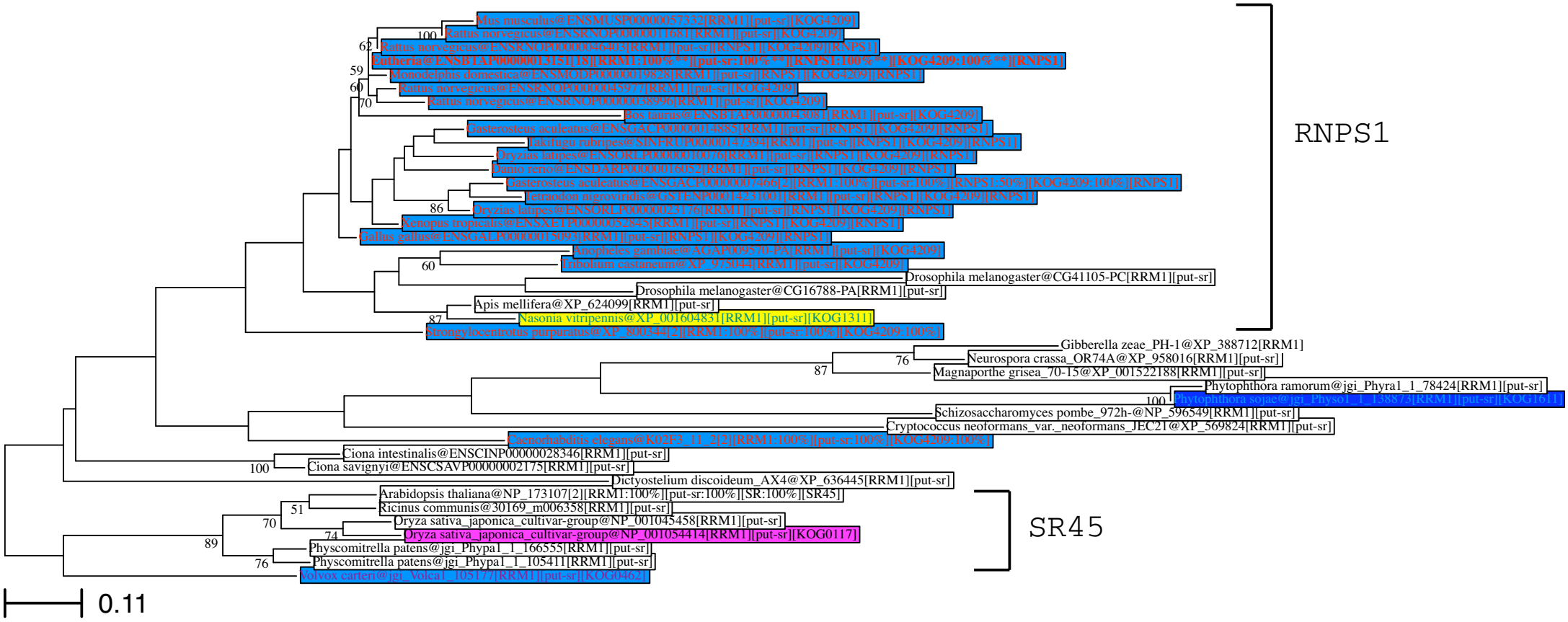

Mus musculus@ENSMUSP00000057332[RRM1][put-sr][KOG4209]
Rattus norvegicus@ENSRNOP00000011663[RRM1][put-sr][KOG4209]
Rattus norvegicus@ENSRNOP00000046403[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Gallaeria@ENSETAP00000015118[RRM1:100%**][put-sr:100%**][RNPS1:100%**][KOG4209:100%**][RNPS1]
Monodelphis domestica@ENSMODP00000010838[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Rattus norvegicus@ENSRNOP00000045977[RRM1][put-sr][KOG4209]
Rattus norvegicus@ENSRNOP00000039990[RRM1][put-sr][KOG4209]
Bos taurus@ENSBTAP00000043985[RRM1][put-sr][KOG4209]
Gasterosteus aculeatus@ENSGACP00000014885[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Maylandia zebra@ENSMZEP00000147394[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Oryzias latipes@ENSORLP00000010076[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Danio rerio@ENSDARP00000016062[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Gasterosteus aculeatus@ENSGACP00000026624[RRM1:100%][put-sr:100%][RNPS1:50%][KOG4209:100%][RNPS1]
Tetraodon nigroviridis@GSTENP00012310001[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Oryzias latipes@ENSORLP00000023176[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Xenopus tropicalis@ENSXETP00000025825[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Gallus gallus@ENSGALP00000015693[RRM1][put-sr][RNPS1][KOG4209][RNPS1]
Anopheles gambiae@AGAP009570-PA[RRM1][put-sr][KOG4209]
Tribolium castaneum@XP_975041[RRM1][put-sr][KOG4209]
Drosophila melanogaster@CG41105-PC[RRM1][put-sr]
Drosophila melanogaster@CG16788-PA[RRM1][put-sr]
Apis mellifera@XP_624099[RRM1][put-sr]
Nasonia vitripennis@XP_001604831[RRM1][put-sr][KOG1311]
Strongylocentrotus purpuratus@XP_800344[2][RRM1:100%][put-sr:100%][KOG4209:100%]
RNPS1
Gibberella zeae_PH-1@XP_388712[RRM1]
Neurospora crassa_OR74A@XP_958016[RRM1][put-sr]
Magnaporthe grisea_70-15@XP_001522188[RRM1][put-sr]
Phytophthora ramorum@jgi_Phyra1_1_78424[RRM1][put-sr]
Phytophthora sojae@jgi_Physo1_1_138875[RRM1][put-sr][KOG1611]
Schizosaccharomyces pombe_972h-@NP_596549[RRM1][put-sr]
Cryptococcus neoformans_var._neoformans_JEC21@XP_569824[RRM1][put-sr]
Caenorhabditis elegans@K02F3_11_2[2][RRM1:100%][put-sr:100%][KOG4209:100%]
Ciona intestinalis@ENSCINP00000028346[RRM1][put-sr]
Ciona savignyi@ENSCSAVP00000002175[RRM1][put-sr]
Dictyostelium discoideum_AX4@XP_636445[RRM1][put-sr]
Arabidopsis thaliana@NP_173107[2][RRM1:100%][put-sr:100%][SR:100%][SR45]
Ricinus communis@30169_m006358[RRM1][put-sr]
Oryza sativa_japonica_cultivar-group@NP_001045458[RRM1][put-sr]
Oryza sativa_japonica_cultivar-group@NP_001054414[RRM1][put-sr][KOG0117]
Physcomitrella patens@jgi_Phypa1_1_166555[RRM1][put-sr]
Physcomitrella patens@jgi_Phypa1_1_105411[RRM1][put-sr]
Volvox carteri@jgi_Volca1_105177[RRM1][put-sr][KOG0462]
SR45
100
62
59
60
70
86
60
87
76
87
100
51
70
74
89
76
100
0.11

Fig. S17

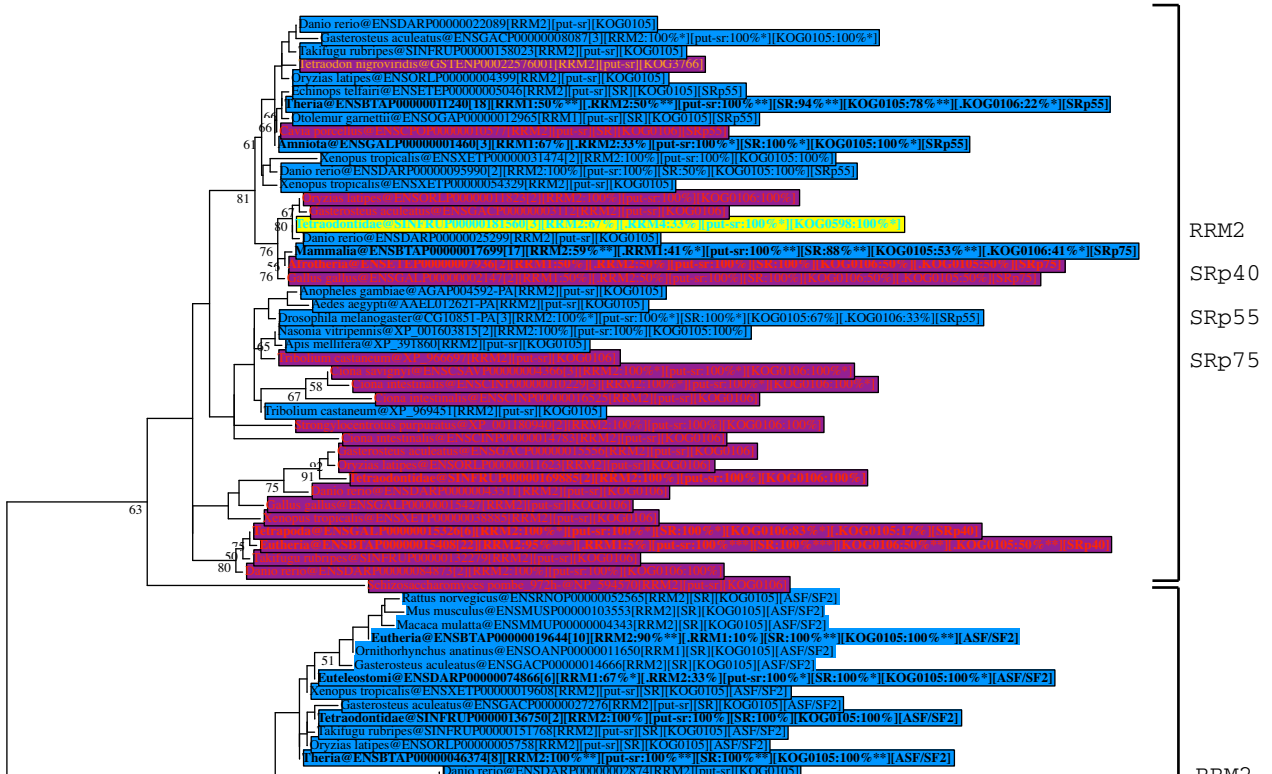

Danio rerio@ENSDARP00000022089[RRM2][put-sr][KOG0105]
Gasterosteus aculeatus@ENSGACP00000008087[3][RRM2:100%*][put-sr:100%*][KOG0105:100%*]
Takifugu rubripes@SINFRUP00000158023[RRM2][put-sr][KOG0105]
Tetraodon nigroviridis@GSTENP00022570001[RRM2][put-sr][KOG3766]
Oryzias latipes@ENSORLP00000004399[RRM2][put-sr][KOG0105]
Echinops telfairi@ENSETEP00000005046[RRM2][put-sr][SR][KOG0105][SRp55]
Theria@ENSBTAP00000011240[18][RRM1:50%**][.RRM2:50%**][put-sr:100%**][SR:94%**][KOG0105:78%**][.KOG0106:22%*][SRp55]
Otolemur garnettii@ENSOGAP00000012965[RRM1][put-sr][SR][KOG0105][SRp55]
Amniota@ENSGALP00000001460[3][RRM1:67%][.RRM2:33%][put-sr:100%*][SR:100%*][KOG0105:100%*][SRp55]
Xenopus tropicalis@ENSXETP00000031474[2][RRM2:100%][put-sr:100%][KOG0105:100%]
Danio rerio@ENSDARP00000095990[2][RRM2:100%][put-sr:100%][SR:50%][KOG0105:100%][SRp55]
Xenopus tropicalis@ENSXETP00000054329[RRM2][put-sr][KOG0105]
Tetraodontidae@SINFRUP00000183607[3][RRM2:67%][.RRM2:33%][put-sr:100%*][KOG0598:100%*]
Danio rerio@ENSDARP00000025299[RRM2][put-sr][KOG0105]
Mammalia@ENSBTAP00000017699[17][RRM2:59%**][.RRM1:41%*][put-sr:100%**][SR:88%**][KOG0105:53%**][.KOG0106:41%*][SRp75]
Anopheles gambiae@AGAP004592-PA[RRM2][put-sr][KOG0105]
Aedes aegypti@AAEL012621-PA[RRM2][put-sr][KOG0105]
Drosophila melanogaster@CG10851-PA[3][RRM2:100%*][put-sr:100%*][SR:100%*][KOG0105:67%][.KOG0106:33%][SRp55]
Nasonia vitripennis@XP_001603815[2][RRM2:100%][put-sr:100%][KOG0105:100%]
Apis mellifera@XP_397860[RRM2][put-sr][KOG0105]
Tribolium castaneum@XP_969451[RRM2][put-sr][KOG0105]
Rattus norvegicus@ENSRNOP00000052565[RRM2][SR][KOG0105][ASF/SF2]
Mus musculus@ENSMUSP00000103553[RRM2][SR][KOG0105][ASF/SF2]
Macaca mulatta@ENSMMUP00000004343[RRM2][SR][KOG0105][ASF/SF2]
Eutheria@ENSBTAP00000019644[10][RRM2:90%**][.RRM1:10%][SR:100%**][KOG0105:100%**][ASF/SF2]
Ornithorhynchus anatinus@ENSOANP00000011650[RRM1][SR][KOG0105][ASF/SF2]
Gasterosteus aculeatus@ENSGACP00000014666[RRM2][SR][KOG0105][ASF/SF2]
Euteleostomi@ENSDARP00000074866[6][RRM1:67%*][.RRM2:33%][put-sr:100%*][SR:100%*][KOG0105:100%*][ASF/SF2]
Xenopus tropicalis@ENSXETP00000019608[RRM2][put-sr][SR][KOG0105][ASF/SF2]
Gasterosteus aculeatus@ENSGACP00000027276[RRM2][put-sr][SR][KOG0105][ASF/SF2]
Tetraodontidae@SINFRUP00000136750[2][RRM2:100%][put-sr:100%][SR:100%][KOG0105:100%][ASF/SF2]
Takifugu rubripes@SINFRUP00000151768[RRM2][put-sr][SR][KOG0105][ASF/SF2]
Oryzias latipes@ENSORLP00000057758[RRM2][put-sr][SR][KOG0105][ASF/SF2]
Theria@ENSBTAP00000046374[8][RRM2:100%**][put-sr:100%**][SR:100%**][KOG0105:100%**][ASF/SF2]
RRM2
SRp40
SRp55
SRp75

Fig. S17

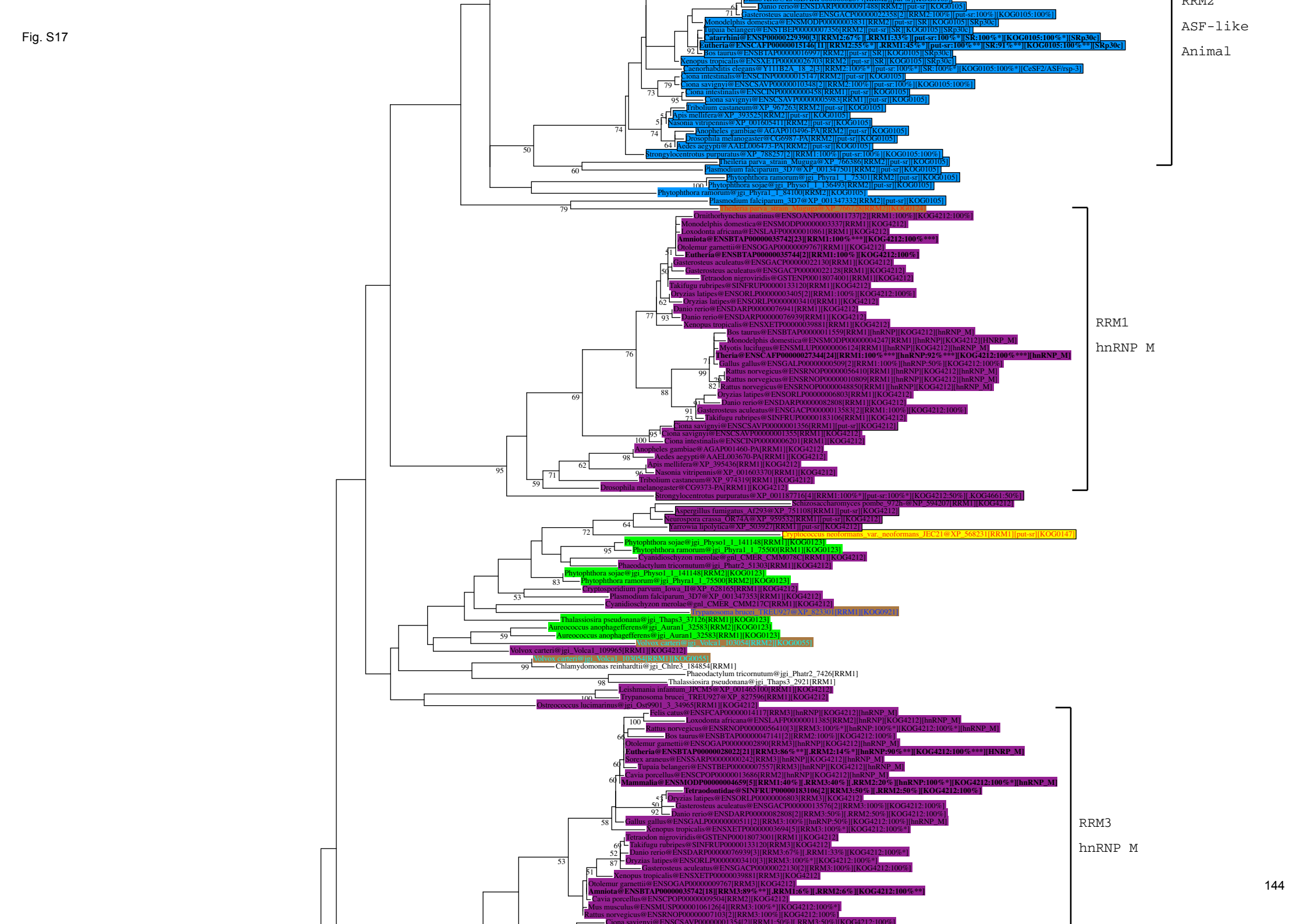


Fig. S17

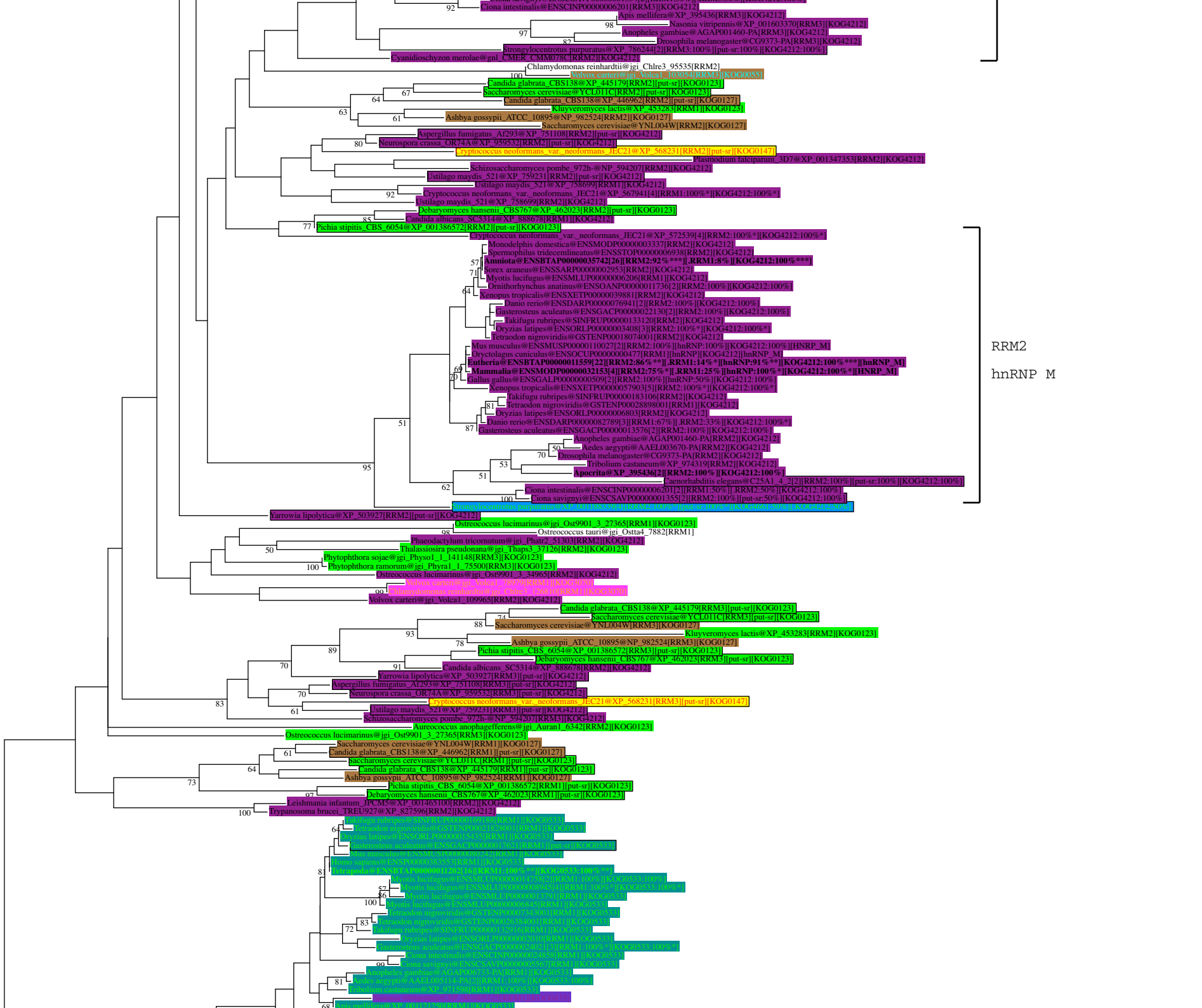
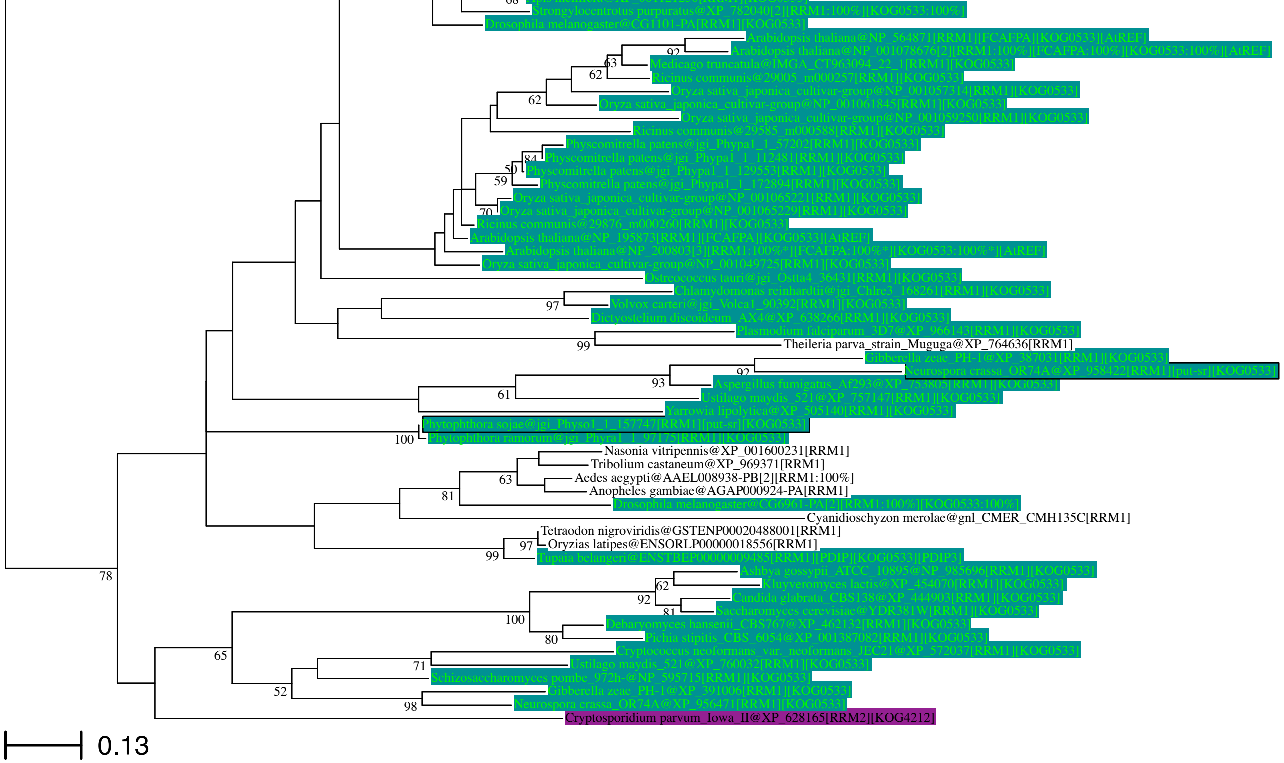


Fig. S17



Strongylocentrotus purpuratus@XP_782040[2][RRM1:100%][KOG0533:100%]
Drosophila melanogaster@CG1101-PA[RRM1][KOG0533]
Arabidopsis thaliana@NP_564871[RRM1][FCAFPA][KOG0533][AtREF]
Arabidopsis thaliana@NP_001078676[2][RRM1:100%][FCAFPA:100%][KOG0533:100%][AtREF]
Medicago truncatula@IMGA_CT963094_22_1[RRM1][KOG0533]
Ricinus communis@29005_m000257[RRM1][KOG0533]
Oryza sativa_japonica_cultivar-group@NP_001057314[RRM1][KOG0533]
Oryza sativa_japonica_cultivar-group@NP_001061845[RRM1][KOG0533]
Oryza sativa_japonica_cultivar-group@NP_001059250[RRM1][KOG0533]
Ricinus communis@29585_m000588[RRM1][KOG0533]
Physcomitrella patens@jgi_Phypa1_1_57202[RRM1][KOG0533]
Physcomitrella patens@jgi_Phypa1_1_112481[RRM1][KOG0533]
Physcomitrella patens@jgi_Phypa1_1_129553[RRM1][KOG0533]
Physcomitrella patens@jgi_Phypa1_1_172894[RRM1][KOG0533]
Oryza sativa_japonica_cultivar-group@NP_001065221[RRM1][KOG0533]
Oryza sativa_japonica_cultivar-group@NP_001065229[RRM1][KOG0533]
Ricinus communis@29876_m000260[RRM1][KOG0533]
Arabidopsis thaliana@NP_195873[RRM1][FCAFPA][KOG0533][AtREF]
Arabidopsis thaliana@NP_200803[3][RRM1:100%*][FCAFPA:100%*][KOG0533:100%*][AtREF]
Oryza sativa_japonica_cultivar-group@NP_001049725[RRM1][KOG0533]
Ostreococcus tauri@jgi_Ostta4_36431[RRM1][KOG0533]
Chlamydomonas reinhardtii@jgi_Chlre3_168261[RRM1][KOG0533]
Volvox carteri@jgi_Volca1_90392[RRM1][KOG0533]
Dictyostelium discoideum_AX4@XP_638266[RRM1][KOG0533]
Plasmodium falciparum_3D7@XP_966143[RRM1][KOG0533]
Theileria parva_strain_Muguga@XP_764636[RRM1]
Gibberella zeae_PH-1@XP_387031[RRM1][KOG0533]
Neurospora crassa_OR74A@XP_958422[RRM1][put-sr][KOG0533]
Aspergillus fumigatus_Af293@XP_753805[RRM1][KOG0533]
Ustilago maydis_521@XP_757147[RRM1][KOG0533]
Yarrowia lipolytica@XP_505140[RRM1][KOG0533]
Phytophthora sojae@jgi_Physo1_1_157747[RRM1][put-sr][KOG0533]
Phytophthora ramorum@jgi_Phyra1_1_97175[RRM1][KOG0533]
Nasonia vitripennis@XP_001600231[RRM1]
Tribolium castaneum@XP_969371[RRM1]
Aedes aegypti@AAEL008938-PB[2][RRM1:100%]
Anopheles gambiae@AGAP000924-PA[RRM1]
Drosophila melanogaster@CG6961-PA[2][RRM1:100%][KOG0533:100%]
Cyanidioschyzon merolae@gnl_CMER_CMH135C[RRM1]
Tetraodon nigroviridis@GSTENP00020488001[RRM1]
Oryzias latipes@ENSORLP00000018556[RRM1]
Tupaia belangeri@ENSTBEP00000009485[RRM1][PDIP][KOG0533][PDIP3]
Ashbya gossypii_ATCC_10895@NP_985696[RRM1][KOG0533]
Kluyveromyces lactis@XP_454070[RRM1][KOG0533]
Candida glabrata_CBS138@XP_444903[RRM1][KOG0533]
Saccharomyces cerevisiae@YDR381W[RRM1][KOG0533]
Debaryomyces hansenii_CBS767@XP_462132[RRM1][KOG0533]
Pichia stipitis_CBS_6054@XP_001387082[RRM1][KOG0533]
Cryptococcus neoformans_var._neoformans_JEC21@XP_572037[RRM1][KOG0533]
Ustilago maydis_521@XP_760032[RRM1][KOG0533]
Schizosaccharomyces pombe_972h-@NP_595715[RRM1][KOG0533]
Gibberella zeae_PH-1@XP_391006[RRM1][KOG0533]
Neurospora crassa_OR74A@XP_956471[RRM1][KOG0533]
Cryptosporidium parvum_Iowa_II@XP_628165[RRM2][KOG4212]
0.13

Fig. S18

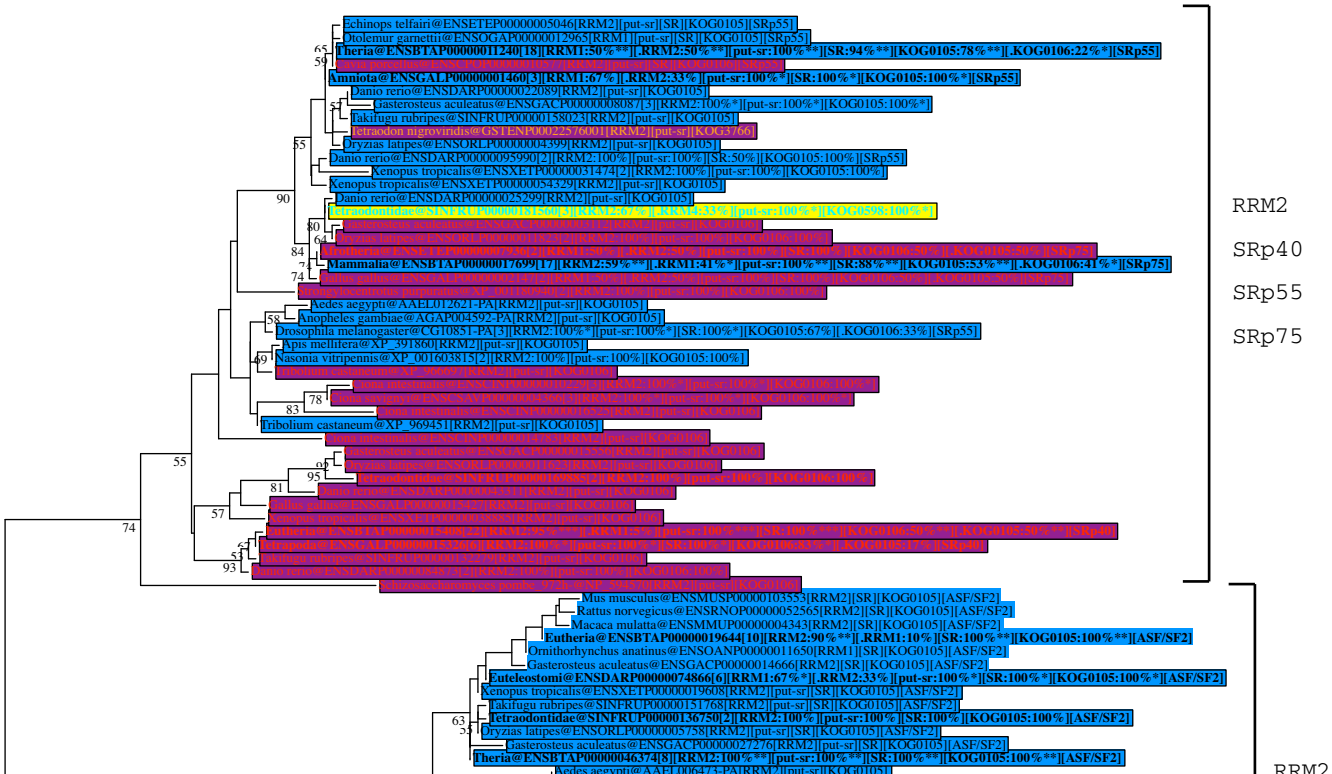

Echinops telfairi@ENSETEP00000005046[RRM2][put-sr][SR][KOG0105][SRp55]
Otolemur garnettii@ENSOGAP00000012965[RRM1][put-sr][SR][KOG0105][SRp55]
Theria@ENSBTAP00000011240[18][RRM1:50%**][.RRM2:50%**][put-sr:100%**][SR:94%**][KOG0105:78%**][.KOG0106:22%*][SRp55]
Amniota@ENSGALP00000001460[3][RRM1:67%*][.RRM2:33%][put-sr:100%*][SR:100%*][KOG0105:100%*][SRp55]
Danio rerio@ENSDARP00000022089[RRM2][put-sr][KOG0105]
Gasterosteus aculeatus@ENSGACP00000008087[3][RRM2:100%*][put-sr:100%*][KOG0105:100%*]
Takifugu rubripes@SINFRUP00000158023[RRM2][put-sr][KOG0105]
Tetraodon nigroviridis@GSTENP00022576001[RRM2][put-sr][KOG3766]
Oryzias latipes@ENSORLP00000004399[RRM2][put-sr][KOG0105]
Danio rerio@ENSDARP00000095990[2][RRM2:100%][put-sr:100%][SR:50%][KOG0105:100%][SRp55]
Xenopus tropicalis@ENSXETP00000031474[2][RRM2:100%][put-sr:100%][KOG0105:100%]
Xenopus tropicalis@ENSXETP00000054329[RRM2][put-sr][KOG0105]
Danio rerio@ENSDARP00000025299[RRM2][put-sr][KOG0105]
Tetraodontidae@SINFRUP00000131561[3][RRM2:67%*][RRM4:33%][put-sr:100%*][KOG0598:100%*]
Mammalia@ENSBTAP00000017699[17][RRM2:59%**][.RRM1:41%*][put-sr:100%**][SR:88%**][KOG0105:53%**][.KOG0106:41%*][SRp75]
Aedes aegypti@AAEL012621-PA[RRM2][put-sr][KOG0105]
Anopheles gambiae@AGAP004592-PA[RRM2][put-sr][KOG0105]
Drosophila melanogaster@CG10851-PA[3][RRM2:100%*][put-sr:100%*][SR:100%*][KOG0105:67%][.KOG0106:33%][SRp55]
Apis mellifera@XP_391860[RRM2][put-sr][KOG0105]
Nasonia vitripennis@XP_001603815[2][RRM2:100%][put-sr:100%][KOG0105:100%]
Tribolium castaneum@XP_969451[RRM2][put-sr][KOG0105]
Mus musculus@ENSMUSP00000103553[RRM2][SR][KOG0105][ASF/SF2]
Rattus norvegicus@ENSRNOP00000052565[RRM2][SR][KOG0105][ASF/SF2]
Macaca mulatta@ENSMMUP00000004343[RRM2][SR][KOG0105][ASF/SF2]
Eutheria@ENSBTAP00000019644[10][RRM2:90%**][.RRM1:10%][SR:100%**][KOG0105:100%**][ASF/SF2]
Ornithorhynchus anatinus@ENSOANP00000011650[RRM1][SR][KOG0105][ASF/SF2]
Gasterosteus aculeatus@ENSGACP00000014666[RRM2][SR][KOG0105][ASF/SF2]
Euteleostomi@ENSDARP00000074866[6][RRM1:67%*][.RRM2:33%][put-sr:100%*][SR:100%*][KOG0105:100%*][ASF/SF2]
Xenopus tropicalis@ENSXETP00000019608[RRM2][put-sr][SR][KOG0105][ASF/SF2]
Takifugu rubripes@SINFRUP00000151768[RRM2][put-sr][SR][KOG0105][ASF/SF2]
Tetraodontidae@SINFRUP00000136750[2][RRM2:100%][put-sr:100%][SR:100%][KOG0105:100%][ASF/SF2]
Oryzias latipes@ENSORLP00000005758[RRM2][put-sr][SR][KOG0105][ASF/SF2]
Gasterosteus aculeatus@ENSGACP00000027276[RRM2][put-sr][SR][KOG0105][ASF/SF2]
Theria@ENSBTAP00000046374[8][RRM2:100%**][put-sr:100%**][SR:100%**][KOG0105:100%**][ASF/SF2]
RRM2
SRp40
SRp55
SRp75
RRM2

Fig. S18

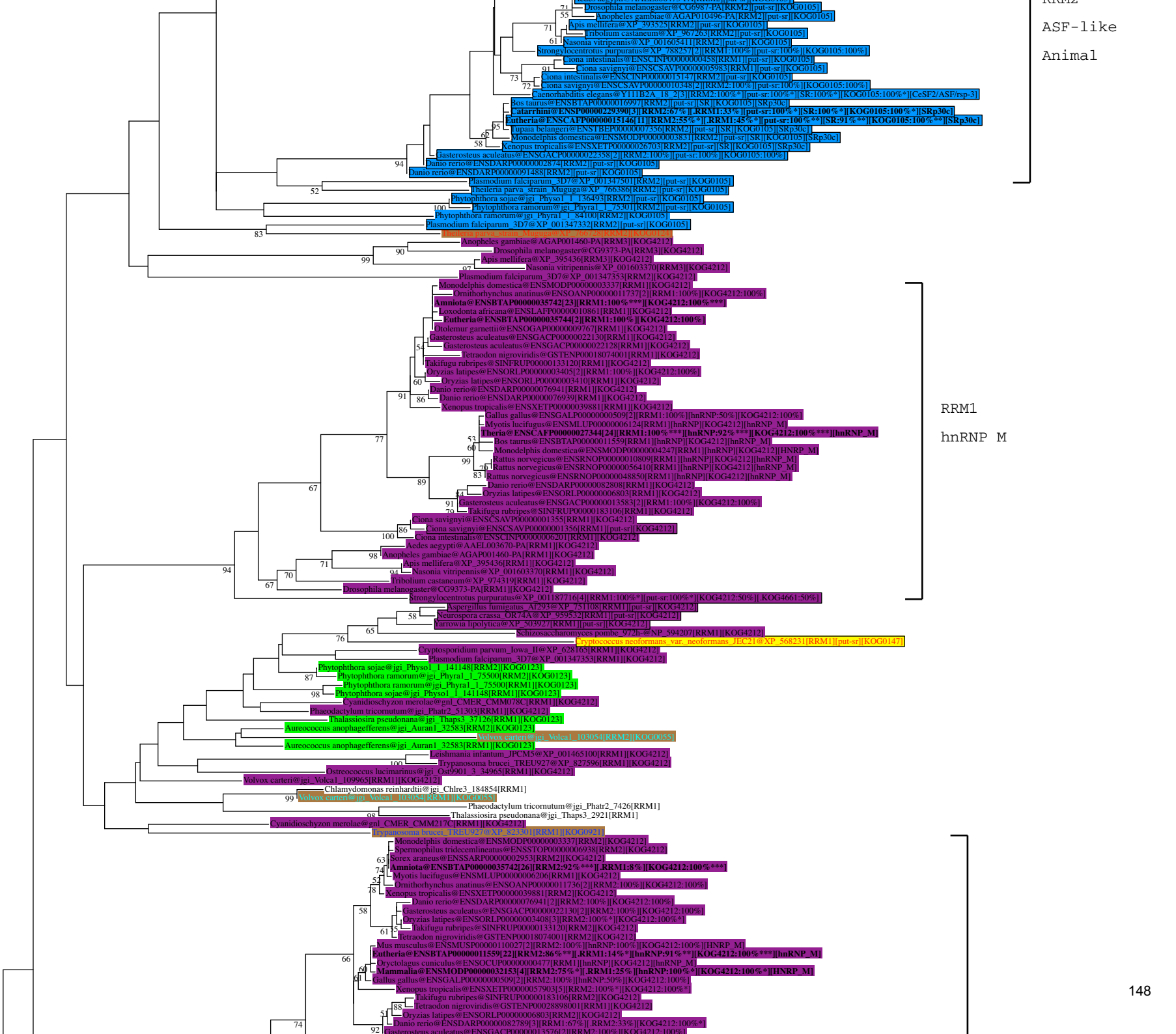


Fig. S18

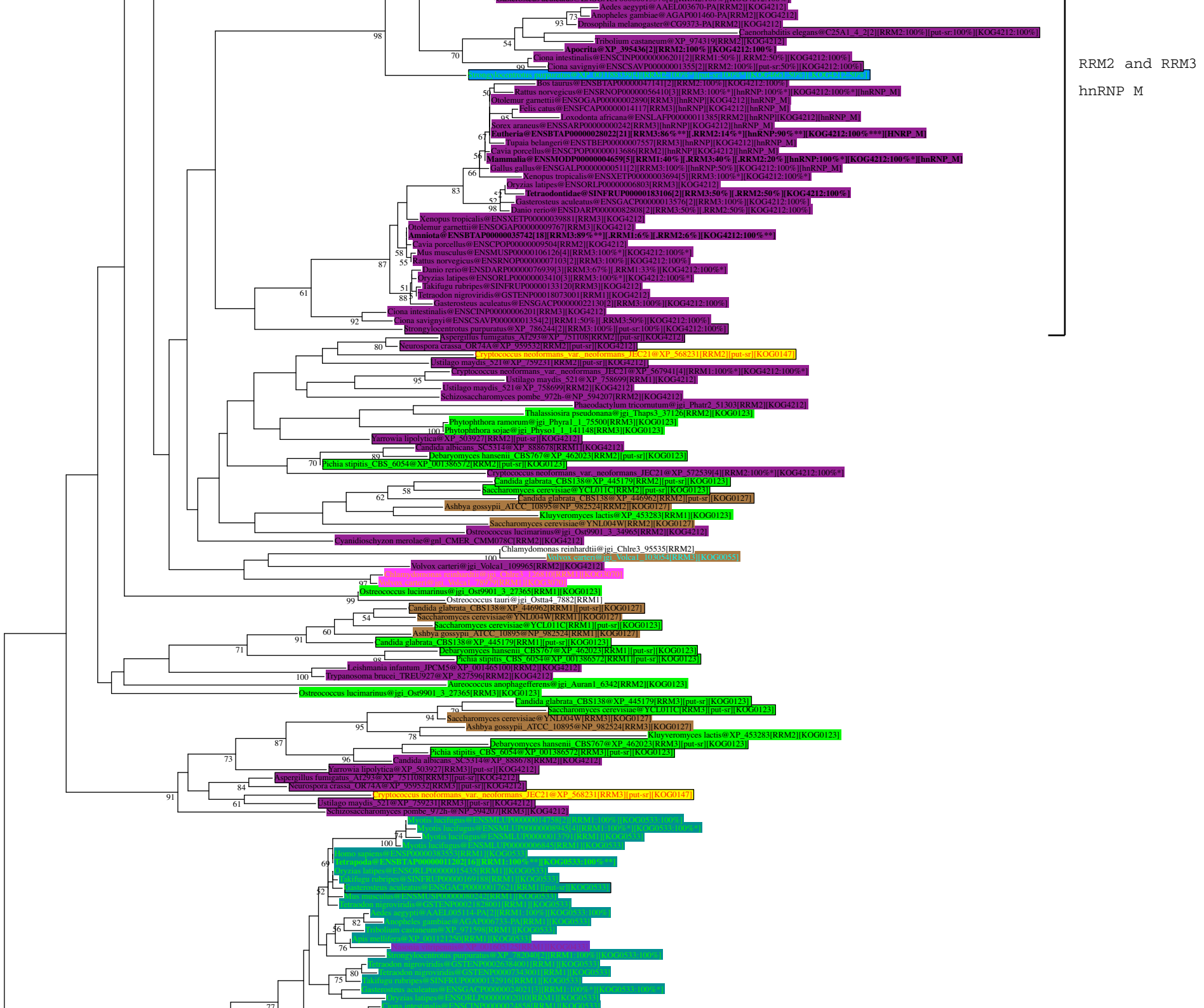


Fig. S18

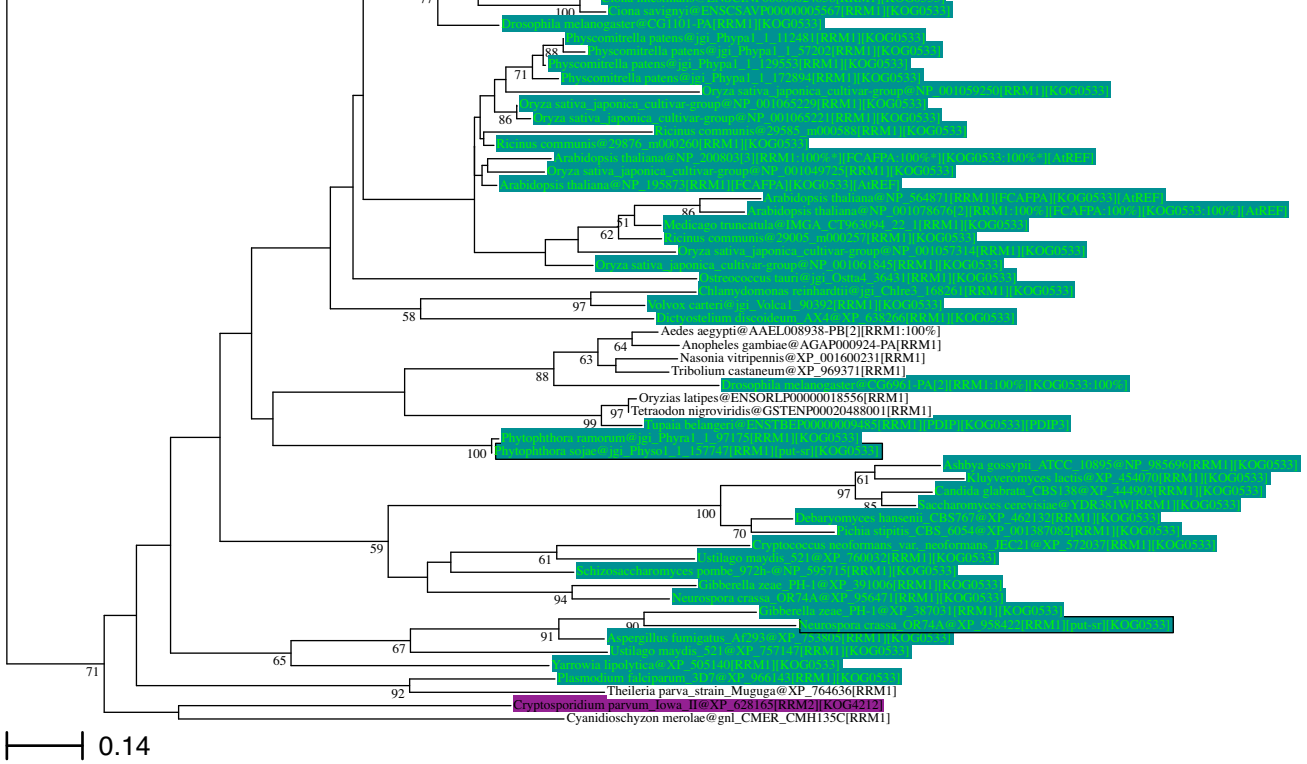

Drosophila_melanogaster@CG11101-PA[RRM1][KOG0533]
Physcomitrella patens@gnl_Phypa1_1_112481[RRM1][KOG0533]
Physcomitrella patens@gnl_Phypa1_1_57202[RRM1][KOG0533]
Physcomitrella patens@gnl_Phypa1_1_129353[RRM1][KOG0533]
Physcomitrella patens@gnl_Phypa1_1_172894[RRM1][KOG0533]
Oryza sativa_japonica_cultivar-group@NP_001059250[RRM1][KOG0533]
Oryza sativa_japonica_cultivar-group@NP_001065220[RRM1][KOG0533]
Oryza sativa_japonica_cultivar-group@NP_001065221[RRM1][KOG0533]
Ricinus communis@29585_m000588[RRM1][KOG0533]
Ricinus communis@29876_m000260[RRM1][KOG0533]
Arabidopsis thaliana@NP_200803[3][RRM1:100%*][FCAFPA:100%*][KOG0533-100%*][ARIF]
Oryza sativa_japonica_cultivar-group@NP_001059725[RRM1][KOG0533]
Arabidopsis thaliana@NP_195874[RRM1][FCAFPA][KOG0533][ARIF]
Arabidopsis thaliana@NP_001078676[2][RRM1:100%][FCAFPA:100%][KOG0533-100%][ARIF]
Medicago truncatula@TMGA_CT963064_22_1[RRM1][KOG0533]
Ricinus communis@29005_m000257[RRM1][KOG0533]
Oryza sativa_japonica_cultivar-group@NP_001057316[RRM1][KOG0533]
Oryza sativa_japonica_cultivar-group@NP_001061845[RRM1][KOG0533]
Chlamydomonas reinhardtii@gnl_Chlre3_168761[RRM1][KOG0533]
Volvox carteri@gnl_Volca1_90392[RRM1][KOG0533]
Dictyostelium discoideum_AX4@XP_638266[RRM1][KOG0533]
Aedes aegypti@AAEL008938-PB[2][RRM1:100%]
Anopheles gambiae@AGAP000924-PA[RRM1]
Nasonia vitripennis@XP_001600231[RRM1]
Tribolium castaneum@XP_969371[RRM1]
Drosophila melanogaster@CG6961-PA[2][RRM1:100%][KOG0533-100%]
Oryzias latipes@ENSORLP00000018556[RRM1]
Tetraodon nigroviridis@GSTENP00020488001[RRM1]
Tupaia belangeri@ENSTBEP00000004453[RRM1][PD][KOG0533][PD][P3]
Phytophthora ramorum@gnl_Phyra1_1_97175[RRM1][KOG0533]
Phytophthora sojae@gnl_Physoj1_1_155747[RRM1][partial][KOG0533]
Ashbya gossypii_ATCC_10895@NP_985696[RRM1][KOG0533]
Kluyveromyces lactis@XP_454020[RRM1][KOG0533]
Candida glabrata_CBS138@XP_447903[RRM1][KOG0533]
Saccharomyces cerevisiae@YDR351W[RRM1][KOG0533]
Schizosaccharomyces japonicus_CBS7796@XP_002172[RRM1][KOG0533]
Pichia stipitis_CBS_6054@XP_001387083[RRM1][KOG0533]
Cryptococcus neoformans_var._neoformans_JEC21@XP_572617[RRM1][KOG0533]
Debaryomyces_hansenii_521@XP_460032[RRM1][KOG0533]
Schizosaccharomyces_pombe_972h-@NP_595715[RRM1][KOG0533]
Gibberella zeae_PH-1@XP_391006[RRM1][KOG0533]
Neurospora crassa_OR74A@XP_956471[RRM1][KOG0533]
Gibberella zeae_PH-1@XP_387031[RRM1][KOG0533]
Neurospora crassa_OR74A@XP_958422[RRM1][partial][KOG0533]
Aspergillus fumigatus_Af293@XP_750665[RRM1][KOG0533]
Ustilago maydis_521@XP_757147[RRM1][KOG0533]
Yarrowia lipolytica@XP_505140[RRM1][KOG0533]
Plasmodium falciparum_3D7@XP_966143[RRM1][KOG0533]
Theileria parva_strain_Muguga@XP_764636[RRM1]
Cryptosporidium parvum_Iowa_II@XP_628165[RRM2][KOG4212]
Cyanidioschyzon merolae@gnl_CMER_CMH135C[RRM1]
0.14

Fig. S19

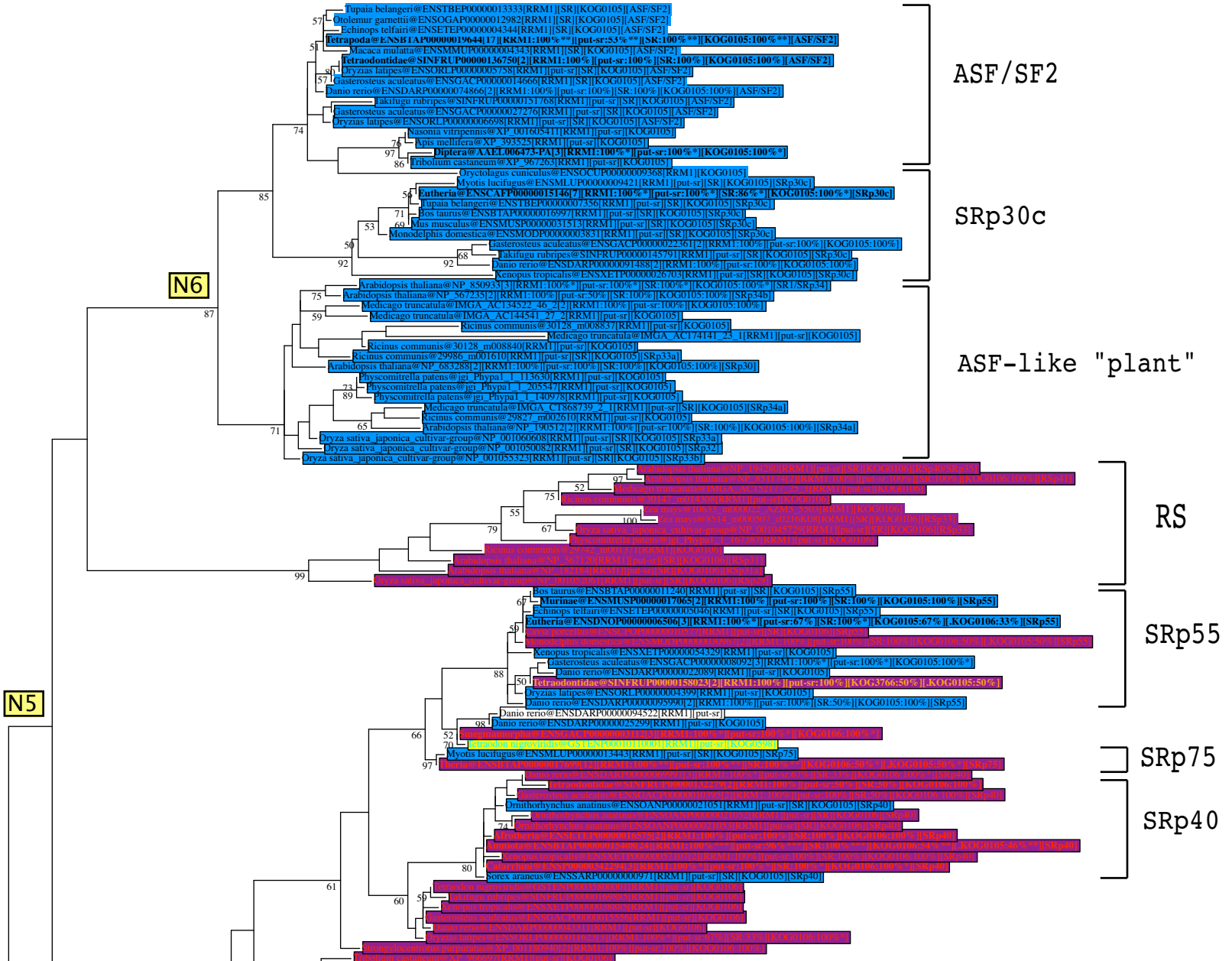

ASF/SF2
SRp30c
ASF-like "plant"
RS
SRp55
SRp75
SRp40
N6
N5
Tupaia belangeri@ENSTBEP00000013333[RRM1][SR][KOG0105][ASF/SF2]
Otolemur garnettii@ENSOGAP00000012982[RRM1][SR][KOG0105][ASF/SF2]
Echinops telfairi@ENSETEP00000004344[RRM1][SR][KOG0105][ASF/SF2]
Tetrapoda@ENSBTAP00000019644[17][RRM1:100%*][put-sr:53%*][SR:100%*][KOG0105:100%*][ASF/SF2]
Macaca mulatta@ENSMMUP00000004343[RRM1][SR][KOG0105][ASF/SF2]
Tetraodontidae@SINFRUP00000136750[2][RRM1:100%][put-sr:100%][SR:100%][KOG0105:100%][ASF/SF2]
Oryzias latipes@ENSORLP00000005758[RRM1][put-sr][SR][KOG0105][ASF/SF2]
Gasterosteus aculeatus@ENSGACP00000014666[RRM1][SR][KOG0105][ASF/SF2]
Danio rerio@ENSDARP00000074866[2][RRM1:100%][put-sr:100%][SR:100%][KOG0105:100%][ASF/SF2]
Takifugu rubripes@SINFRUP00000151768[RRM1][put-sr][SR][KOG0105][ASF/SF2]
Gasterosteus aculeatus@ENSGACP00000027276[RRM1][put-sr][SR][KOG0105][ASF/SF2]
Oryzias latipes@ENSORLP00000006698[RRM1][put-sr][SR][KOG0105][ASF/SF2]
Nasonia vitripennis@XP_001605411[RRM1][put-sr][KOG0105]
Apis mellifera@XP_393525[RRM1][put-sr][KOG0105]
Diptera@AAEL006473-PA[3][RRM1:100%*][put-sr:100%*][KOG0105:100%*]
Tribolium castaneum@XP_967263[RRM1][put-sr][KOG0105]
Oryctolagus cuniculus@ENSOCUP00000009368[RRM1][KOG0105]
Myotis lucifugus@ENSMLUP00000009421[RRM1][put-sr][SR][KOG0105][SRp30c]
Eutheria@ENSCAFP00000015146[7][RRM1:100%*][put-sr:100%*][SR:86%*][KOG0105:100%*][SRp30c]
Tupaia belangeri@ENSTBEP00000007356[RRM1][put-sr][SR][KOG0105][SRp30c]
Bos taurus@ENSBTAP00000016997[RRM1][put-sr][SR][KOG0105][SRp30c]
Mus musculus@ENSMUSP00000031513[RRM1][put-sr][SR][KOG0105][SRp30c]
Monodelphis domestica@ENSMODP00000003831[RRM1][put-sr][SR][KOG0105][SRp30c]
Gasterosteus aculeatus@ENSGACP00000022361[2][RRM1:100%][put-sr:100%][KOG0105:100%]
Takifugu rubripes@SINFRUP00000145791[RRM1][put-sr][SR][KOG0105][SRp30c]
Danio rerio@ENSDARP00000091488[2][RRM1:100%][put-sr:100%][KOG0105:100%]
Xenopus tropicalis@ENSXETP00000026703[RRM1][put-sr][SR][KOG0105][SRp30c]
Arabidopsis thaliana@NP_850933[3][RRM1:100%*][put-sr:100%*][SR:100%*][KOG0105:100%*][SR1/SRp34]
Arabidopsis thaliana@NP_567235[2][RRM1:100%][put-sr:50%][SR:100%][KOG0105:100%][SRp34b]
Medicago truncatula@IMGA_AC134522_46_2[2][RRM1:100%][put-sr:100%][KOG0105:100%]
Medicago truncatula@IMGA_AC144541_27_2[RRM1][put-sr][KOG0105]
Ricinus communis@30128_m008837[RRM1][put-sr][KOG0105]
Medicago truncatula@IMGA_AC174141_23_1[RRM1][put-sr][KOG0105]
Ricinus communis@30128_m008840[RRM1][put-sr][KOG0105]
Ricinus communis@29986_m00160[RRM1][put-sr][SR][KOG0105][SRp33a]
Arabidopsis thaliana@NP_683288[2][RRM1:100%][put-sr:100%][SR:100%][KOG0105:100%][SRp30]
Physcomitrella patens@jgi_Phypa1_1_113630[RRM1][put-sr][KOG0105]
Physcomitrella patens@jgi_Phypa1_1_205547[RRM1][put-sr][KOG0105]
Physcomitrella patens@jgi_Phypa1_1_140978[RRM1][put-sr][KOG0105]
Medicago truncatula@IMGA_CT868739_2_1[RRM1][put-sr][SR][KOG0105][SRp34a]
Ricinus communis@29827_m002610[RRM1][put-sr][KOG0105]
Arabidopsis thaliana@NP_190512[2][RRM1:100%][put-sr:100%][SR:100%][KOG0105:100%][SRp34a]
Oryza sativa_japonica_cultivar-group@NP_001060608[RRM1][put-sr][SR][KOG0105][SRp33a]
Oryza sativa_japonica_cultivar-group@NP_001050082[RRM1][put-sr][SR][KOG0105][SRp32]
Oryza sativa_japonica_cultivar-group@NP_001055323[RRM1][put-sr][SR][KOG0105][SRp33b]
Bos taurus@ENSBTAP00000011240[RRM1][put-sr][SR][KOG0105][SRp55]
Murinae@ENSMUSP00000017065[2][RRM1:100%][put-sr:100%][SR:100%][KOG0105:100%][SRp55]
Echinops telfairi@ENSETEP00000005046[RRM1][put-sr][SR][KOG0105][SRp55]
Eutheria@ENSDNOP00000006506[3][RRM1:100%*][put-sr:67%][SR:100%*][KOG0105:67%][.KOG0106:33%][SRp55]
Xenopus tropicalis@ENSXETP00000054329[RRM1][put-sr][KOG0105]
Gasterosteus aculeatus@ENSGACP00000008092[3][RRM1:100%*][put-sr:100%*][KOG0105:100%*]
Danio rerio@ENSDARP00000022089[RRM1][put-sr][KOG0105]
Tetraodontidae@SINFRUP00000158023[2][RRM1:100%][put-sr:100%][KOG3766:50%][.KOG0105:50%]
Oryzias latipes@ENSORLP00000004399[RRM1][put-sr][KOG0105]
Danio rerio@ENSDARP00000095990[2][RRM1:100%][put-sr:100%][SR:50%][KOG0105:100%][SRp55]
Danio rerio@ENSDARP00000094522[RRM1][put-sr]
Danio rerio@ENSDARP00000025299[RRM1][put-sr][KOG0105]
Myotis lucifugus@ENSMLUP00000013443[RRM1][put-sr][SR][KOG0105][SRp75]
Ornithorhynchus anatinus@ENSOANP00000021051[RRM1][put-sr][SR][KOG0105][SRp40]
Sorex araneus@ENSSARP00000000971[RRM1][put-sr][SR][KOG0105][SRp40]

Fig. S19

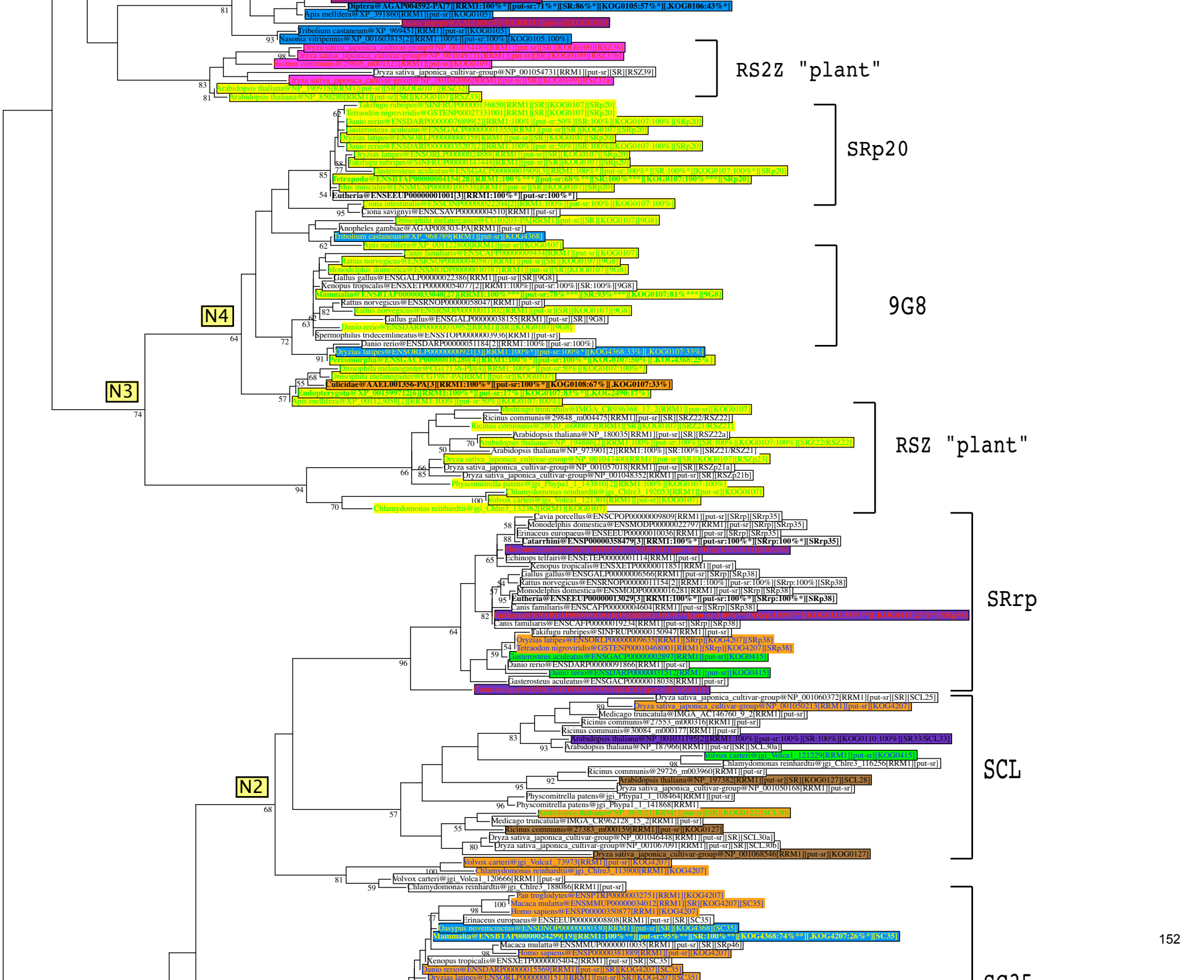


Fig. S19

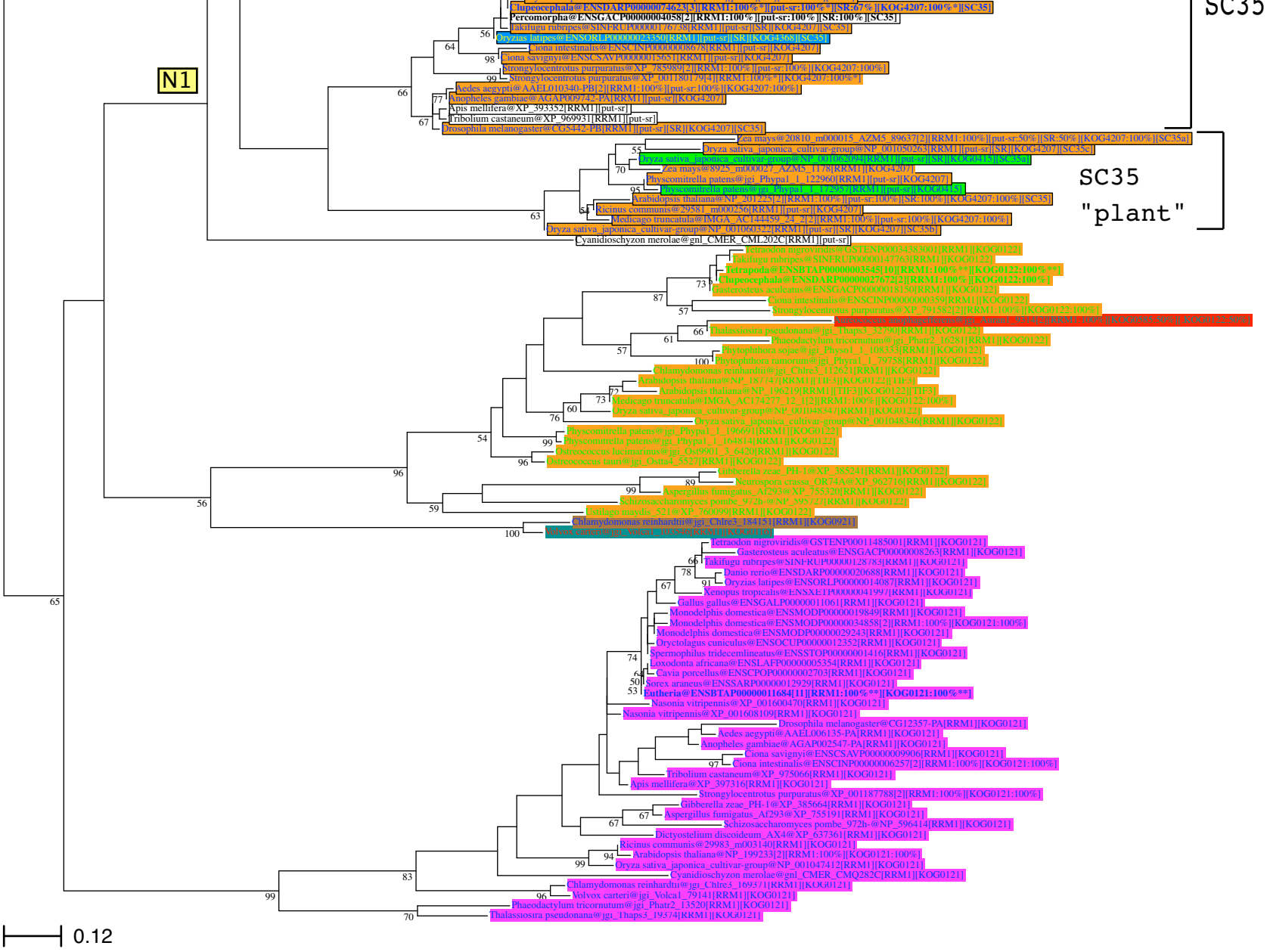


Fig. S20

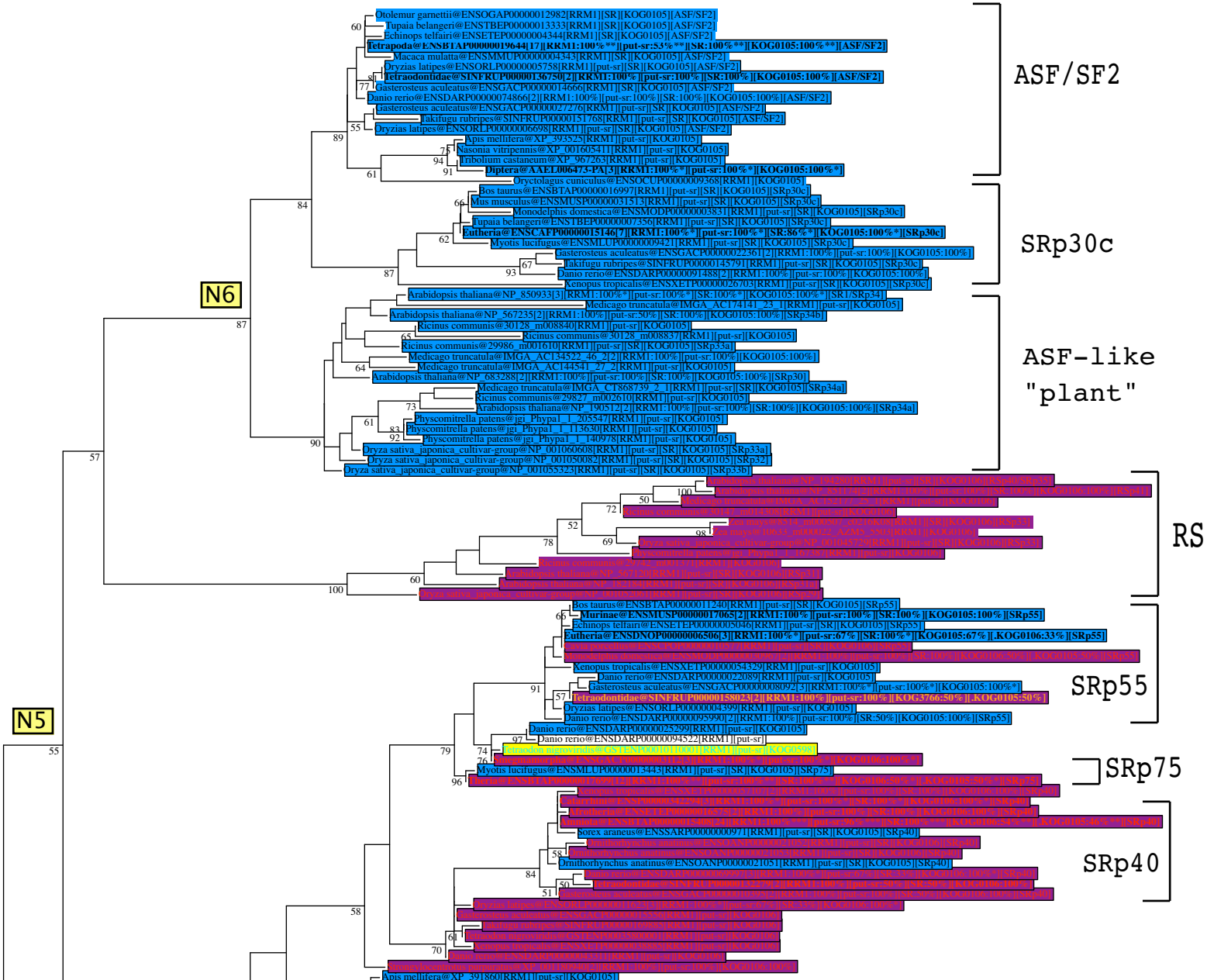

Otolemur garnettii@ENSOGAP00000012982[RRM1][SR][KOG0105][ASF/SF2]
Tupaia belangeri@ENSTBEP00000013333[RRM1][SR][KOG0105][ASF/SF2]
Echinops telfairi@ENSETEP00000004344[RRM1][SR][KOG0105][ASF/SF2]
Tetrapoda@ENSBTAP00000019644[17][RRM1:100%*][put-sr:53%**][SR:100%**][KOG0105:100%**][ASF/SF2]
Macaca mulatta@ENSMMUP00000004343[RRM1][SR][KOG0105][ASF/SF2]
Oryzias latipes@ENSORLP00000005758[RRM1][put-sr][SR][KOG0105][ASF/SF2]
Tetraodontidae@SINFRUP00000136750[2][RRM1:100%][put-sr:100%*][SR:100%*][KOG0105:100%][ASF/SF2]
Gasterosteus aculeatus@ENSGACP00000014666[RRM1][SR][KOG0105][ASF/SF2]
Danio rerio@ENSDARP00000074866[2][RRM1:100%][put-sr:100%][SR:100%][KOG0105:100%][ASF/SF2]
Gasterosteus aculeatus@ENSGACP00000027276[RRM1][put-sr][SR][KOG0105][ASF/SF2]
Takifugu rubripes@SINFRUP00000151768[RRM1][put-sr][SR][KOG0105][ASF/SF2]
Oryzias latipes@ENSORLP00000006698[RRM1][put-sr][SR][KOG0105][ASF/SF2]
Apis mellifera@XP_393525[RRM1][put-sr][KOG0105]
Nasonia vitripennis@XP_001605411[RRM1][put-sr][KOG0105]
Tribolium castaneum@XP_967263[RRM1][put-sr][KOG0105]
Diptera@AAEL006473-PA[3][RRM1:100%*][put-sr:100%*][KOG0105:100%*]
Oryctolagus cuniculus@ENSOCUP00000009368[RRM1][KOG0105]
Bos taurus@ENSBTAP00000016997[RRM1][put-sr][SR][KOG0105][SRp30c]
Mus musculus@ENSMUSP00000031513[RRM1][put-sr][SR][KOG0105][SRp30c]
Monodelphis domestica@ENSMODP00000003831[RRM1][put-sr][SR][KOG0105][SRp30c]
Tupaia belangeri@ENSTBEP00000007356[RRM1][put-sr][SR][KOG0105][SRp30c]
Eutheria@ENSCAFP00000015146[7][RRM1:100%*][put-sr:100%*][SR:86%*][KOG0105:100%*][SRp30c]
Myotis lucifugus@ENSMLUP00000009421[RRM1][put-sr][SR][KOG0105][SRp30c]
Gasterosteus aculeatus@ENSGACP00000022361[2][RRM1:100%][put-sr:100%][KOG0105:100%]
Takifugu rubripes@SINFRUP00000145791[RRM1][put-sr][SR][KOG0105][SRp30c]
Danio rerio@ENSDARP00000091488[2][RRM1:100%][put-sr:100%][KOG0105:100%]
Xenopus tropicalis@ENSXETP00000026703[RRM1][put-sr][SR][KOG0105][SRp30c]
Arabidopsis thaliana@NP_850933[3][RRM1:100%*][put-sr:100%*][SR:100%*][KOG0105:100%*][SRT/SRp34]
Medicago truncatula@IMGA_AC174141_23_1[RRM1][put-sr][KOG0105]
Arabidopsis thaliana@NP_567235[2][RRM1:100%][put-sr:50%][SR:100%][KOG0105:100%][SRp34b]
Ricinus communis@30128_m008840[RRM1][put-sr][KOG0105]
Ricinus communis@30128_m008837[RRM1][put-sr][KOG0105]
Ricinus communis@29986_m001610[RRM1][put-sr][SR][KOG0105][SRp33a]
Medicago truncatula@IMGA_AC134522_46_2[2][RRM1:100%][put-sr:100%*][KOG0105:100%]
Medicago truncatula@IMGA_AC14541_27_2[RRM1][put-sr][KOG0105]
Arabidopsis thaliana@NP_683285[2][RRM1:100%][put-sr:100%*][SR:100%][KOG0105:100%][SRp30]
Medicago truncatula@IMGA_CT868739_2_1[RRM1][put-sr][SR][KOG0105][SRp34a]
Ricinus communis@29827_m002610[RRM1][put-sr][KOG0105]
Arabidopsis thaliana@NP_190512[2][RRM1:100%][put-sr:100%][SR:100%][KOG0105:100%][SRp34a]
Physcomitrella patens@jgi_Phypa1_1_205547[RRM1][put-sr][KOG0105]
Physcomitrella patens@jgi_Phypa1_1_113630[RRM1][put-sr][KOG0105]
Physcomitrella patens@jgi_Phypa1_1_140978[RRM1][put-sr][KOG0105]
Oryza sativa_japonica_cultivar-group@NP_001060608[RRM1][put-sr][SR][KOG0105][SRp33a]
Oryza sativa_japonica_cultivar-group@NP_001050082[RRM1][put-sr][SR][KOG0105][SRp32]
Oryza sativa_japonica_cultivar-group@NP_001055323[RRM1][put-sr][SR][KOG0105][SRp33b]
Bos taurus@ENSBTAP00000011240[RRM1][put-sr][SR][KOG0105][SRp55]
Murinae@ENSMUSP00000017065[2][RRM1:100%*][put-sr:100%*][SR:100%*][KOG0105:100%][SRp55]
Echinops telfairi@ENSETEP00000005046[RRM1][put-sr][SR][KOG0105][SRp55]
Eutheria@ENSDNOP00000006506[3][RRM1:100%*][put-sr:67%*][SR:100%*][KOG0105:67%][.KOG0106:33%][SRp55]
Xenopus tropicalis@ENSXETP00000054329[RRM1][put-sr][KOG0105]
Danio rerio@ENSDARP00000022089[RRM1][put-sr][KOG0105]
Gasterosteus aculeatus@ENSGACP00000008092[3][RRM1:100%*][put-sr:100%*][KOG0105:100%*]
Tetraodontidae@SINFRUP00000158023[2][RRM1:100%][put-sr:100%][KOG3766:50%][.KOG0105:50%]
Oryzias latipes@ENSORLP00000004399[RRM1][put-sr][KOG0105]
Danio rerio@ENSDARP00000095990[2][RRM1:100%][put-sr:100%][SR:50%][KOG0105:100%][SRp55]
Danio rerio@ENSDARP00000025299[RRM1][put-sr][KOG0105]
Danio rerio@ENSDARP00000094522[RRM1][put-sr]
Myotis lucifugus@ENSMLUP00000013443[RRM1][put-sr][SR][KOG0105][SRp75]
Sorex araneus@ENSSARP00000000971[RRM1][put-sr][SR][KOG0105][SRp40]
Ornithorhynchus anatinus@ENSOANP00000021051[RRM1][put-sr][SR][KOG0105][SRp40]
Apis mellifera@XP_391860[RRM1][put-sr][KOG0105]
N6
N5
ASF/SF2
SRp30c
ASF-like
"plant"
RS
SRp55
SRp75
SRp40

Fig. S20

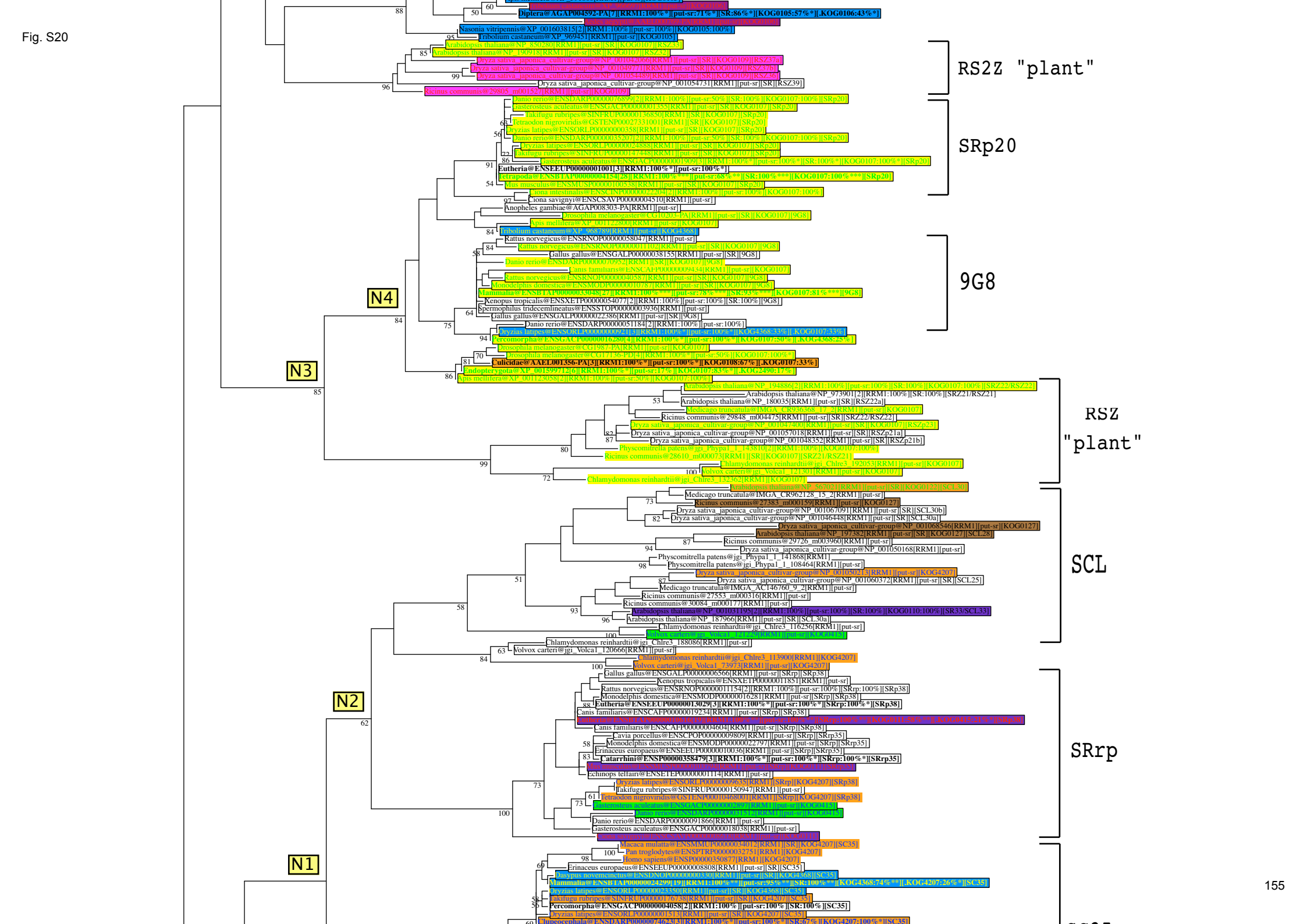


Fig. S20

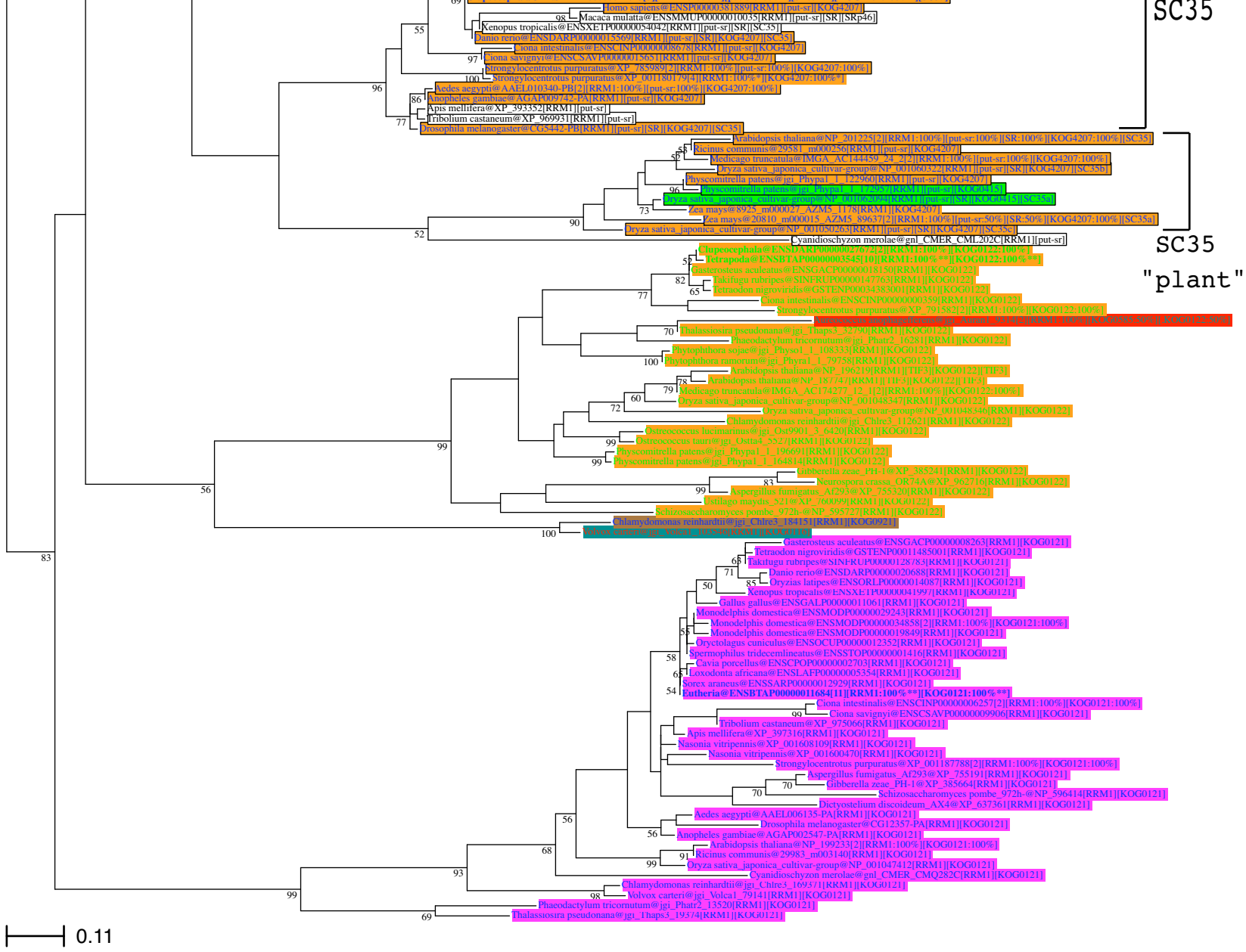


Fig. S21

A. RRM1 of Single RRM SR proteins

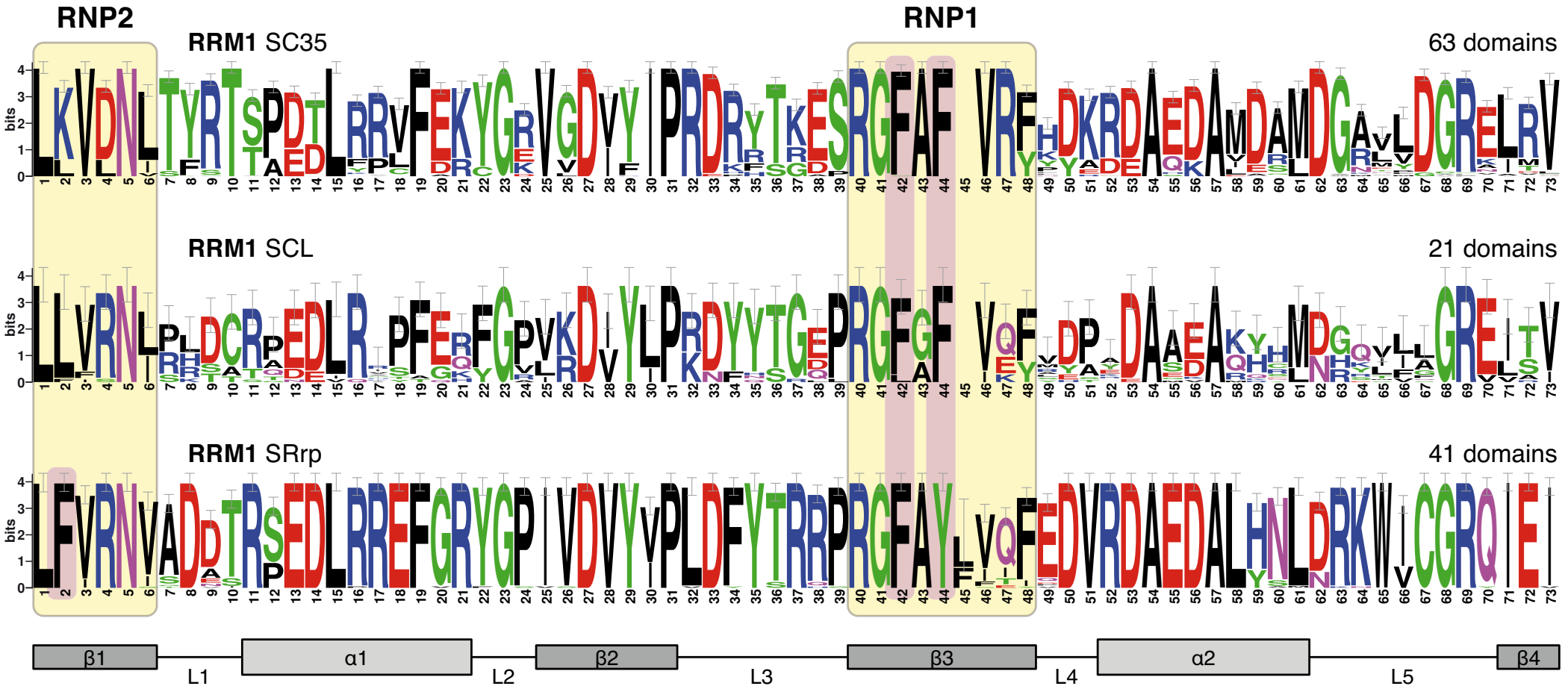


Fig. S21

B. RRM1 of Single RRM ZnK-like SR proteins

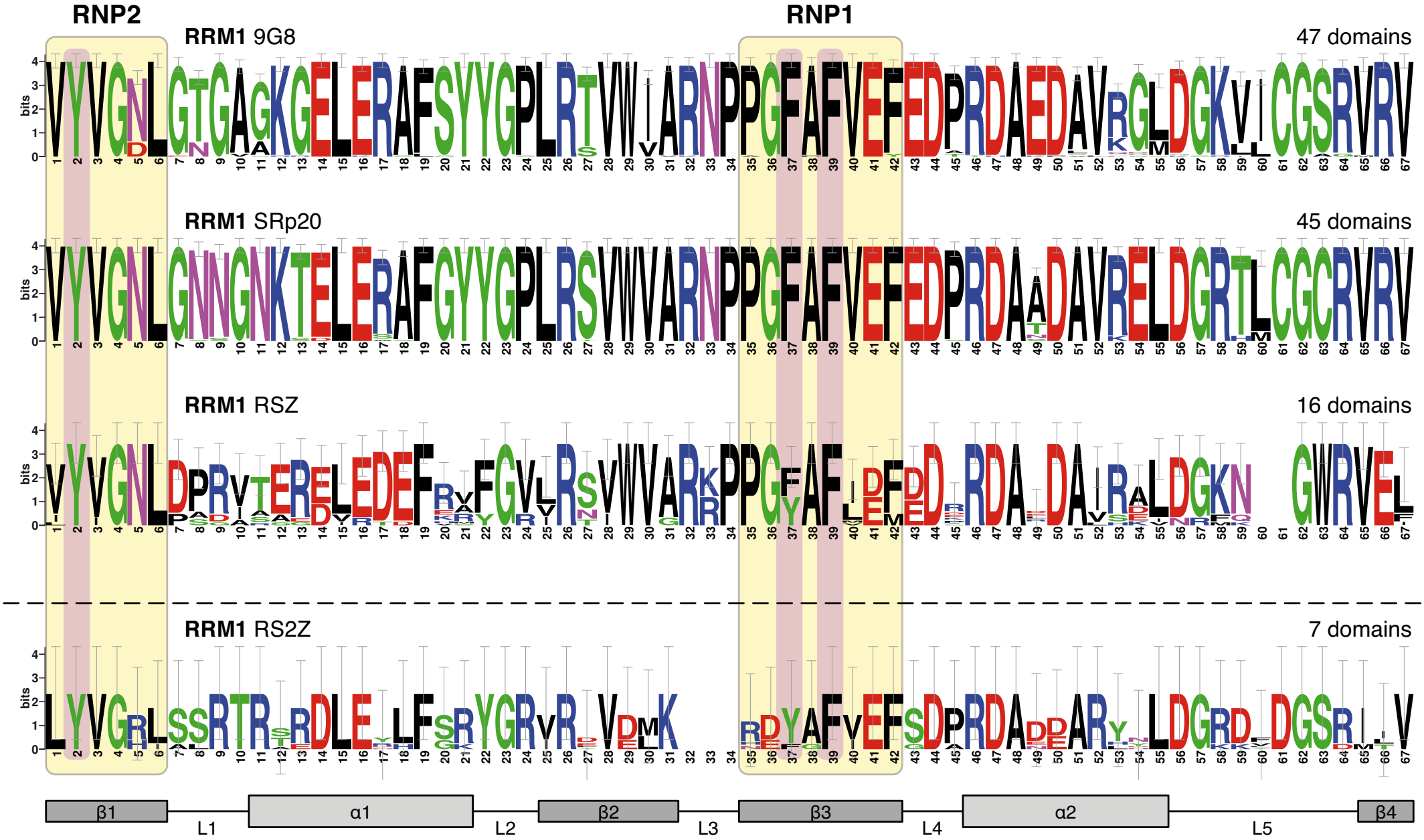


Fig. S21

C. RRM1 of Dual RRM SR proteins

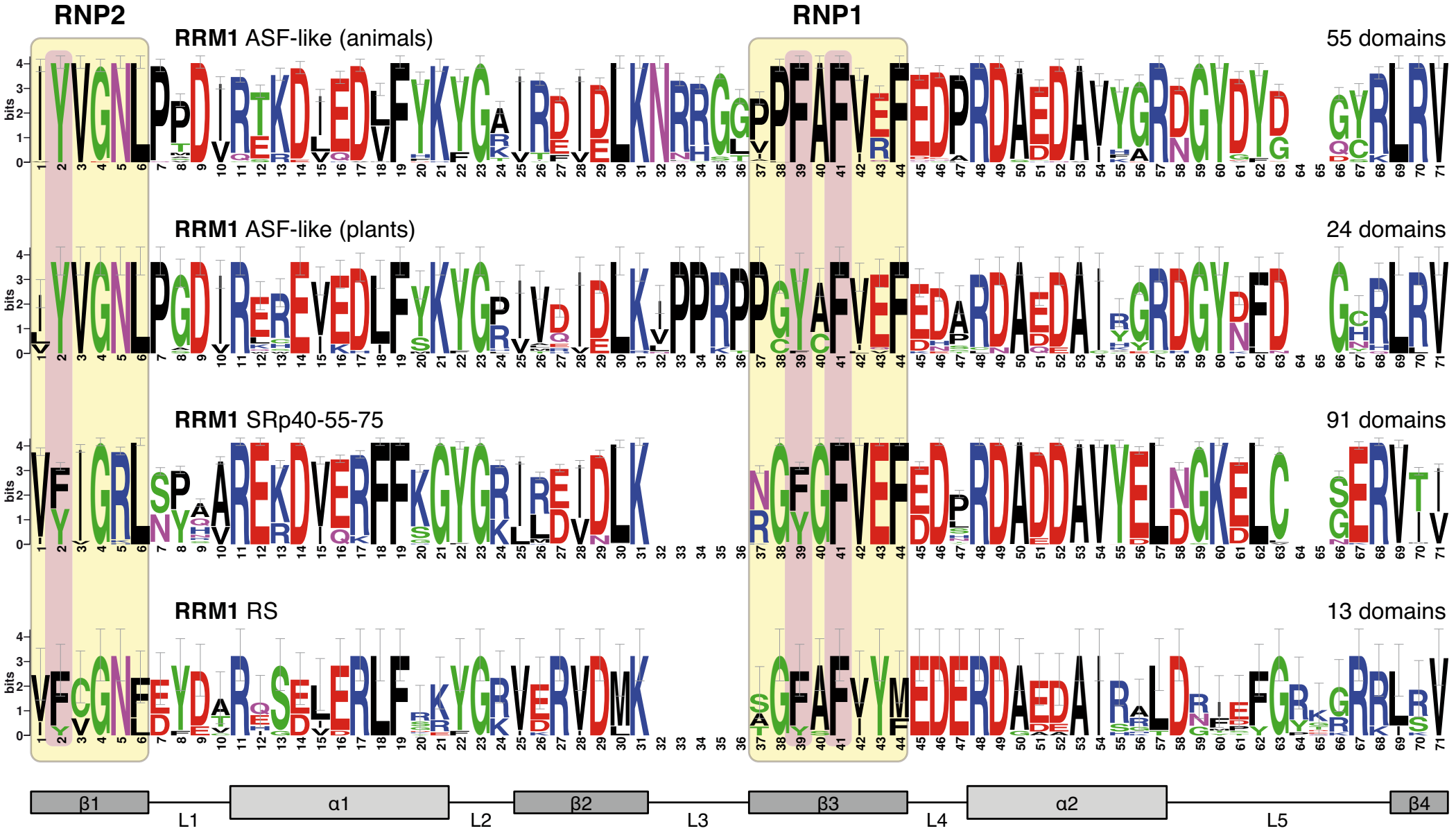


Fig. S21

D. RRM1 of RNPS1-like proteins

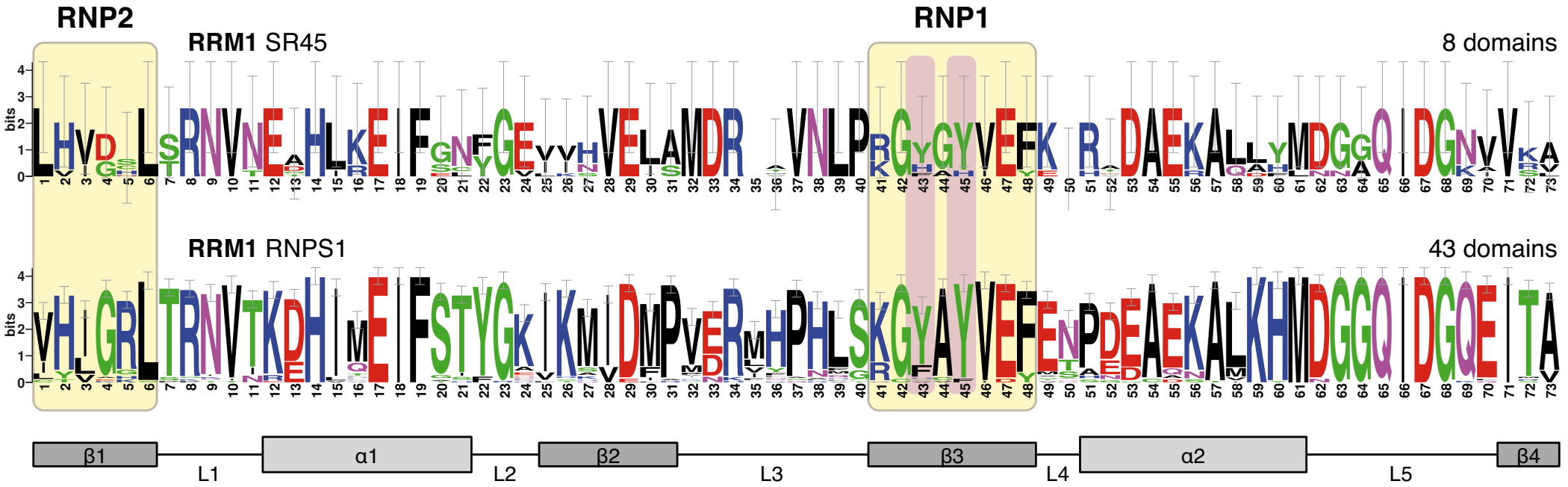


Fig. S22

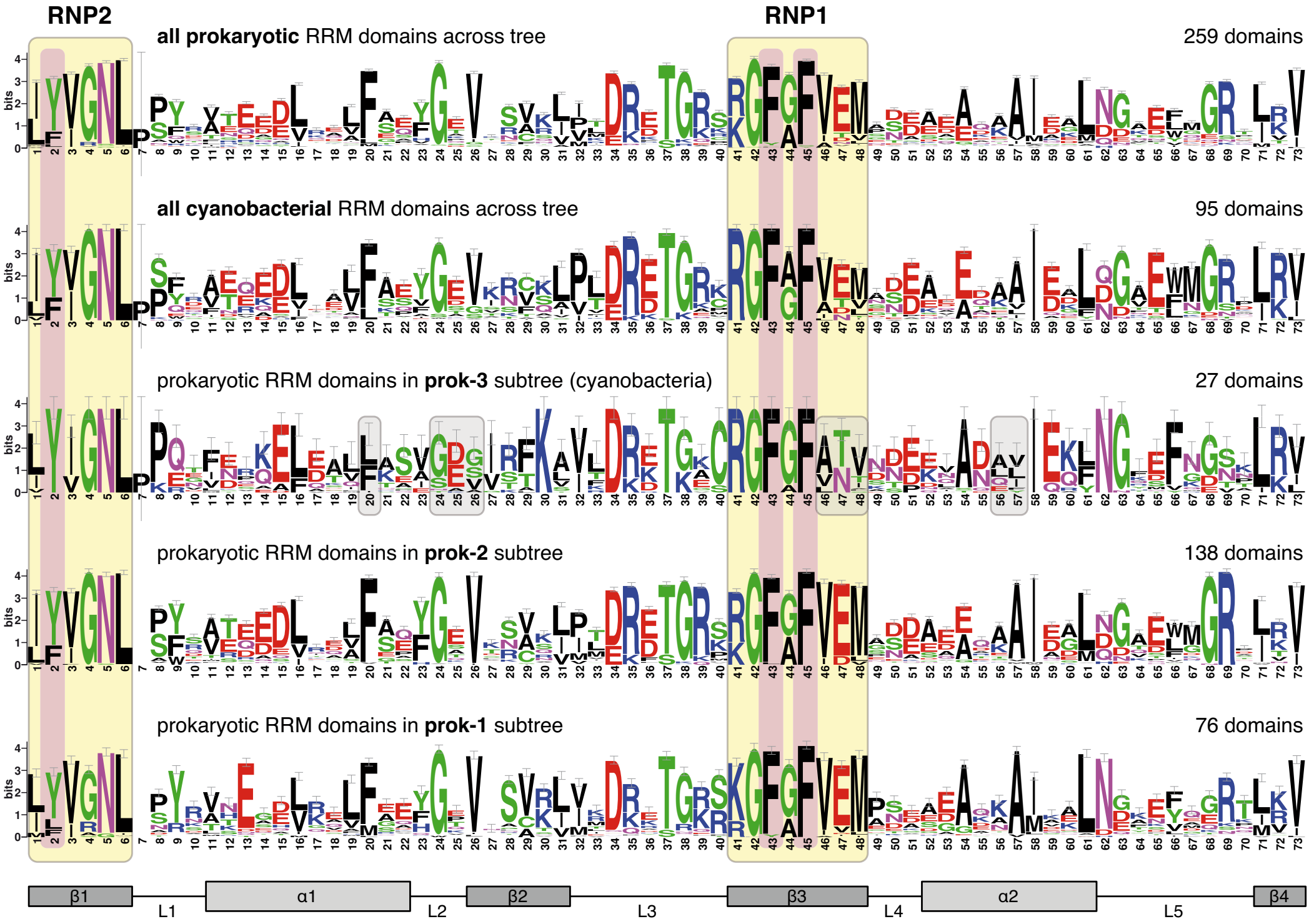
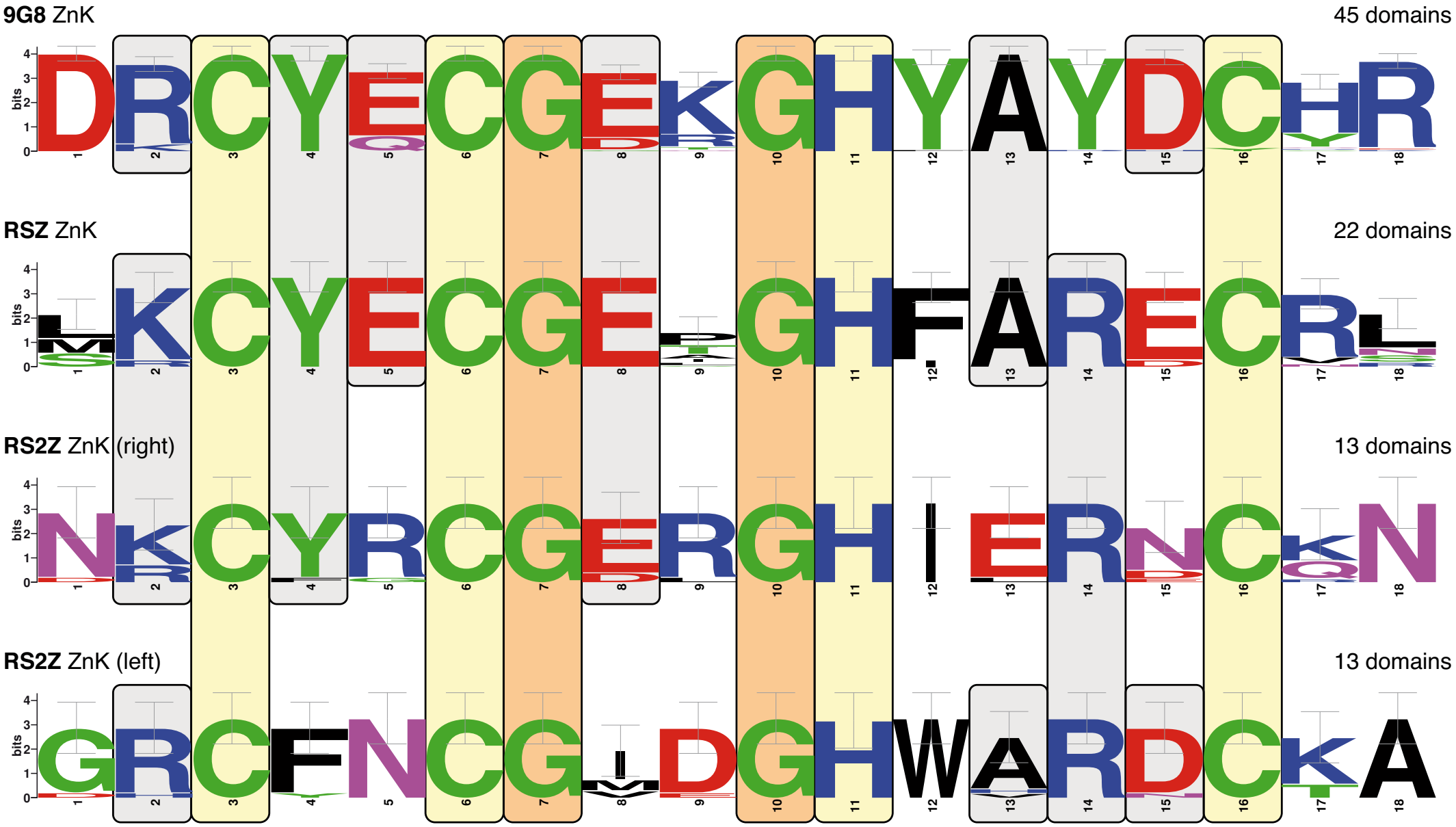
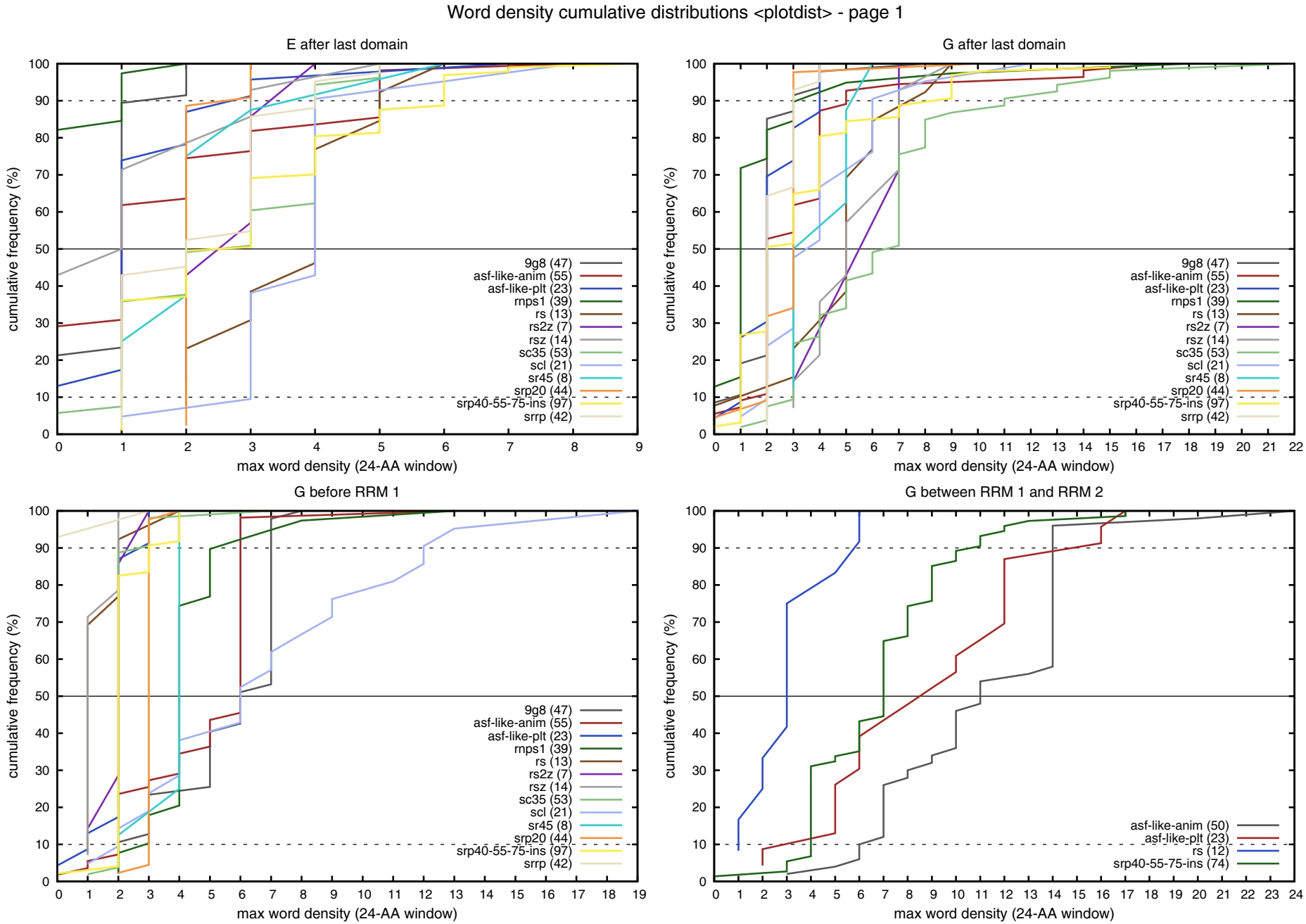
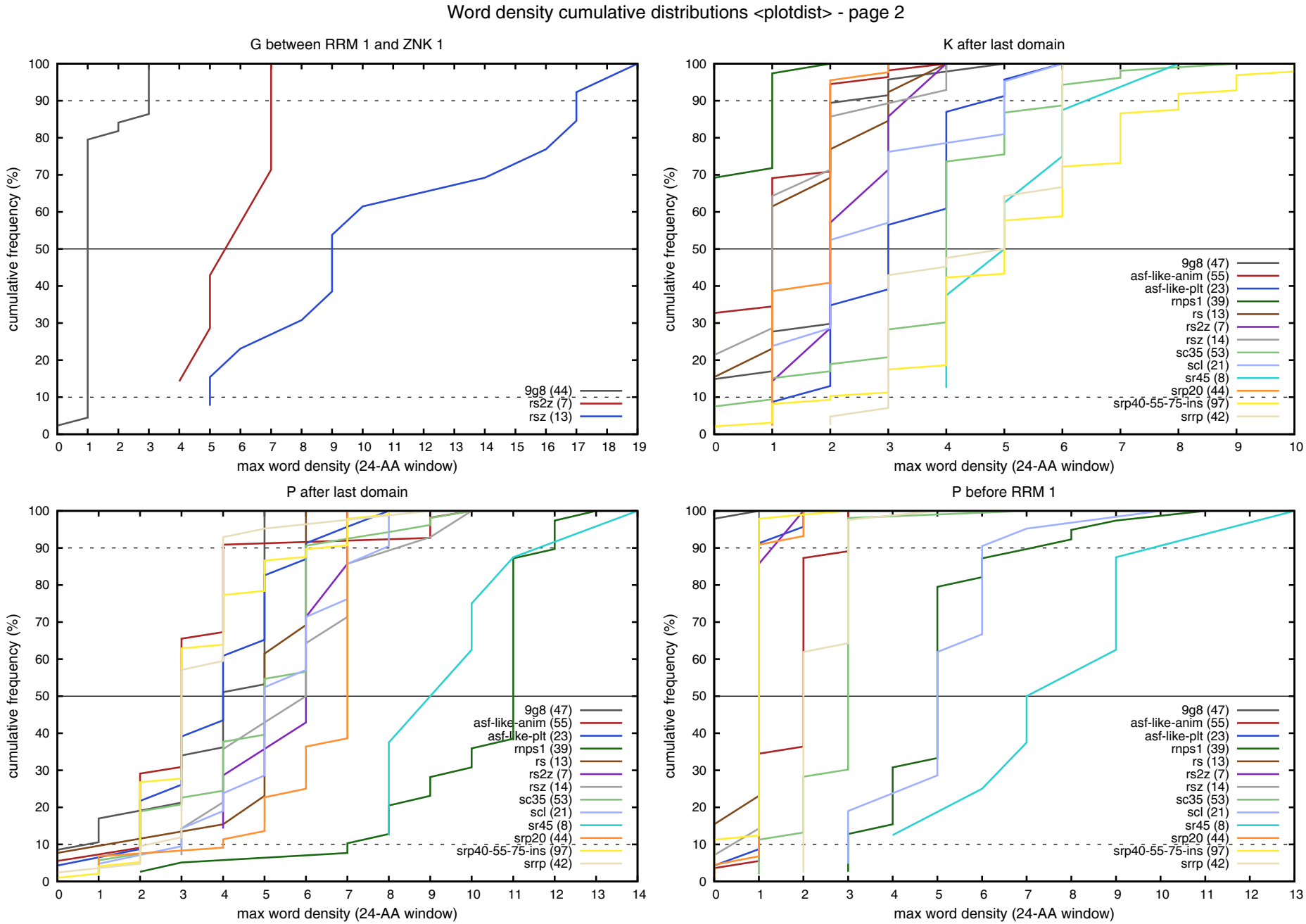

RNP2
RNP1
all prokaryotic RRM domains across tree
259 domains
all cyanobacterial RRM domains across tree
95 domains
prokaryotic RRM domains in prok-3 subtree (cyanobacteria)
27 domains
prokaryotic RRM domains in prok-2 subtree
138 domains
prokaryotic RRM domains in prok-1 subtree
76 domains
bits
β1
α1
β2
β3
α2
β4
L1
L2
L3
L4
L5

Fig. S23

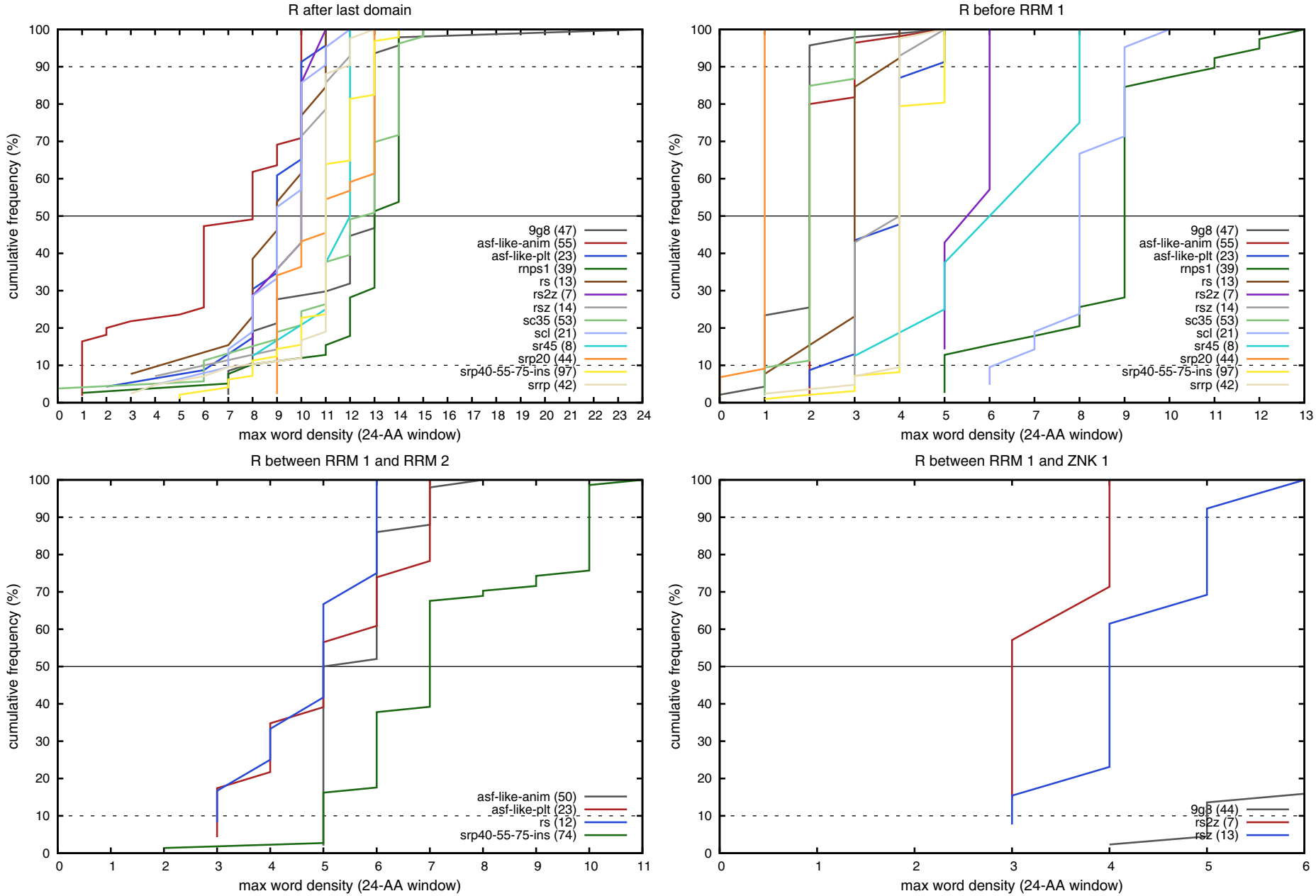


C-Φ-X-C-G-±-X-G-H-X-X-X-δ-C

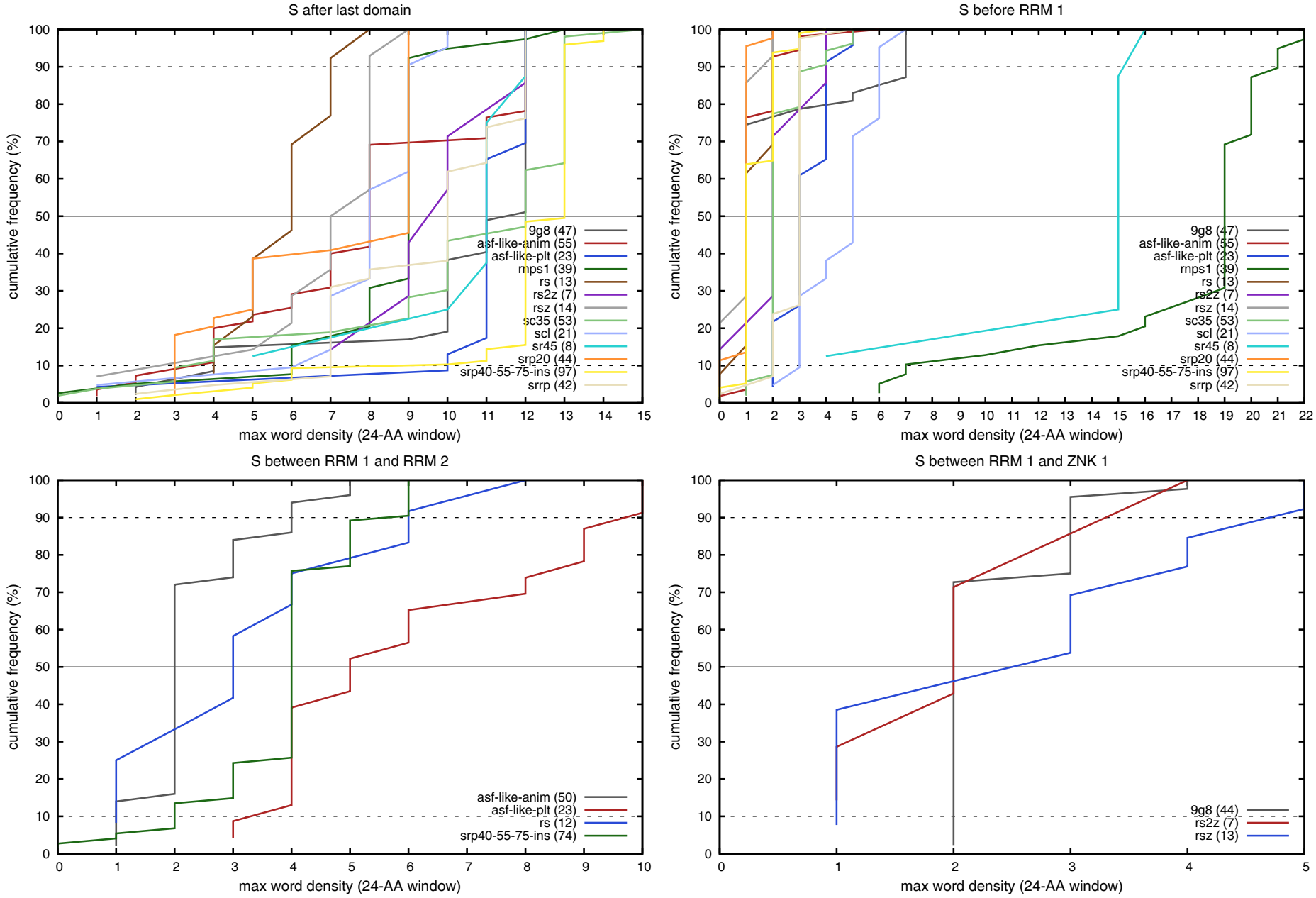

Word density cumulative distributions <plotdist> - page 1
E after last domain
G after last domain
G before RRM 1
G between RRM 1 and RRM 2
cumulative frequency (%)
max word density (24-AA window)
9g8 (47)
asf-like-anim (55)
asf-like-plt (23)
rnps1 (39)
rs (13)
rs2z (7)
rsz (14)
sc35 (53)
scl (21)
sr45 (8)
srp20 (44)
srp40-55-75-ins (97)
srrp (42)
asf-like-anim (50)
asf-like-plt (23)
rs (12)
srp40-55-75-ins (74)


Word density cumulative distributions <plotdist> - page 2
G between RRM 1 and ZNK 1
cumulative frequency (%)
max word density (24-AA window)
9g8 (44)
rs2z (7)
rsz (13)
K after last domain
cumulative frequency (%)
max word density (24-AA window)
9g8 (47)
asf-like-anim (55)
asf-like-plt (23)
rnps1 (39)
rs (13)
rs2z (7)
rsz (14)
sc35 (53)
scl (21)
sr45 (8)
srp20 (44)
srp40-55-75-ins (97)
srrp (42)
P after last domain
cumulative frequency (%)
max word density (24-AA window)
9g8 (47)
asf-like-anim (55)
asf-like-plt (23)
rnps1 (39)
rs (13)
rs2z (7)
rsz (14)
sc35 (53)
scl (21)
sr45 (8)
srp20 (44)
srp40-55-75-ins (97)
srrp (42)
P before RRM 1
cumulative frequency (%)
max word density (24-AA window)
9g8 (47)
asf-like-anim (55)
asf-like-plt (23)
rnps1 (39)
rs (13)
rs2z (7)
rsz (14)
sc35 (53)
scl (21)
sr45 (8)
srp20 (44)
srp40-55-75-ins (97)
srrp (42)

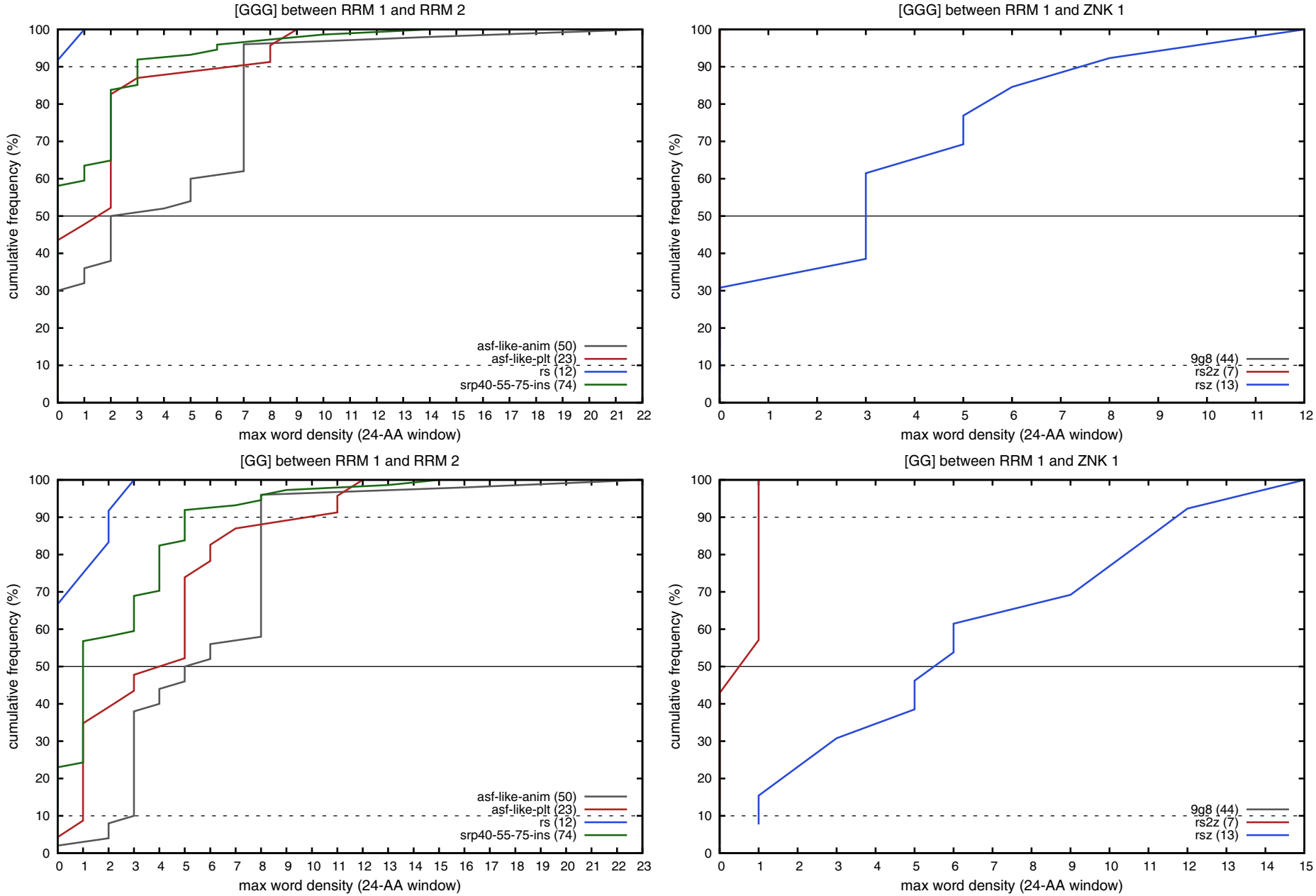
Word density cumulative distributions <plotdist> - page 3



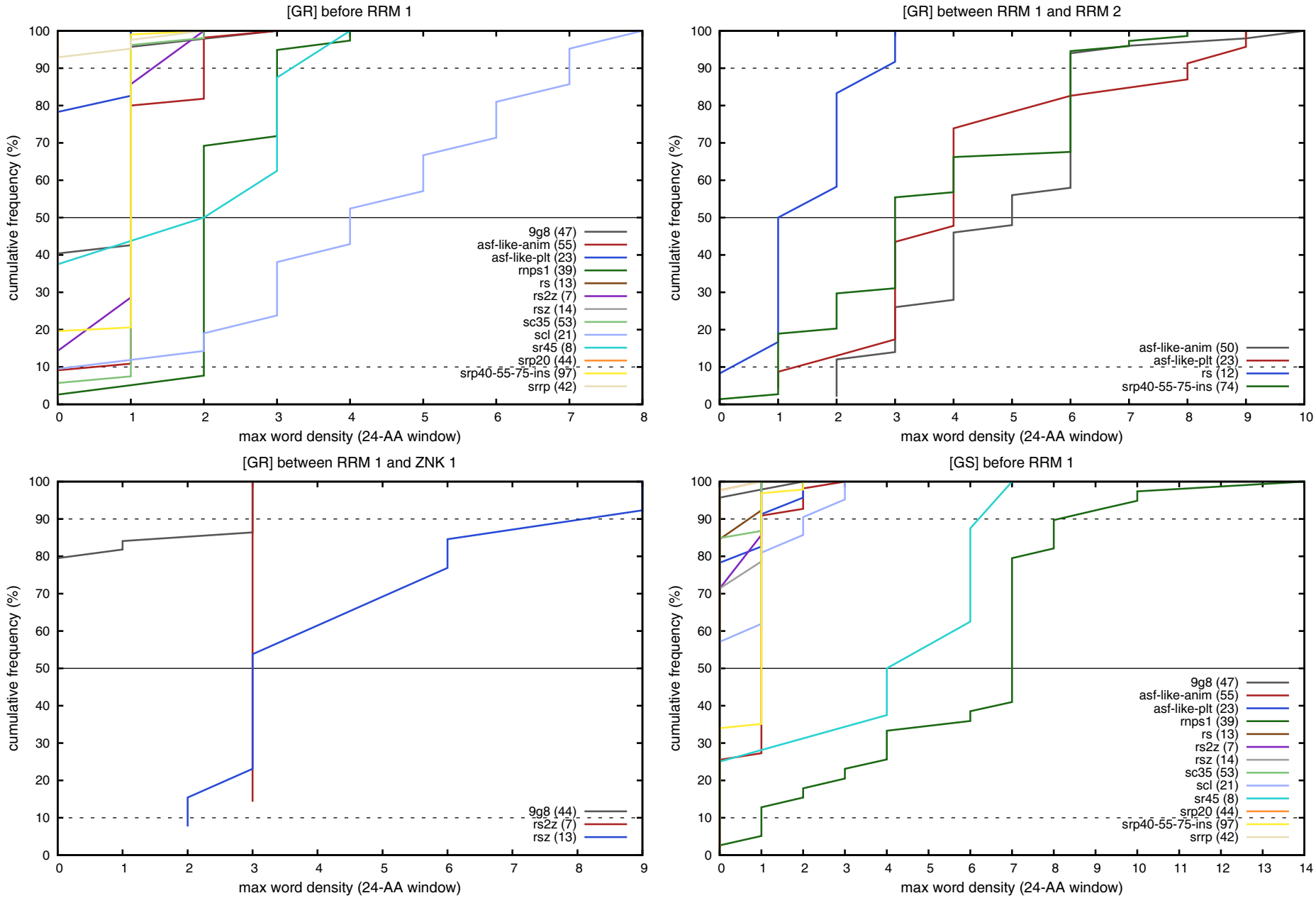
Word density cumulative distributions <plotdist> - page 4



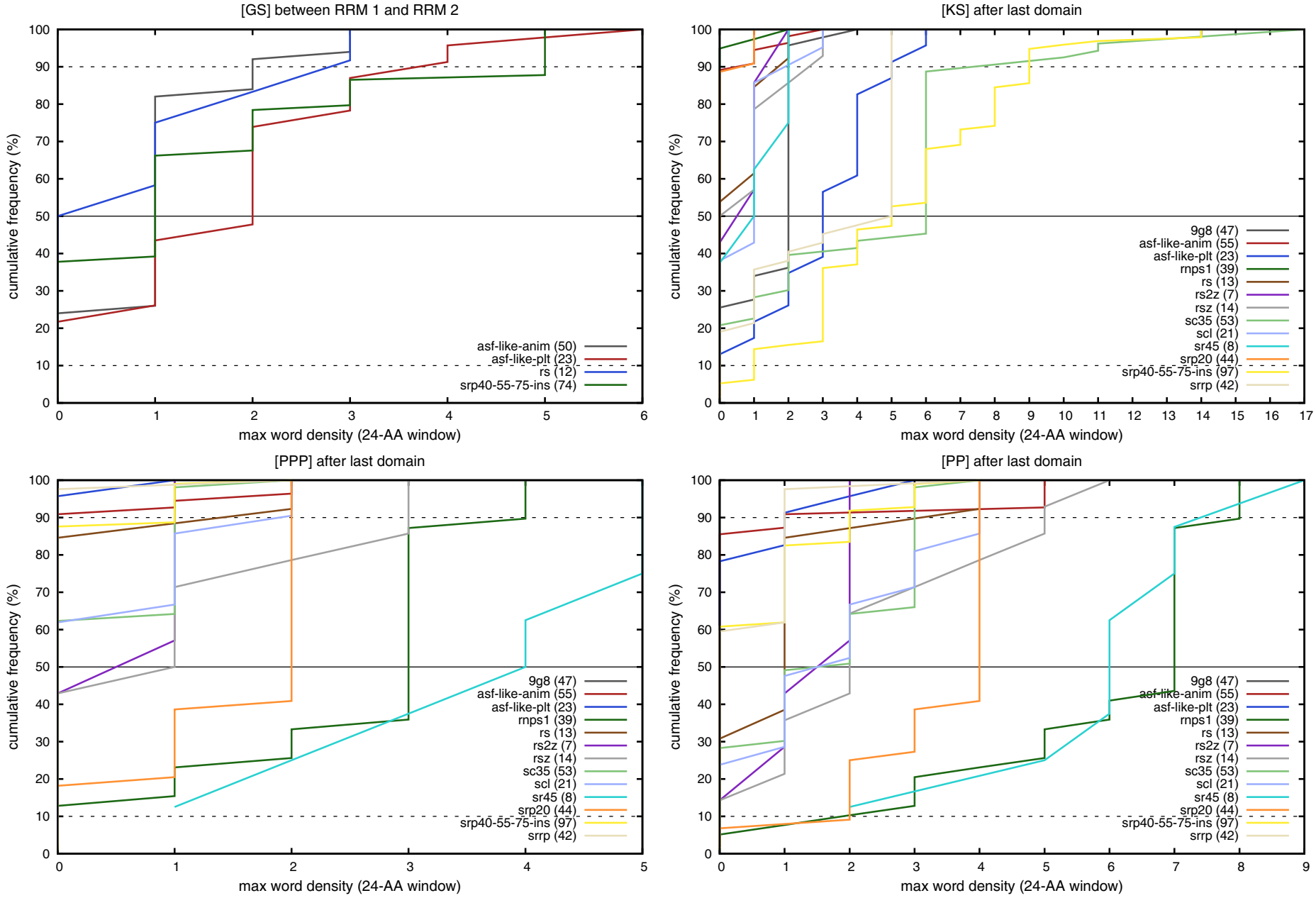
Word density cumulative distributions <plotdist> - page 5



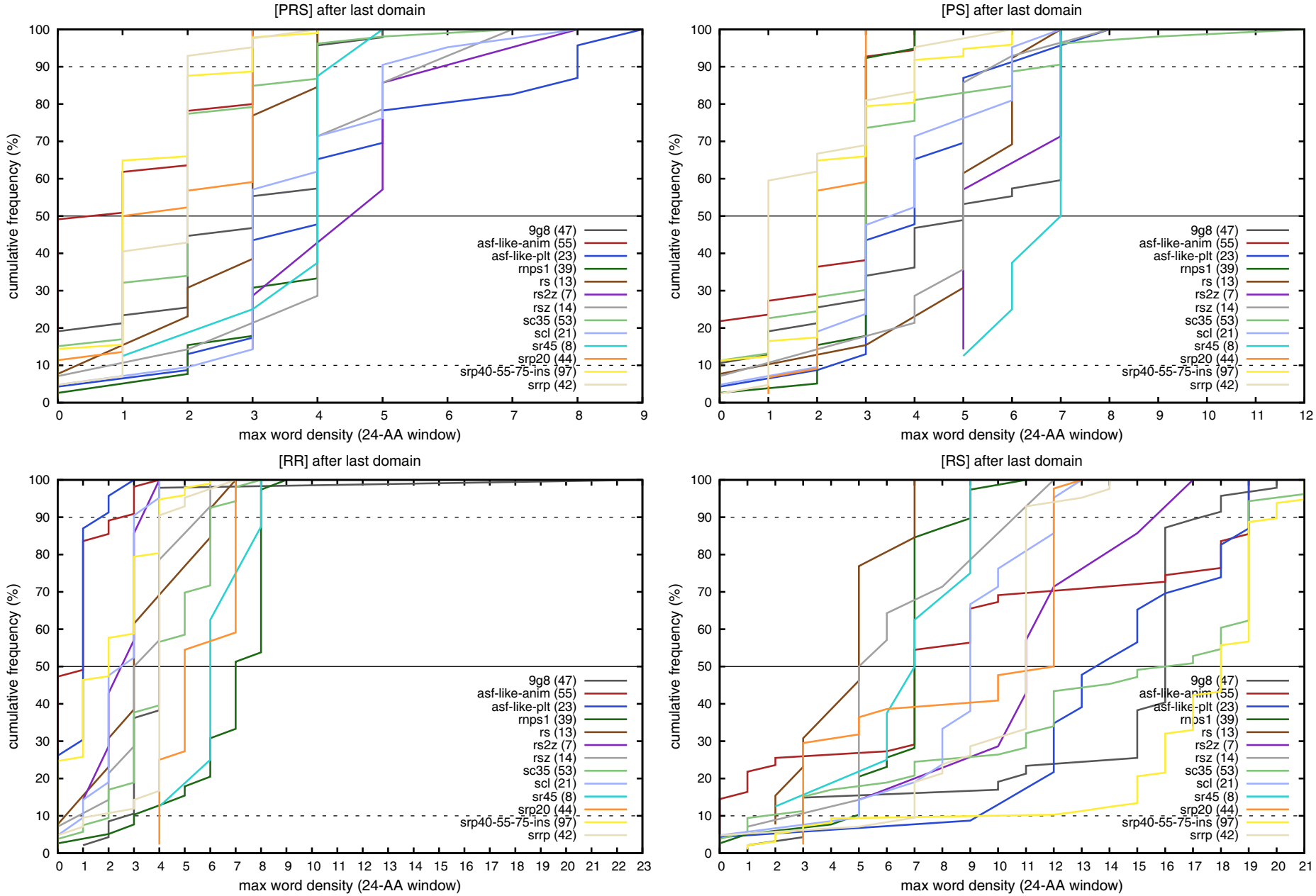
Word density cumulative distributions <plotdist> - page 6



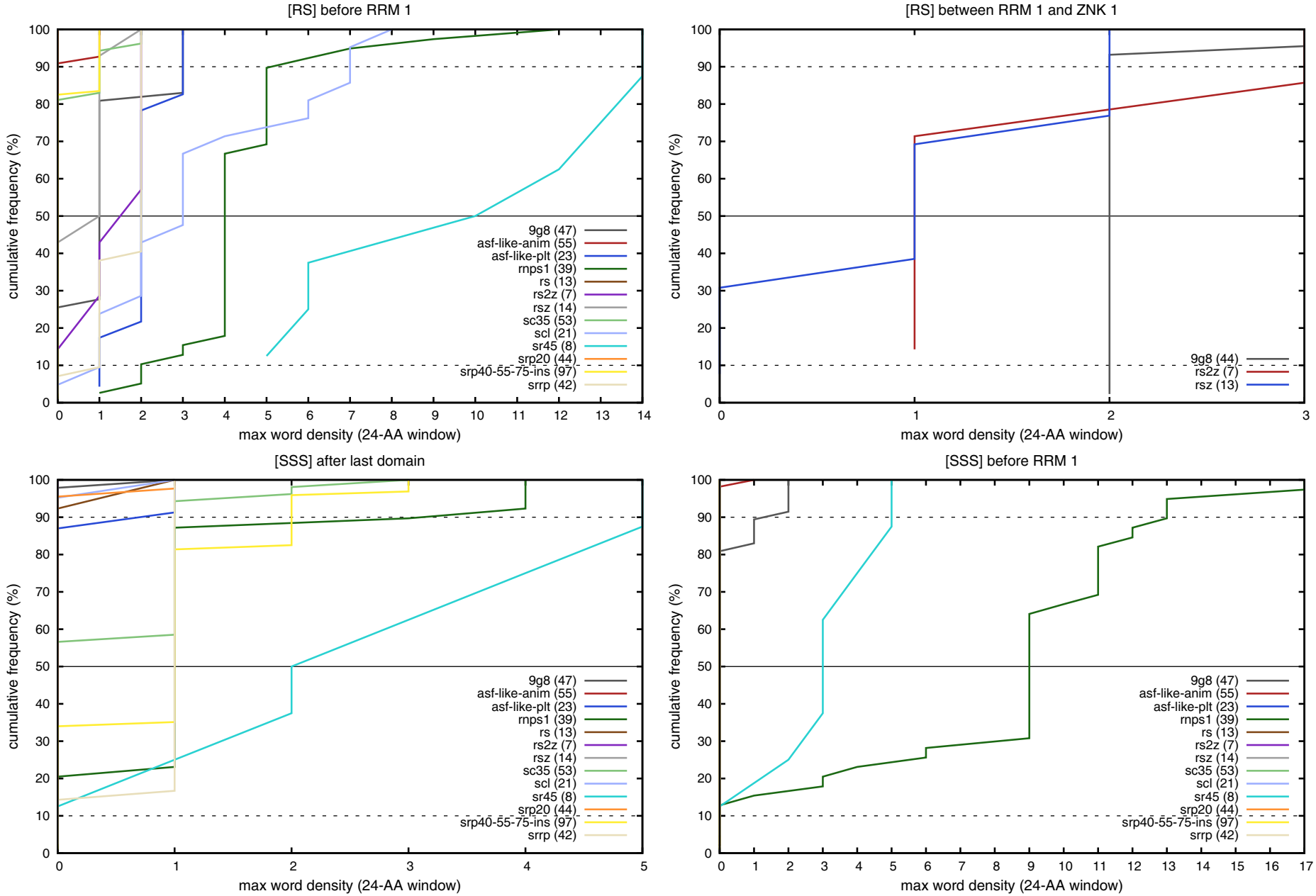
Word density cumulative distributions <plotdist> - page 7



Word density cumulative distributions <plotdist> - page 8



Word density cumulative distributions <plotdist> - page 9



Word density cumulative distributions <plotdist> - page 10

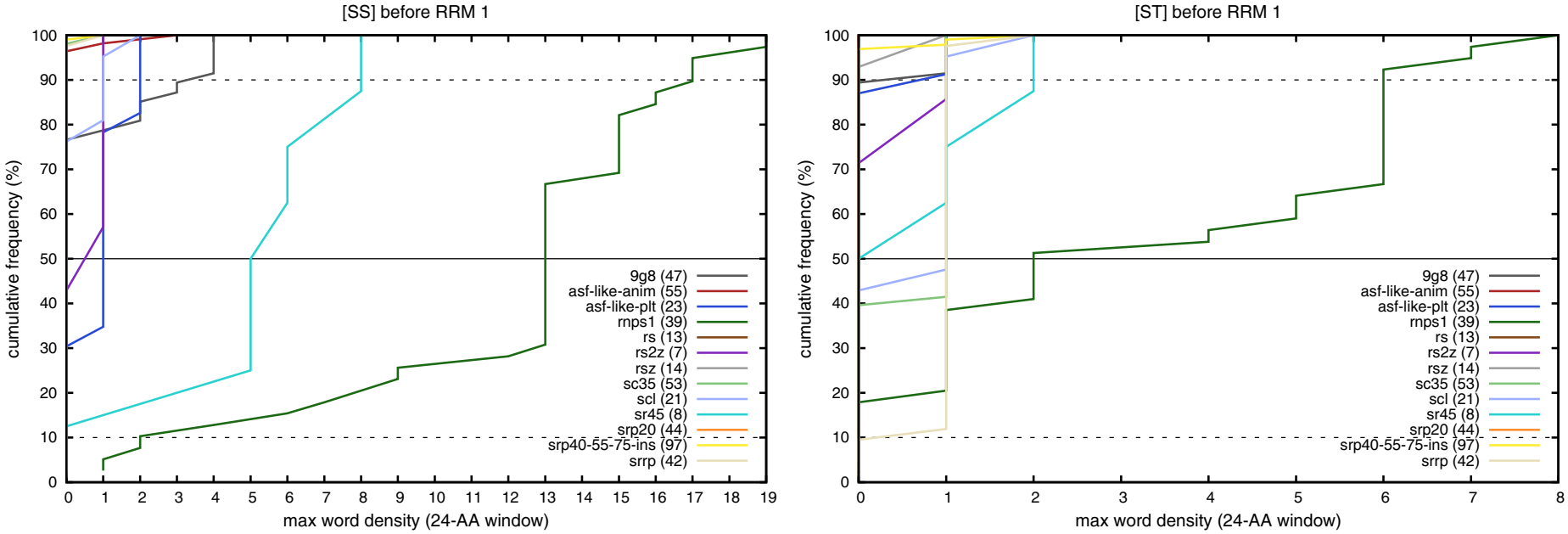


Fig. S25

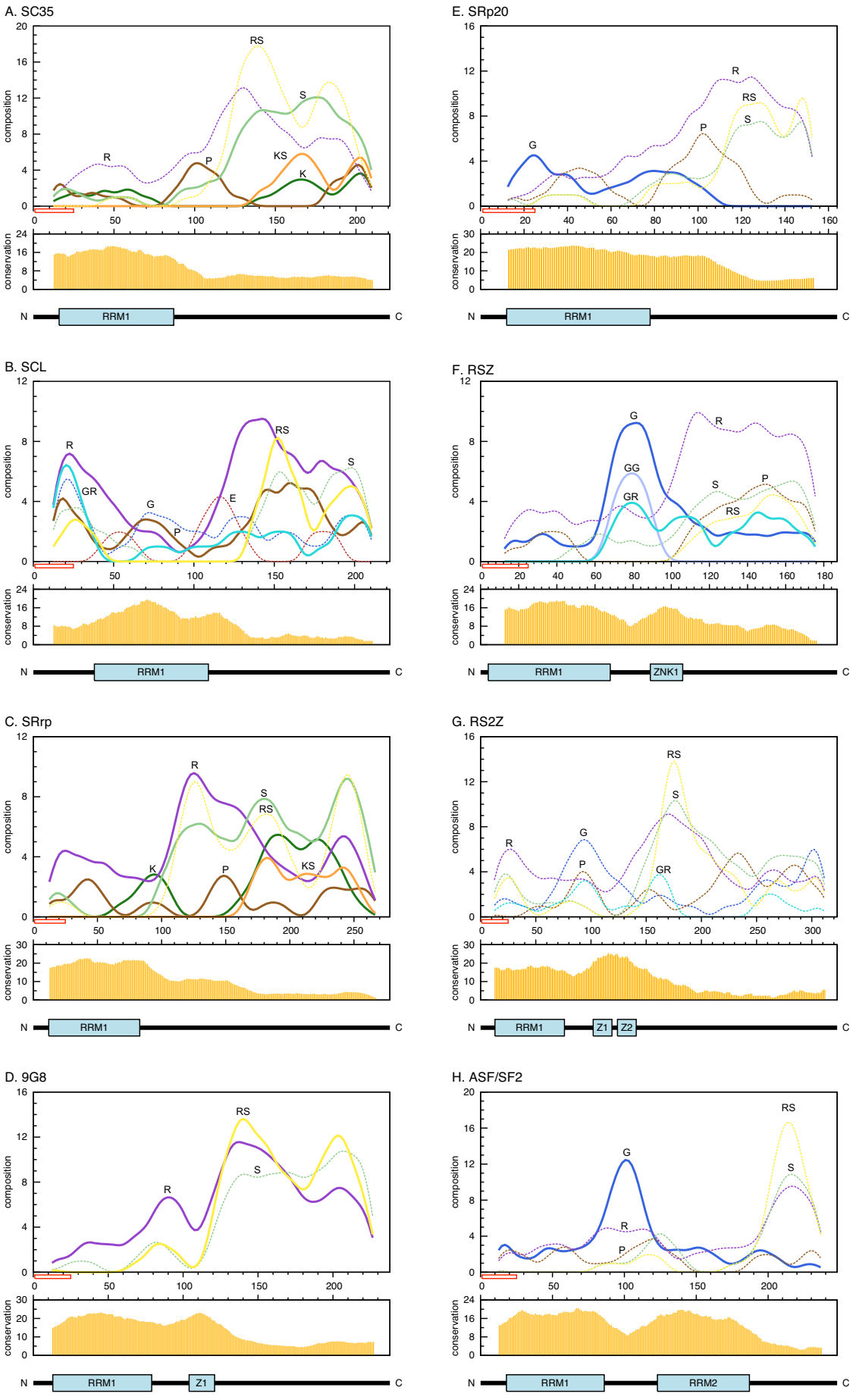

A. SC35
composition
conservation
RS
S
R
P
KS
K
N
RRM1
C
E. SRp20
G
R
RS
S
P
B. SCL
GR
G
E
F. RSZ
GG
ZNK1
C. SRrp
D. 9G8
G. RS2Z
Z1
Z2
H. ASF/SF2
RRM2

Fig. S25

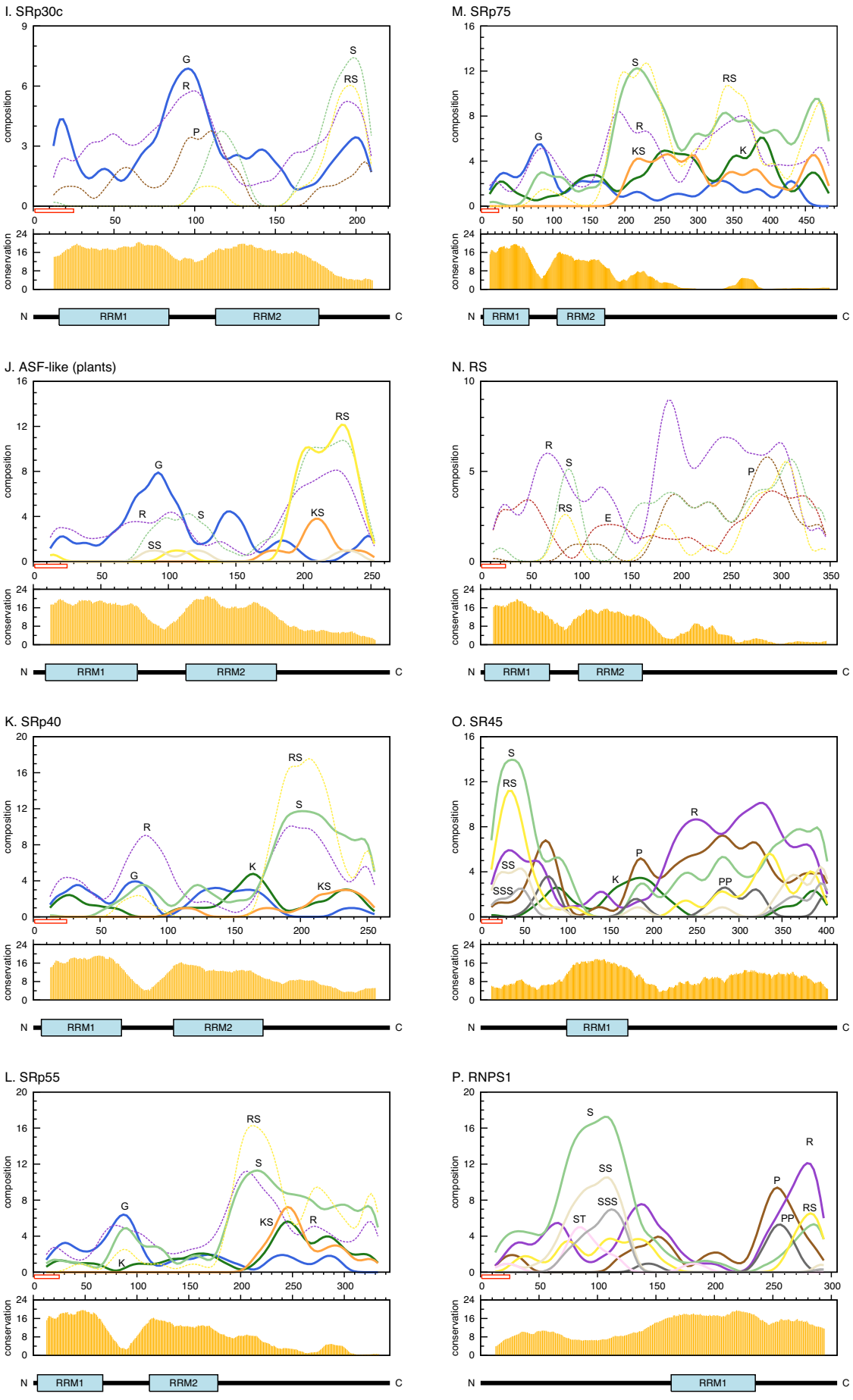

I. SRp30c
M. SRp75
J. ASF-like (plants)
N. RS
K. SRp40
O. SR45
L. SRp55
P. RNPS1
composition
conservation
N
C
RRM1
RRM2

Fig. S26

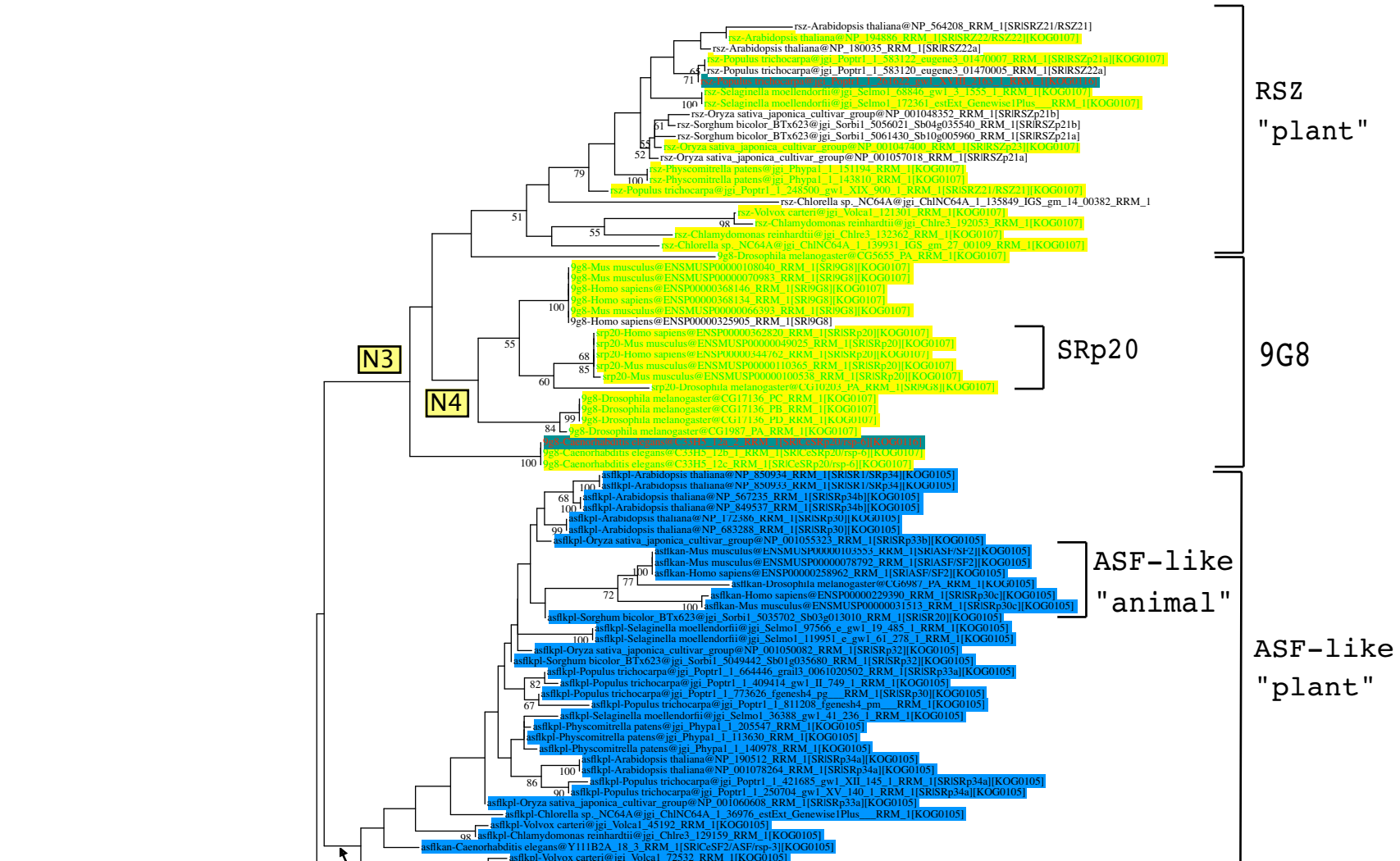

rsz-Arabidopsis thaliana@NP_564208_RRM_1[SRlSRZ21/RSZ21]
rsz-Arabidopsis thaliana@NP_194886_RRM_1[SRlSRZ22/RSZ22][KOG0107]
rsz-Arabidopsis thaliana@NP_180035_RRM_1[SRlRSZ22a]
rsz-Populus trichocarpa@jgi_Poptr1_1_583122_eugene3_01470007_RRM_1[SRlRSZp21a][KOG0107]
rsz-Populus trichocarpa@jgi_Poptr1_1_583120_eugene3_01470005_RRM_1[SRlRSZ22a]
rsz-Selaginella moellendorfii@jgi_Selmo1_68846_gw1_3_1555_1_RRM_1[KOG0107]
rsz-Selaginella moellendorfii@jgi_Selmo1_172361_estExt_Genewise1Plus___RRM_1[KOG0107]
rsz-Oryza sativa_japonica_cultivar_group@NP_001048352_RRM_1[SRlRSZp21b]
rsz-Sorghum bicolor_BTx623@jgi_Sorbi1_5056021_Sb04g035540_RRM_1[SRlRSZp21b]
rsz-Sorghum bicolor_BTx623@jgi_Sorbi1_5061430_Sb10g005960_RRM_1[SRlRSZp21a]
rsz-Oryza sativa_japonica_cultivar_group@NP_001047400_RRM_1[SRlRSZp23][KOG0107]
rsz-Oryza sativa_japonica_cultivar_group@NP_001057018_RRM_1[SRlRSZp21a]
rsz-Physcomitrella patens@jgi_Phypa1_1_151194_RRM_1[KOG0107]
rsz-Physcomitrella patens@jgi_Phypa1_1_143810_RRM_1[KOG0107]
rsz-Populus trichocarpa@jgi_Poptr1_1_248500_gw1_XIX_900_1_RRM_1[SRlSRZ21/RSZ21][KOG0107]
rsz-Chlorella sp._NC64A@jgi_ChlNC64A_1_135849_IGS_gm_14_00382_RRM_1
rsz-Volvox carteri@jgi_Volca1_121301_RRM_1[KOG0107]
rsz-Chlamydomonas reinhardtii@jgi_Chlre3_132362_RRM_1[KOG0107]
rsz-Chlorella sp._NC64A@jgi_ChlNC64A_1_139931_IGS_gm_27_00109_RRM_1[KOG0107]
9g8-Drosophila melanogaster@CG5655_PA_RRM_1[KOG0107]
9g8-Mus musculus@ENSMUSP00000108040_RRM_1[SRl9G8][KOG0107]
9g8-Mus musculus@ENSMUSP00000070983_RRM_1[SRl9G8][KOG0107]
9g8-Homo sapiens@ENSP00000368146_RRM_1[SRl9G8][KOG0107]
9g8-Homo sapiens@ENSP00000368134_RRM_1[SRl9G8][KOG0107]
9g8-Mus musculus@ENSMUSP00000066393_RRM_1[SRl9G8][KOG0107]
9g8-Homo sapiens@ENSP00000325905_RRM_1[SRl9G8]
srp20-Homo sapiens@ENSP00000362820_RRM_1[SRlSRp20][KOG0107]
srp20-Mus musculus@ENSMUSP00000049025_RRM_1[SRlSRp20][KOG0107]
srp20-Homo sapiens@ENSP00000344762_RRM_1[SRlSRp20][KOG0107]
srp20-Mus musculus@ENSMUSP00000110365_RRM_1[SRlSRp20][KOG0107]
srp20-Mus musculus@ENSMUSP00000100538_RRM_1[SRlSRp20][KOG0107]
srp20-Drosophila melanogaster@CG10203_PA_RRM_1[SRl9G8][KOG0107]
9g8-Drosophila melanogaster@CG17136_PC_RRM_1[KOG0107]
9g8-Drosophila melanogaster@CG17136_PB_RRM_1[KOG0107]
9g8-Drosophila melanogaster@CG17136_PD_RRM_1[KOG0107]
9g8-Drosophila melanogaster@CG1987_PA_RRM_1[KOG0107]
9g8-Caenorhabditis elegans@C33H5_12b_1_RRM_1[SRlCeSRp20/rsp-6][KOG0107]
9g8-Caenorhabditis elegans@C33H5_12c_RRM_1[SRlCeSRp20/rsp-6][KOG0107]
asflkpl-Arabidopsis thaliana@NP_850934_RRM_1[SRlSR1/SRp34][KOG0105]
asflkpl-Arabidopsis thaliana@NP_850933_RRM_1[SRlSR1/SRp34][KOG0105]
asflkpl-Arabidopsis thaliana@NP_567235_RRM_1[SRlSRp34b][KOG0105]
asflkpl-Arabidopsis thaliana@NP_849537_RRM_1[SRlSRp34b][KOG0105]
asflkpl-Arabidopsis thaliana@NP_172386_RRM_1[SRlSRp30][KOG0105]
asflkpl-Arabidopsis thaliana@NP_683288_RRM_1[SRlSRp30][KOG0105]
asflkpl-Oryza sativa_japonica_cultivar_group@NP_001055323_RRM_1[SRlSRp33b][KOG0105]
asflkan-Mus musculus@ENSMUSP00000103553_RRM_1[SRlASF/SF2][KOG0105]
asflkan-Mus musculus@ENSMUSP00000078792_RRM_1[SRlASF/SF2][KOG0105]
asflkan-Homo sapiens@ENSP00000258962_RRM_1[SRlASF/SF2][KOG0105]
asflkan-Drosophila melanogaster@CG6987_PA_RRM_1[KOG0105]
asflkan-Homo sapiens@ENSP00000229390_RRM_1[SRlSRp30c][KOG0105]
asflkan-Mus musculus@ENSMUSP00000031513_RRM_1[SRlSRp30c][KOG0105]
asflkpl-Sorghum bicolor_BTx623@jgi_Sorbi1_5035702_Sb03g013010_RRM_1[SRlSR20][KOG0105]
asflkpl-Selaginella moellendorfii@jgi_Selmo1_97566_e_gw1_19_485_1_RRM_1[KOG0105]
asflkpl-Selaginella moellendorfii@jgi_Selmo1_119951_e_gw1_61_278_1_RRM_1[KOG0105]
asflkpl-Oryza sativa_japonica_cultivar_group@NP_001050082_RRM_1[SRlSRp32][KOG0105]
asflkpl-Sorghum bicolor_BTx623@jgi_Sorbi1_5049442_Sb01g035680_RRM_1[SRlSRp32][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_664446_grail3_0061020502_RRM_1[SRlSRp33a][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_409414_gw1_II_749_1_RRM_1[KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_773626_fgenesh4_pg___RRM_1[SRlSRp30][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_811208_fgenesh4_pm___RRM_1[KOG0105]
asflkpl-Selaginella moellendorfii@jgi_Selmo1_36388_gw1_41_236_1_RRM_1[KOG0105]
asflkpl-Physcomitrella patens@jgi_Phypa1_1_205547_RRM_1[KOG0105]
asflkpl-Physcomitrella patens@jgi_Phypa1_1_113630_RRM_1[KOG0105]
asflkpl-Physcomitrella patens@jgi_Phypa1_1_140978_RRM_1[KOG0105]
asflkpl-Arabidopsis thaliana@NP_190512_RRM_1[SRlSRp34a][KOG0105]
asflkpl-Arabidopsis thaliana@NP_001078264_RRM_1[SRlSRp34a][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_421685_gw1_XII_145_1_RRM_1[SRlSRp34a][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_250704_gw1_XV_140_1_RRM_1[SRlSRp34a][KOG0105]
asflkpl-Oryza sativa_japonica_cultivar_group@NP_001060608_RRM_1[SRlSRp33a][KOG0105]
asflkpl-Chlorella sp._NC64A@jgi_ChlNC64A_1_36976_estExt_Genewise1Plus___RRM_1[KOG0105]
asflkpl-Volvox carteri@jgi_Volca1_45192_RRM_1[KOG0105]
asflkpl-Chlamydomonas reinhardtii@jgi_Chlre3_129159_RRM_1[KOG0105]
asflkan-Caenorhabditis elegans@Y111B2A_18_3_RRM_1[SRlCeSF2/ASF/rsp-3][KOG0105]
asflkpl-Volvox carteri@jgi_Volca1_72532_RRM_1[KOG0105]
N3
N4
RSZ
"plant"
9G8
SRp20
ASF-like
"animal"
ASF-like
"plant"

Fig. S26

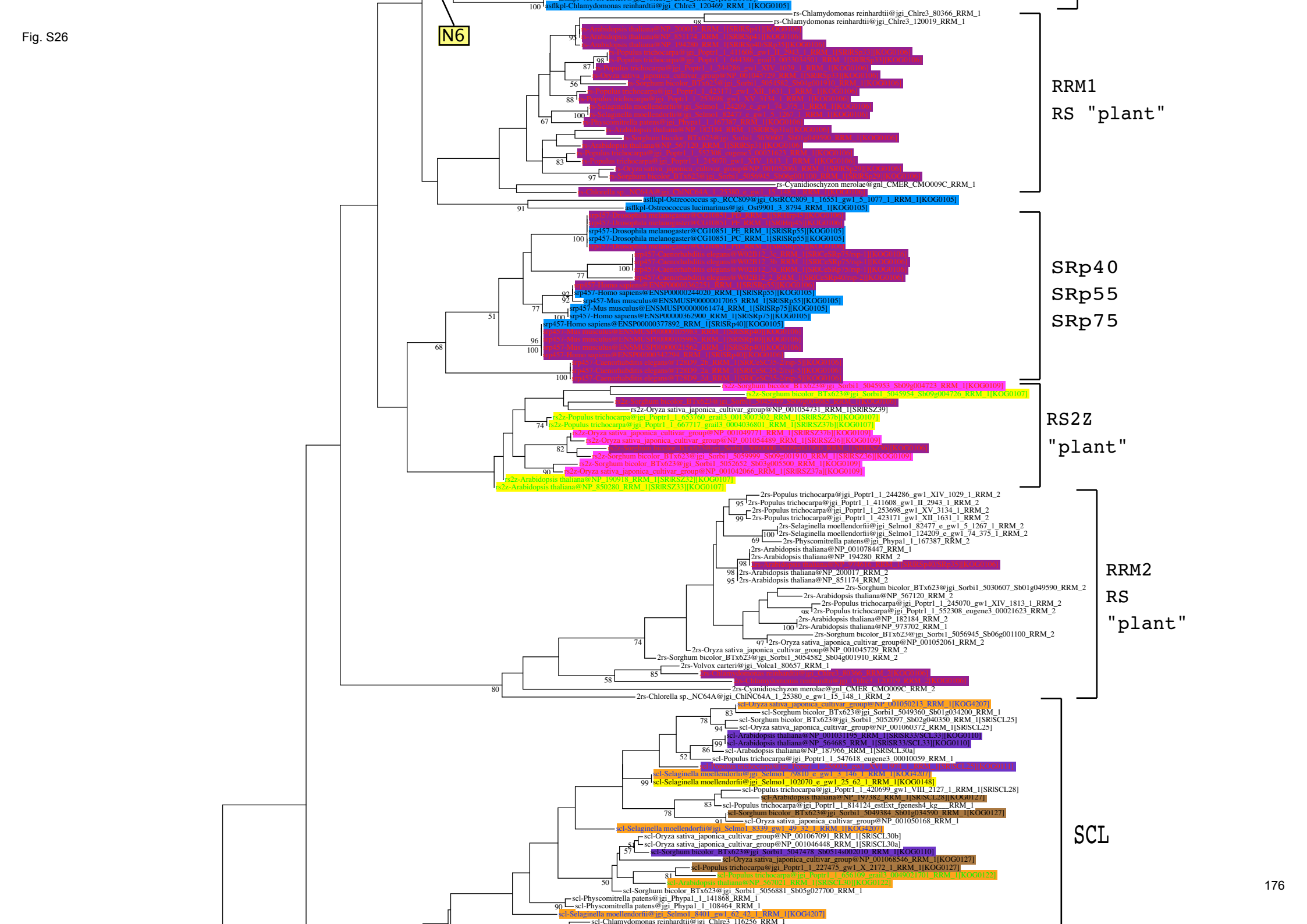

N6
RRM1
RS "plant"
SRp40
SRp55
SRp75
RS2Z
"plant"
RRM2
RS
"plant"
SCL

Fig. S26

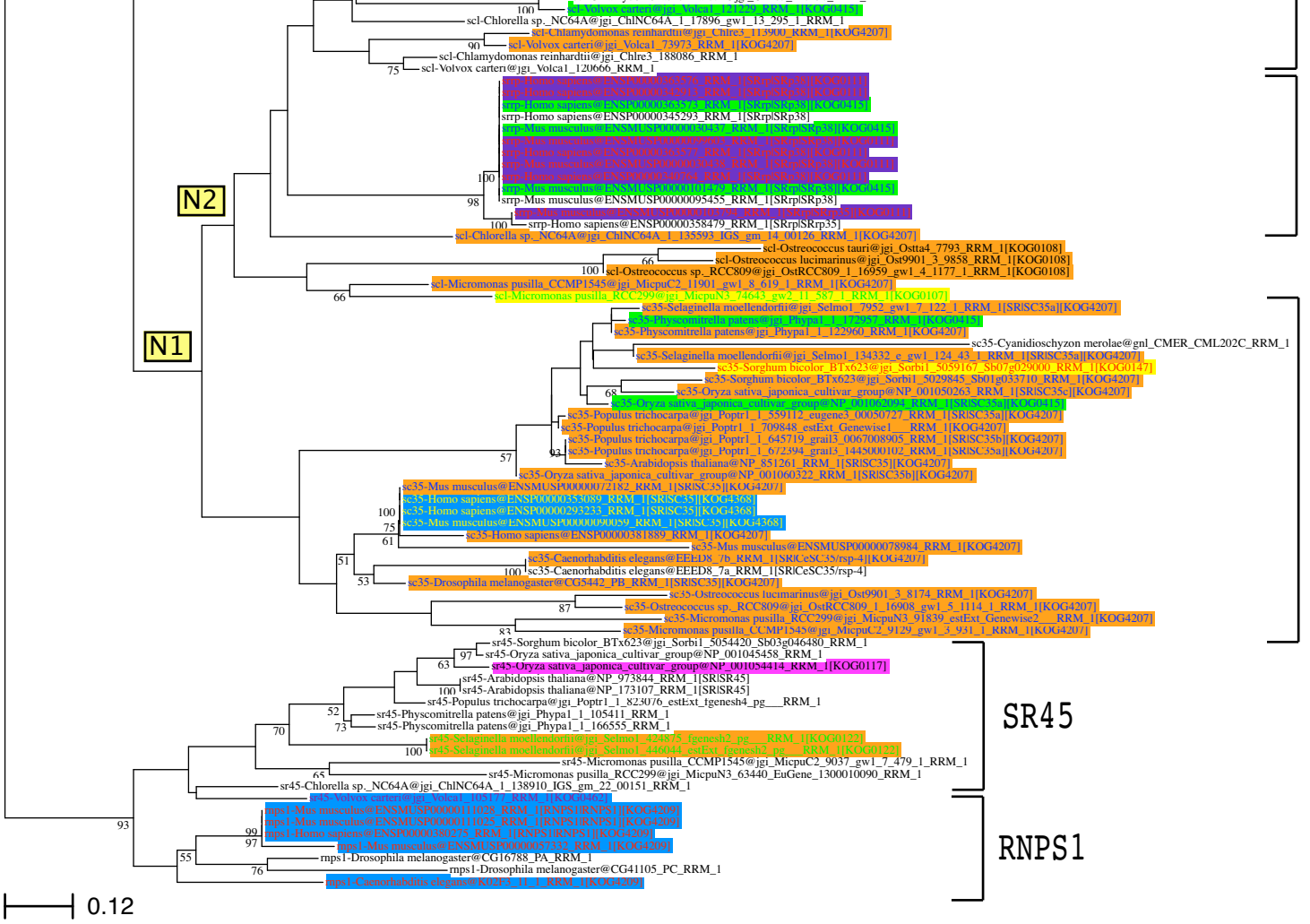

N2
N1
srrp-Homo sapiens@ENSP00000345293_RRM_1[SRrpISRp38]
srrp-Mus musculus@ENSMUSP00000095455_RRM_1[SRrpISRp38]
srrp-Homo sapiens@ENSP00000358479_RRM_1[SRrpISRp35]
scl-Chlorella sp._NC64A@jgi_ChlNC64A_1_135593_IGS_gm_14_00126_RRM_1[KOG4207]
scl-Ostreococcus tauri@jgi_Ostta4_7793_RRM_1[KOG0108]
scl-Ostreococcus lucimarinus@jgi_Ost9901_3_9858_RRM_1[KOG0108]
scl-Ostreococcus sp._RCC809@jgi_OstRCC809_1_16959_gw1_4_1177_1_RRM_1[KOG0108]
scl-Micromonas pusilla_CCMP1545@jgi_MicpuC2_11901_gw1_8_619_1_RRM_1[KOG4207]
scl-Micromonas pusilla_RCC299@jgi_MicpuN3_74643_gw2_11_587_1_RRM_1[KOG0107]
sc35-Selaginella moellendorfii@jgi_Selmo1_7952_gw1_7_122_1_RRM_1[SRiSC35a][KOG4207]
sc35-Physcomitrella patens@jgi_Phypa1_1_172957_RRM_1[KOG0415]
sc35-Physcomitrella patens@jgi_Phypa1_1_122960_RRM_1[KOG4207]
sc35-Cyanidioschyzon merolae@gnl_CMER_CML202C_RRM_1
sc35-Selaginella moellendorfii@jgi_Selmo1_134332_e_gw1_124_43_1_RRM_1[SRiSC35a][KOG4207]
sc35-Sorghum bicolor_BTx623@jgi_Sorbi1_5059167_Sb07g029000_RRM_1[KOG0147]
sc35-Sorghum bicolor_BTx623@jgi_Sorbi1_5029845_Sb01g033710_RRM_1[KOG4207]
sc35-Oryza sativa_japonica_cultivar_group@NP_001050263_RRM_1[SRiSC35c][KOG4207]
sc35-Oryza sativa_japonica_cultivar_group@NP_001062094_RRM_1[SRiSC35a][KOG0415]
sc35-Populus trichocarpa@jgi_Poptr1_1_559112_eugene3_00050727_RRM_1[SRiSC35a][KOG4207]
sc35-Populus trichocarpa@jgi_Poptr1_1_709848_estExt_Genewise1___RRM_1[KOG4207]
sc35-Populus trichocarpa@jgi_Poptr1_1_645719_grail3_0067008905_RRM_1[SRiSC35b][KOG4207]
sc35-Populus trichocarpa@jgi_Poptr1_1_672394_grail3_1445001102_RRM_1[SRiSC35a][KOG4207]
sc35-Arabidopsis thaliana@NP_851261_RRM_1[SRiSC35][KOG4207]
sc35-Oryza sativa_japonica_cultivar_group@NP_001060322_RRM_1[SRiSC35b][KOG4207]
sc35-Mus musculus@ENSMUSP00000072182_RRM_1[SRiSC35][KOG4207]
sc35-Homo sapiens@ENSP00000353089_RRM_1[SRiSC35][KOG4368]
sc35-Homo sapiens@ENSP00000293233_RRM_1[SRiSC35][KOG4368]
sc35-Mus musculus@ENSMUSP00000090059_RRM_1[SRiSC35][KOG4368]
sc35-Homo sapiens@ENSP00000381889_RRM_1[KOG4207]
sc35-Mus musculus@ENSMUSP00000078984_RRM_1[KOG4207]
sc35-Caenorhabditis elegans@EEED8_7b_RRM_1[SRiCeSC35/rsp-4][KOG4207]
sc35-Caenorhabditis elegans@EEED8_7a_RRM_1[SRiCeSC35/rsp-4]
sc35-Drosophila melanogaster@CG5442_PB_RRM_1[SRiSC35][KOG4207]
sc35-Ostreococcus lucimarinus@jgi_Ost9901_3_8174_RRM_1[KOG4207]
sc35-Ostreococcus sp._RCC809@jgi_OstRCC809_1_16908_gw1_5_1114_1_RRM_1[KOG4207]
sc35-Micromonas pusilla_RCC299@jgi_MicpuN3_91839_estExt_Genewise2___RRM_1[KOG4207]
sc35-Micromonas pusilla_CCMP1545@jgi_MicpuC2_9129_gw1_3_931_1_RRM_1[KOG4207]
sr45-Sorghum bicolor_BTx623@jgi_Sorbi1_5054420_Sb03g046480_RRM_1
sr45-Oryza sativa_japonica_cultivar_group@NP_001045458_RRM_1
sr45-Oryza sativa_japonica_cultivar_group@NP_001054414_RRM_1[KOG0117]
sr45-Arabidopsis thaliana@NP_973844_RRM_1[SRiSR45]
sr45-Arabidopsis thaliana@NP_173107_RRM_1[SRiSR45]
sr45-Populus trichocarpa@jgi_Poptr1_1_825076_estExt_fgenesh4_pg___RRM_1
sr45-Physcomitrella patens@jgi_Phypa1_1_105411_RRM_1
sr45-Physcomitrella patens@jgi_Phypa1_1_166555_RRM_1
sr45-Selaginella moellendorfii@jgi_Selmo1_424875_fgenesh2_pg___RRM_1[KOG0122]
sr45-Selaginella moellendorfii@jgi_Selmo1_446044_estExt_fgenesh2_pg___RRM_1[KOG0122]
sr45-Micromonas pusilla_CCMP1545@jgi_MicpuC2_9037_gw1_7_479_1_RRM_1
sr45-Micromonas pusilla_RCC299@jgi_MicpuN3_63440_EuGene_1300010090_RRM_1
sr45-Chlorella sp._NC64A@jgi_ChlNC64A_1_138910_IGS_gm_22_00151_RRM_1
sr45-Volvox carteri@jgi_Volca1_105177_RRM_1[KOG0117]
rnps1-Mus musculus@ENSMUSP00000111028_RRM_1[RNPS1SRNPS1][KOG4209]
rnps1-Mus musculus@ENSMUSP00000111025_RRM_1[RNPS1SRNPS1][KOG4209]
rnps1-Homo sapiens@ENSP00000380274_RRM_1[RNPS1SRNPS1][KOG4209]
rnps1-Mus musculus@ENSMUSP00000057533_RRM_1[KOG4209]
rnps1-Drosophila melanogaster@CG16788_PA_RRM_1
rnps1-Drosophila melanogaster@CG41105_PC_RRM_1
rnps1-Caenorhabditis elegans@K02F3_11_1_RRM_1[KOG4209]
SRrp
SC35
SR45
RNPS1
0.12

Fig. S27

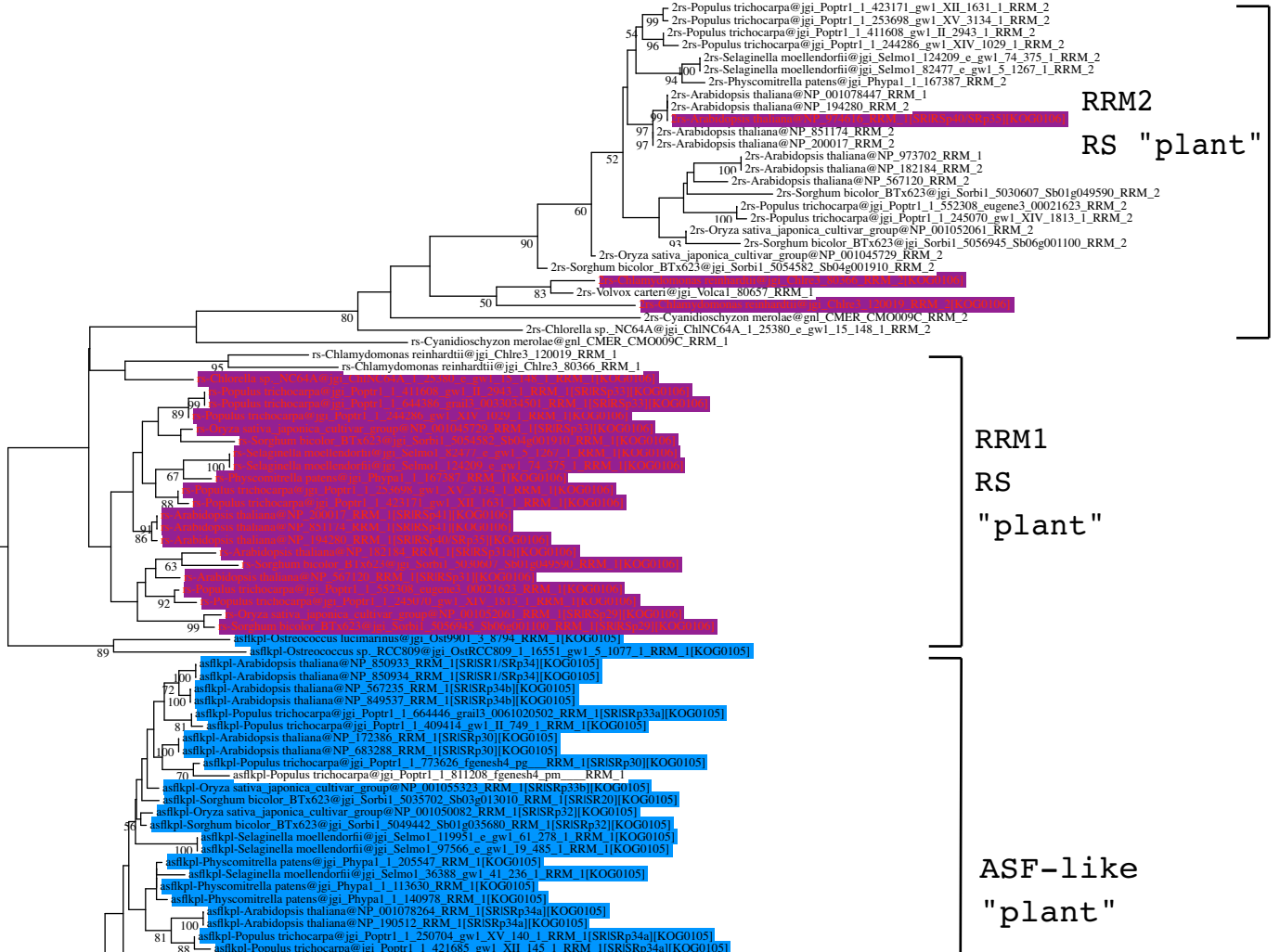

2rs-Populus trichocarpa@jgi_Poptr1_1_423171_gw1_XII_1631_1_RRM_2
2rs-Populus trichocarpa@jgi_Poptr1_1_253698_gw1_XV_3134_1_RRM_2
2rs-Populus trichocarpa@jgi_Poptr1_1_411608_gw1_II_2943_1_RRM_2
2rs-Populus trichocarpa@jgi_Poptr1_1_244286_gw1_XIV_1029_1_RRM_2
2rs-Selaginella moellendorfii@jgi_Selmo1_124209_e_gw1_74_375_1_RRM_2
2rs-Selaginella moellendorfii@jgi_Selmo1_82477_e_gw1_5_1267_1_RRM_2
2rs-Physcomitrella patens@jgi_Phypa1_1_167387_RRM_2
2rs-Arabidopsis thaliana@NP_001078447_RRM_1
2rs-Arabidopsis thaliana@NP_194280_RRM_2
2rs-Arabidopsis thaliana@NP_851174_RRM_2
2rs-Arabidopsis thaliana@NP_200017_RRM_2
2rs-Arabidopsis thaliana@NP_973702_RRM_1
2rs-Arabidopsis thaliana@NP_182184_RRM_2
2rs-Arabidopsis thaliana@NP_567120_RRM_2
2rs-Sorghum bicolor_BTx623@jgi_Sorbi1_5030607_Sb01g049590_RRM_2
2rs-Populus trichocarpa@jgi_Poptr1_1_552308_eugene3_00021623_RRM_2
2rs-Populus trichocarpa@jgi_Poptr1_1_245070_gw1_XIV_1813_1_RRM_2
2rs-Oryza sativa_japonica_cultivar_group@NP_001052061_RRM_2
2rs-Sorghum bicolor_BTx623@jgi_Sorbi1_5056945_Sb06g001100_RRM_2
2rs-Oryza sativa_japonica_cultivar_group@NP_001045729_RRM_2
2rs-Sorghum bicolor_BTx623@jgi_Sorbi1_5054582_Sb04g001910_RRM_2
2rs-Volvox carteri@jgi_Volca1_80657_RRM_1
2rs-Cyanidioschyzon merolae@gnl_CMER_CMO009C_RRM_2
2rs-Chlorella sp._NC64A@jgi_ChlNC64A_1_25380_e_gw1_15_148_1_RRM_2
rs-Cyanidioschyzon merolae@gnl_CMER_CMO009C_RRM_1
rs-Chlamydomonas reinhardtii@jgi_Chlre3_120019_RRM_1
rs-Chlamydomonas reinhardtii@jgi_Chlre3_80366_RRM_1
asflkpl-Ostreococcus lucimarinus@jgi_Ost9901_3_8794_RRM_1[KOG0105]
asflkpl-Ostreococcus sp._RCC809@jgi_OstRCC809_1_16551_gw1_5_1077_1_RRM_1[KOG0105]
asflkpl-Arabidopsis thaliana@NP_850933_RRM_1[SRlSR1/SRp34][KOG0105]
asflkpl-Arabidopsis thaliana@NP_850934_RRM_1[SRlSR1/SRp34][KOG0105]
asflkpl-Arabidopsis thaliana@NP_567235_RRM_1[SRlSRp34b][KOG0105]
asflkpl-Arabidopsis thaliana@NP_849537_RRM_1[SRlSRp34b][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_664446_grail3_0061020502_RRM_1[SRlSRp33a][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_409414_gw1_II_749_1_RRM_1[KOG0105]
asflkpl-Arabidopsis thaliana@NP_172386_RRM_1[SRlSRp30][KOG0105]
asflkpl-Arabidopsis thaliana@NP_683288_RRM_1[SRlSRp30][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_773626_fgenesh4_pg___RRM_1[SRlSRp30][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_811208_fgenesh4_pm___RRM_1
asflkpl-Oryza sativa_japonica_cultivar_group@NP_001055323_RRM_1[SRlSRp33b][KOG0105]
asflkpl-Sorghum bicolor_BTx623@jgi_Sorbi1_5035702_Sb03g013010_RRM_1[SRlSR20][KOG0105]
asflkpl-Oryza sativa_japonica_cultivar_group@NP_001050082_RRM_1[SRlSRp32][KOG0105]
asflkpl-Sorghum bicolor_BTx623@jgi_Sorbi1_5049442_Sb01g035680_RRM_1[SRlSRp32][KOG0105]
asflkpl-Selaginella moellendorfii@jgi_Selmo1_119951_e_gw1_61_278_1_RRM_1[KOG0105]
asflkpl-Selaginella moellendorfii@jgi_Selmo1_97566_e_gw1_19_485_1_RRM_1[KOG0105]
asflkpl-Physcomitrella patens@jgi_Phypa1_1_205547_RRM_1[KOG0105]
asflkpl-Selaginella moellendorfii@jgi_Selmo1_36388_gw1_41_236_1_RRM_1[KOG0105]
asflkpl-Physcomitrella patens@jgi_Phypa1_1_113630_RRM_1[KOG0105]
asflkpl-Physcomitrella patens@jgi_Phypa1_1_140978_RRM_1[KOG0105]
asflkpl-Arabidopsis thaliana@NP_001078264_RRM_1[SRlSRp34a][KOG0105]
asflkpl-Arabidopsis thaliana@NP_190512_RRM_1[SRlSRp34a][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_250704_gw1_XV_140_1_RRM_1[SRlSRp34a][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_421685_gw1_XII_145_1_RRM_1[SRlSRp34a][KOG0105]
RRM2
RS "plant"
RRM1
RS
"plant"
ASF-like
"plant"

Fig. S27

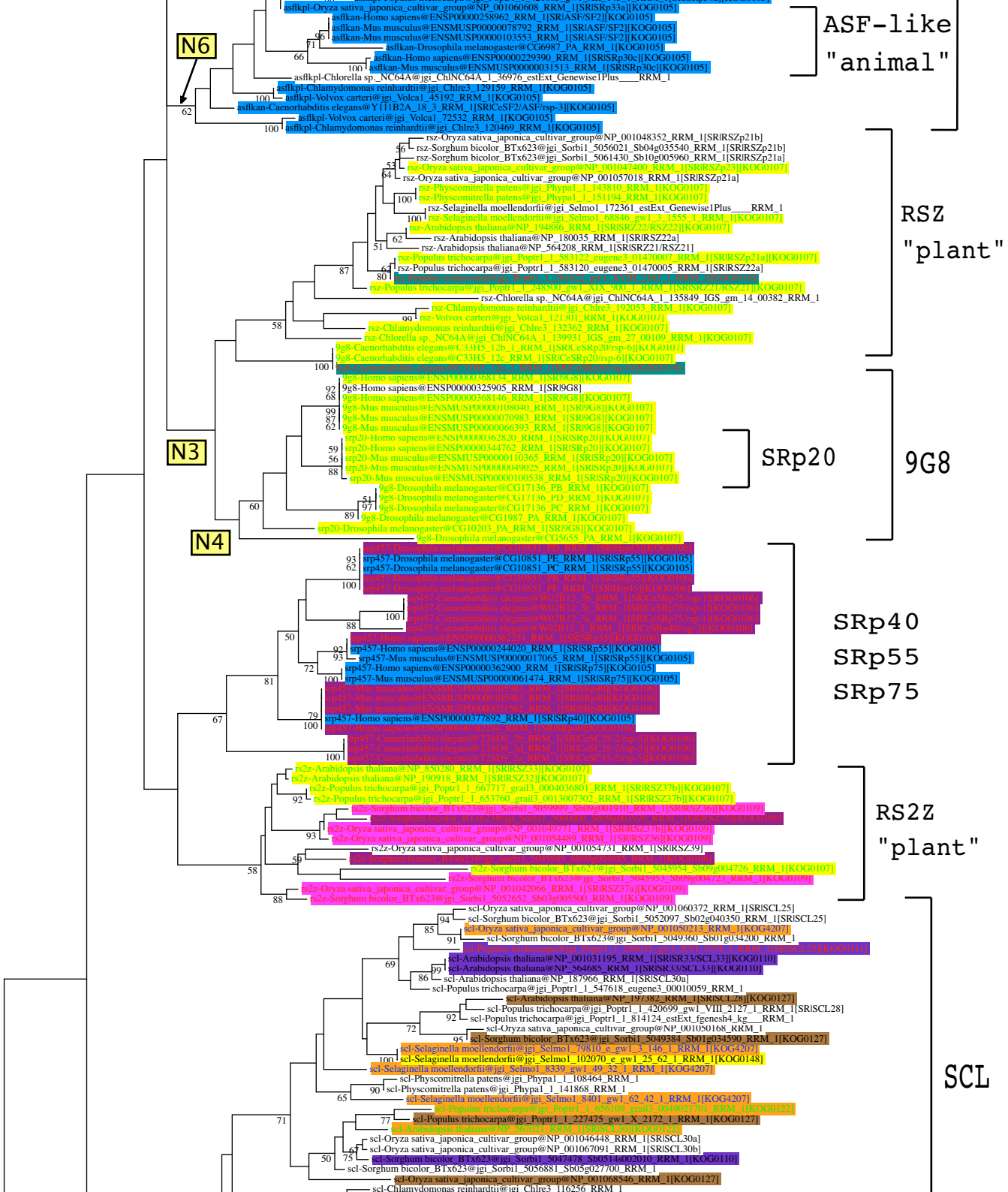

N6
N3
N4
ASF-like
"animal"
RSZ
"plant"
9G8
SRp20
SRp40
SRp55
SRp75
RS2Z
"plant"
SCL

Fig. S27

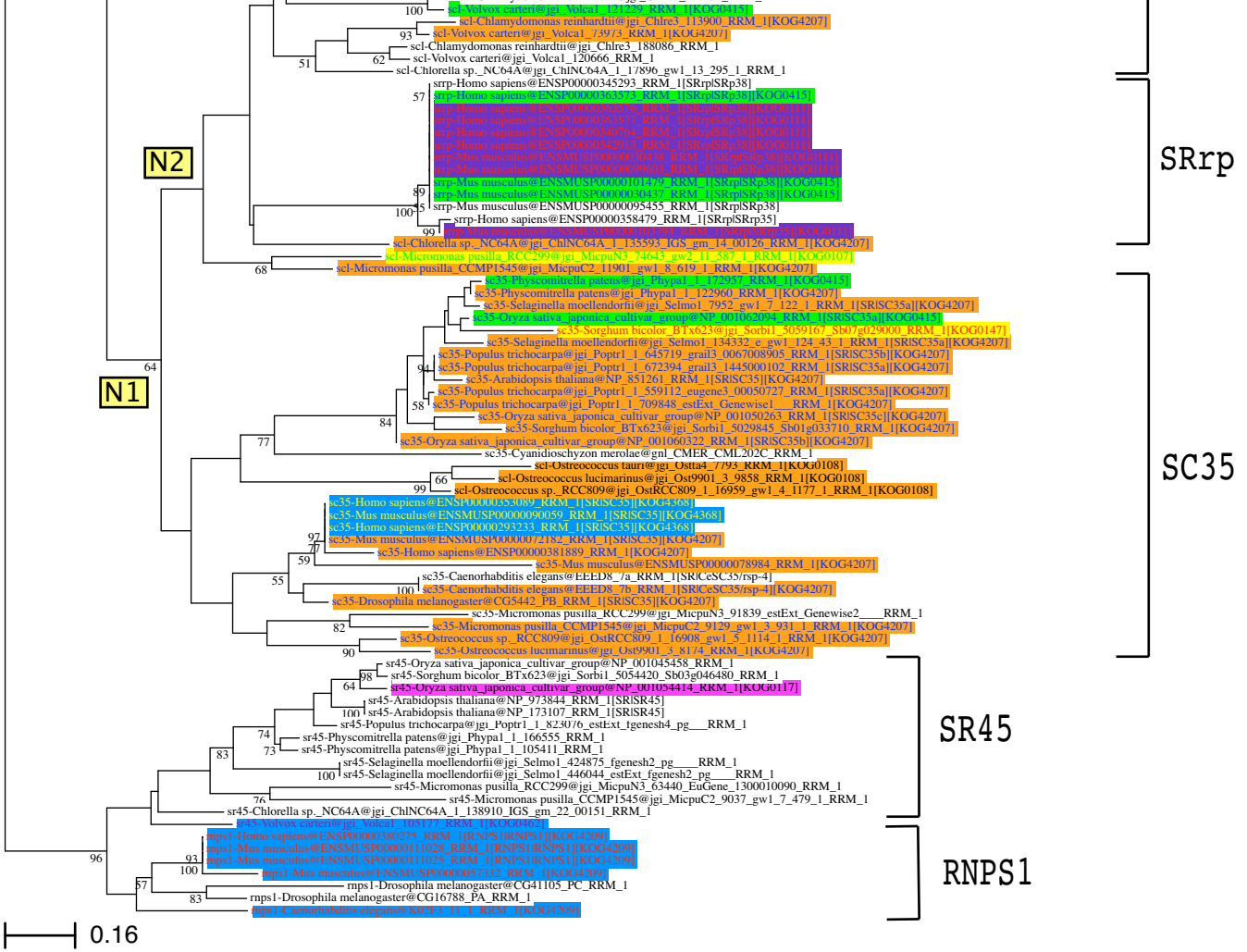


Fig. S28

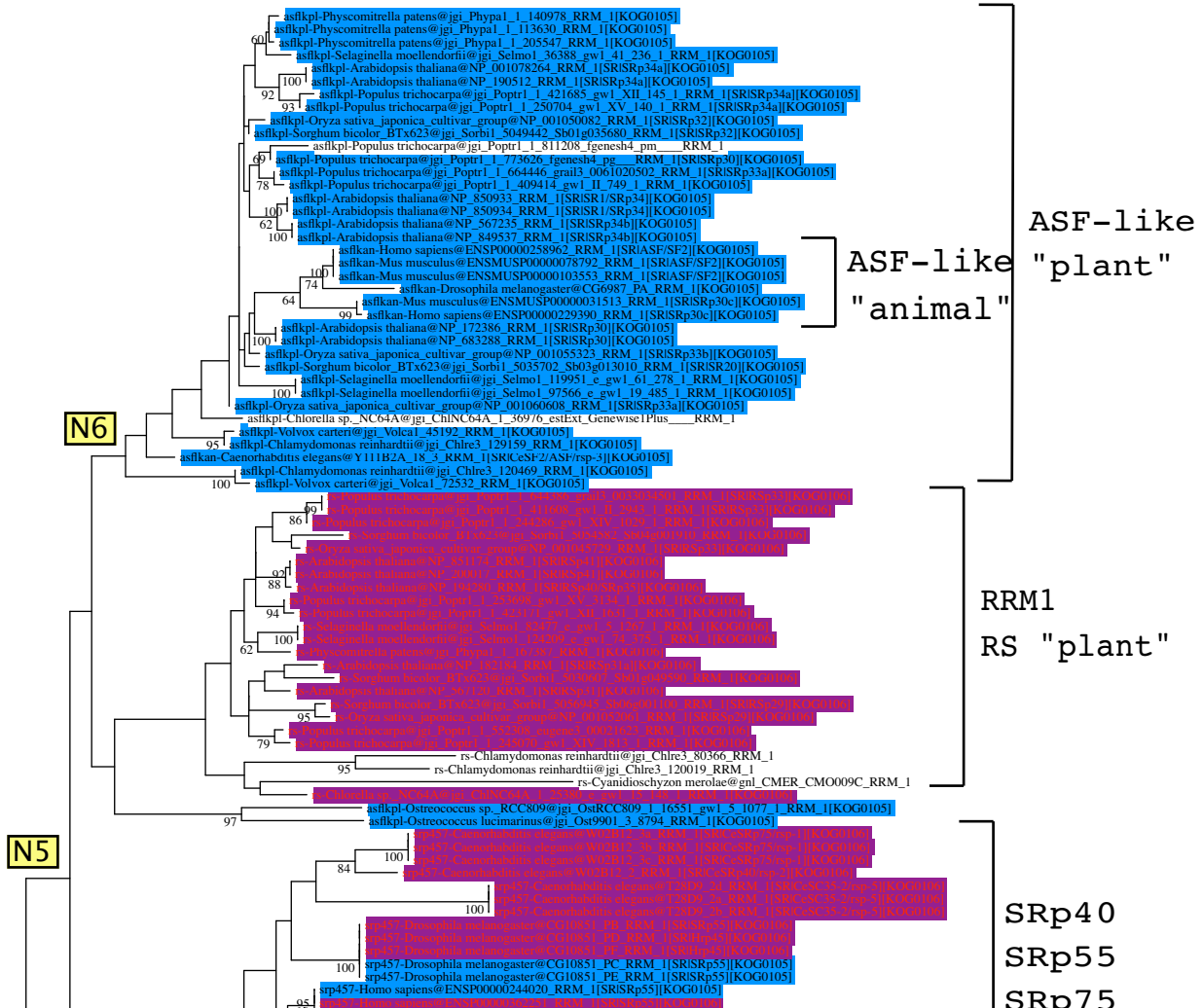

asflkpl-Physcomitrella patens@jgi_Phypa1_1_140978_RRM_1[KOG0105]
asflkpl-Physcomitrella patens@jgi_Phypa1_1_113630_RRM_1[KOG0105]
asflkpl-Physcomitrella patens@jgi_Phypa1_1_205547_RRM_1[KOG0105]
asflkpl-Selaginella moellendorfii@jgi_Selmo1_36388_gw1_41_236_1_RRM_1[KOG0105]
asflkpl-Arabidopsis thaliana@NP_001078264_RRM_1[SRlSRp34a][KOG0105]
asflkpl-Arabidopsis thaliana@NP_190512_RRM_1[SRlSRp34a][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_421685_gw1_XII_145_1_RRM_1[SRlSRp34a][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_250704_gw1_XV_140_1_RRM_1[SRlSRp34a][KOG0105]
asflkpl-Oryza sativa_japonica_cultivar_group@NP_001050082_RRM_1[SRlSRp32][KOG0105]
asflkpl-Sorghum bicolor_BTx623@jgi_Sorbi1_5049442_Sb01g035680_RRM_1[SRlSRp32][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_811208_fgenesh4_pm____RRM_1
asflkpl-Populus trichocarpa@jgi_Poptr1_1_773626_fgenesh4_pg____RRM_1[SRlSRp30][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_664446_grail3_0061020502_RRM_1[SRlSRp33a][KOG0105]
asflkpl-Populus trichocarpa@jgi_Poptr1_1_409414_gw1_II_749_1_RRM_1[KOG0105]
asflkpl-Arabidopsis thaliana@NP_850933_RRM_1[SRlSR1/SRp34][KOG0105]
asflkpl-Arabidopsis thaliana@NP_850934_RRM_1[SRlSR1/SRp34][KOG0105]
asflkpl-Arabidopsis thaliana@NP_567235_RRM_1[SRlSRp34b][KOG0105]
asflkpl-Arabidopsis thaliana@NP_849537_RRM_1[SRlSRp34b][KOG0105]
asflkan-Homo sapiens@ENSP00000258962_RRM_1[SRlASF/SF2][KOG0105]
asflkan-Mus musculus@ENSMUSP00000078792_RRM_1[SRlASF/SF2][KOG0105]
asflkan-Mus musculus@ENSMUSP00000103553_RRM_1[SRlASF/SF2][KOG0105]
asflkan-Drosophila melanogaster@CG6987_PA_RRM_1[KOG0105]
asflkan-Mus musculus@ENSMUSP00000031513_RRM_1[SRlSRp30c][KOG0105]
asflkan-Homo sapiens@ENSP00000229390_RRM_1[SRlSRp30c][KOG0105]
asflkpl-Arabidopsis thaliana@NP_172386_RRM_1[SRlSRp30][KOG0105]
asflkpl-Arabidopsis thaliana@NP_683288_RRM_1[SRlSRp30][KOG0105]
asflkpl-Oryza sativa_japonica_cultivar_group@NP_001055323_RRM_1[SRlSRp33b][KOG0105]
asflkpl-Sorghum bicolor_BTx623@jgi_Sorbi1_5035702_Sb03g013010_RRM_1[SRlSR20][KOG0105]
asflkpl-Selaginella moellendorfii@jgi_Selmo1_119951_e_gw1_61_278_1_RRM_1[KOG0105]
asflkpl-Selaginella moellendorfii@jgi_Selmo1_97566_e_gw1_19_485_1_RRM_1[KOG0105]
asflkpl-Oryza sativa_japonica_cultivar_group@NP_001060608_RRM_1[SRlSRp33a][KOG0105]
asflkpl-Chlorella sp._NC64A@jgi_ChlNC64A_1_36976_estExt_Genewise1Plus____RRM_1
asflkpl-Volvox carteri@jgi_Volca1_45192_RRM_1[KOG0105]
asflkpl-Chlamydomonas reinhardtii@jgi_Chlre3_129159_RRM_1[KOG0105]
asflkan-Caenorhabditis elegans@Y111B2A_18_5_RRM_1[SRlCeSF2/ASF/rsp-5][KOG0105]
asflkpl-Chlamydomonas reinhardtii@jgi_Chlre3_120469_RRM_1[KOG0105]
asflkpl-Volvox carteri@jgi_Volca1_72532_RRM_1[KOG0105]
rs-Chlamydomonas reinhardtii@jgi_Chlre3_80366_RRM_1
rs-Chlamydomonas reinhardtii@jgi_Chlre3_120019_RRM_1
rs-Cyanidioschyzon merolae@gnl_CMER_CMO009C_RRM_1
asflkpl-Ostreococcus sp._RCC809@jgi_OstRCC809_1_16551_gw1_5_1077_1_RRM_1[KOG0105]
asflkpl-Ostreococcus lucimarinus@jgi_Ost9901_3_8794_RRM_1[KOG0105]
srp457-Drosophila melanogaster@CG10851_PC_RRM_1[SRlSRp55][KOG0105]
srp457-Drosophila melanogaster@CG10851_PE_RRM_1[SRlSRp55][KOG0105]
srp457-Homo sapiens@ENSP00000244020_RRM_1[SRlSRp55][KOG0105]
N6
N5
ASF-like "plant"
ASF-like "animal"
RRM1 RS "plant"
SRp40
SRp55
SRp75

Fig. S28

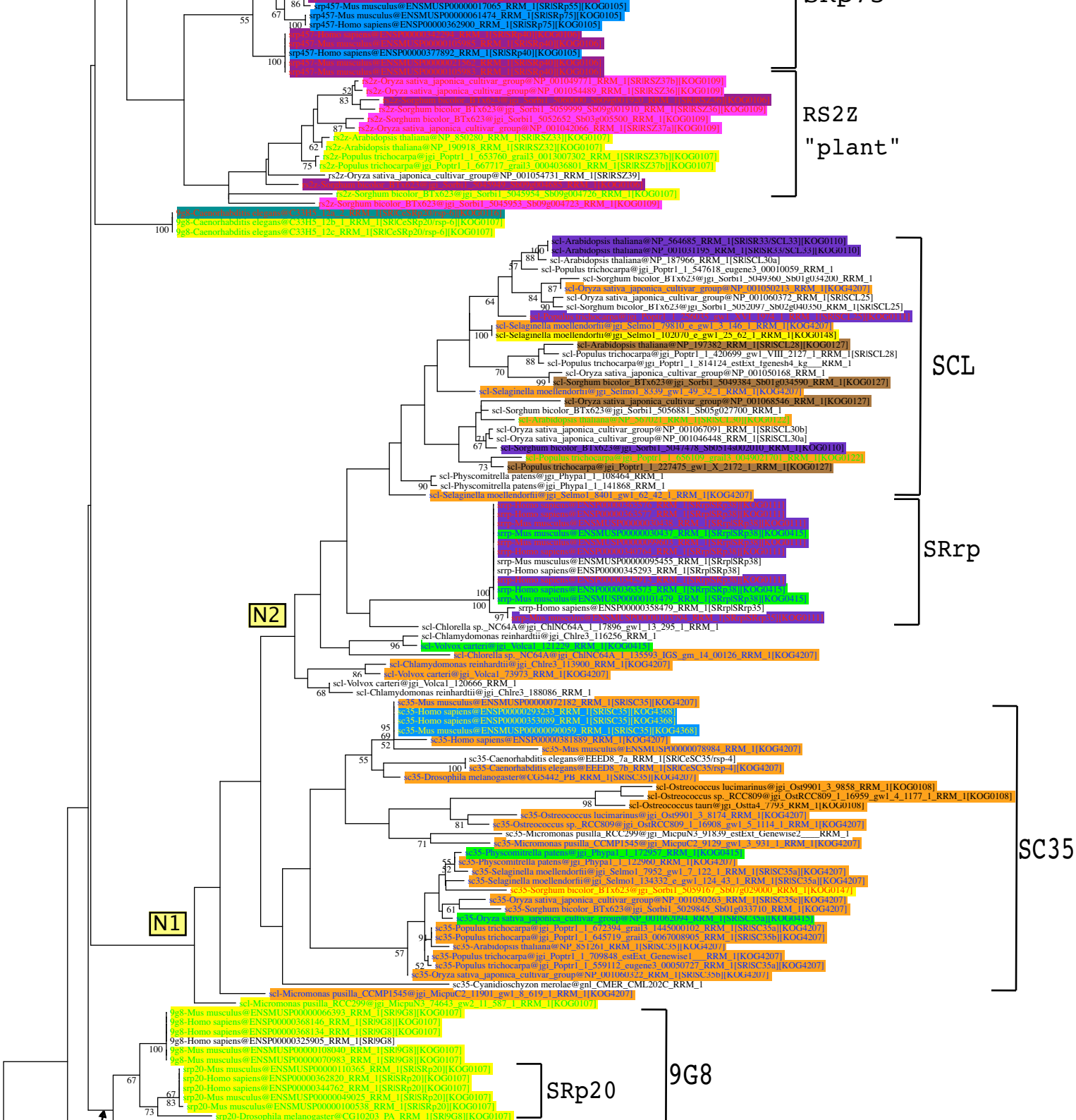

srp457-Mus musculus@ENSMUSP00000017065_RRM_1[SRlSRp55][KOG01051]
srp457-Mus musculus@ENSMUSP00000061474_RRM_1[SRlSRp75][KOG0105]
srp457-Homo sapiens@ENSP00000362900_RRM_1[SRlSRp75][KOG01051]
srp457-Homo sapiens@ENSP00000377892_RRM_1[SRlSRp40][KOG0105]
rs2z-Oryza sativa_japonica_cultivar_group@NP_001049771_RRM_1[SRlRSZ37b][KOG0109]
rs2z-Oryza sativa_japonica_cultivar_group@NP_001054489_RRM_1[SRlRSZ36][KOG0109]
rs2z-Sorghum bicolor_BTx623@jgi_Sorbi1_5059999_Sb09g001910_RRM_1[SRlRSZ36][KOG0109]
rs2z-Sorghum bicolor_BTx623@jgi_Sorbi1_5052652_Sb03g005500_RRM_1[KOG0109]
rs2z-Oryza sativa_japonica_cultivar_group@NP_001012066_RRM_1[SRlRSZ37a][KOG0109]
rs2z-Arabidopsis thaliana@NP_850280_RRM_1[SRlRSZ33][KOG0107]
rs2z-Arabidopsis thaliana@NP_190918_RRM_1[SRlRSZ32][KOG0107]
rs2z-Populus trichocarpa@jgi_Poptr1_1_653760_grail3_0013007302_RRM_1[SRlRSZ37b][KOG0107]
rs2z-Populus trichocarpa@jgi_Poptr1_1_667717_grail3_0004036801_RRM_1[SRlRSZ37b][KOG0107]
rs2z-Oryza sativa_japonica_cultivar_group@NP_001054731_RRM_1[SRlRSZ39]
rs2z-Sorghum bicolor_BTx623@jgi_Sorbi1_5045954_Sb09g004726_RRM_1[KOG0107]
rs2z-Sorghum bicolor_BTx623@jgi_Sorbi1_5045953_Sb09g004723_RRM_1[KOG0109]
RS2Z "plant"
9g8-Caenorhabditis elegans@C33H5_12b_1_RRM_1[SRlCeSRp20/rsp-6][KOG0107]
9g8-Caenorhabditis elegans@C33H5_12c_RRM_1[SRlCeSRp20/rsp-6][KOG0107]
scl-Arabidopsis thaliana@NP_564685_RRM_1[SRlSR33/SCL33][KOG0110]
scl-Arabidopsis thaliana@NP_001031195_RRM_1[SRlSR33/SCL33][KOG0110]
scl-Arabidopsis thaliana@NP_187966_RRM_1[SRlSCL30a]
scl-Populus trichocarpa@jgi_Poptr1_1_547618_eugene3_00010059_RRM_1
scl-Sorghum bicolor_BTx623@jgi_Sorbi1_5049360_Sb01g034200_RRM_1
scl-Oryza sativa_japonica_cultivar_group@NP_001050213_RRM_1[KOG4207]
scl-Oryza sativa_japonica_cultivar_group@NP_001060372_RRM_1[SRlSCL25]
scl-Sorghum bicolor_BTx623@jgi_Sorbi1_5052097_Sb02g040350_RRM_1[SRlSCL25]
scl-Selaginella moellendorfii@jgi_Selmo1_79810_e_gw1_3_146_1_RRM_1[KOG4207]
scl-Selaginella moellendorfii@jgi_Selmo1_10210_0_e_gw1_25_62_1_RRM_1[KOG0148]
scl-Arabidopsis thaliana@NP_197382_RRM_1[SRlSCL28][KOG0127]
scl-Populus trichocarpa@jgi_Poptr1_1_420699_gw1_VIII_2127_1_RRM_1[SRlSCL28]
scl-Populus trichocarpa@jgi_Poptr1_1_814124_estExt_fgenesh4_kg___RRM_1
scl-Oryza sativa_japonica_cultivar_group@NP_001050168_RRM_1
scl-Sorghum bicolor_BTx623@jgi_Sorbi1_5049384_Sb01g034590_RRM_1[KOG0127]
scl-Selaginella moellendorfii@jgi_Selmo1_8339_gw1_49_32_1_RRM_1[KOG4207]
scl-Oryza sativa_japonica_cultivar_group@NP_001068546_RRM_1[KOG0127]
scl-Sorghum bicolor_BTx623@jgi_Sorbi1_5056881_Sb05g027700_RRM_1
scl-Arabidopsis thaliana@NP_567021_RRM_1[SRlSCL30][KOG0122]
scl-Oryza sativa_japonica_cultivar_group@NP_001067091_RRM_1[SRlSCL30b]
scl-Oryza sativa_japonica_cultivar_group@NP_001046448_RRM_1[SRlSCL30a]
scl-Sorghum bicolor_BTx623@jgi_Sorbi1_5047478_Sb05145002010_RRM_1[KOG0110]
scl-Populus trichocarpa@jgi_Poptr1_1_636109_grail3_0049021701_RRM_1[KOG0122]
scl-Populus trichocarpa@jgi_Poptr1_1_227475_gw1_X_2172_1_RRM_1[KOG0127]
scl-Physcomitrella patens@jgi_Phypa1_1_108464_RRM_1
scl-Physcomitrella patens@jgi_Phypa1_1_141868_RRM_1
scl-Selaginella moellendorfii@jgi_Selmo1_8401_gw1_62_42_1_RRM_1[KOG4207]
SCL
srrp-Mus musculus@ENSMUSP00000095455_RRM_1[SRrplSRp38]
srrp-Homo sapiens@ENSP00000345293_RRM_1[SRrplSRp38]
srrp-Homo sapiens@ENSP00000358479_RRM_1[SRrplSRp35]
SRrp
N2
scl-Chlorella sp._NC64A@jgi_ChlNC64A_1_17896_gw1_13_295_1_RRM_1
scl-Chlamydomonas reinhardtii@jgi_Chlre3_116256_RRM_1
scl-Volvox carteri@jgi_Volca1_121229_RRM_1[KOG0415]
scl-Chlorella sp._NC64A@jgi_ChlNC64A_1_135593_IGS_gm_14_00126_RRM_1[KOG4207]
scl-Chlamydomonas reinhardtii@jgi_Chlre3_113900_RRM_1[KOG4207]
scl-Volvox carteri@jgi_Volca1_73973_RRM_1[KOG4207]
scl-Volvox carteri@jgi_Volca1_120666_RRM_1
scl-Chlamydomonas reinhardtii@jgi_Chlre3_188086_RRM_1
sc35-Mus musculus@ENSMUSP00000072182_RRM_1[SRlSC35][KOG4207]
sc35-Homo sapiens@ENSP00000293253_RRM_1[SRlSC35][KOG4368]
sc35-Homo sapiens@ENSP00000353089_RRM_1[SRlSC35][KOG4368]
sc35-Mus musculus@ENSMUSP00000091059_RRM_1[SRlSC35][KOG4368]
sc35-Homo sapiens@ENSP00000381839_RRM_1[KOG4207]
sc35-Mus musculus@ENSMUSP00000078984_RRM_1[KOG4207]
sc35-Caenorhabditis elegans@EEED8_7a_RRM_1[SRlCeSC35/rsp-4]
sc35-Caenorhabditis elegans@EEED8_7b_RRM_1[SRlCeSC35/rsp-4][KOG4207]
sc35-Drosophila melanogaster@CG5442_PB_RRM_1[SRlSC35][KOG4207]
scl-Ostreococcus lucimarinus@jgi_Ost9901_3_9858_RRM_1[KOG0108]
scl-Ostreococcus sp._RCC809@jgi_OstRCC809_1_16959_gw1_4_1177_1_RRM_1[KOG0108]
scl-Ostreococcus tauri@jgi_Ostta4_7795_RRM_1[KOG0108]
sc35-Ostreococcus lucimarinus@jgi_Ost9901_3_8174_RRM_1[KOG4207]
sc35-Ostreococcus sp._RCC809@jgi_OstRCC809_1_16908_gw1_5_1114_1_RRM_1[KOG4207]
sc35-Micromonas pusilla_RCC299@jgi_MicpuN3_91839_estExt_Genewise2____RRM_1
sc35-Micromonas pusilla_CCMP1545@jgi_MicpuC2_9129_gw1_3_931_1_RRM_1[KOG4207]
sc35-Physcomitrella patens@jgi_Phypa1_1_172957_RRM_1[KOG0415]
sc35-Physcomitrella patens@jgi_Phypa1_1_122960_RRM_1[KOG4207]
sc35-Selaginella moellendorfii@jgi_Selmo1_7952_gw1_7_122_1_RRM_1[SRlSC35a][KOG4207]
sc35-Selaginella moellendorfii@jgi_Selmo1_134332_e_gw1_124_43_1_RRM_1[SRlSC35a][KOG4207]
sc35-Sorghum bicolor_BTx623@jgi_Sorbi1_5059167_Sb07g029000_RRM_1[KOG0147]
sc35-Oryza sativa_japonica_cultivar_group@NP_001050263_RRM_1[SRlSC35c][KOG4207]
sc35-Sorghum bicolor_BTx623@jgi_Sorbi1_5029845_Sb01g033710_RRM_1[KOG4207]
sc35-Oryza sativa_japonica_cultivar_group@NP_001062994_RRM_1[SRlSC35a][KOG0415]
sc35-Populus trichocarpa@jgi_Poptr1_1_672394_grail3_1445000102_RRM_1[SRlSC35a][KOG4207]
sc35-Populus trichocarpa@jgi_Poptr1_1_645719_grail3_0067008905_RRM_1[SRlSC35b][KOG4207]
sc35-Arabidopsis thaliana@NP_851261_RRM_1[SRlSC35][KOG4207]
sc35-Populus trichocarpa@jgi_Poptr1_1_709848_estExt_Genewise1___RRM_1[KOG4207]
sc35-Populus trichocarpa@jgi_Poptr1_1_559112_eugene3_00050727_RRM_1[SRlSC35a][KOG4207]
sc35-Oryza sativa_japonica_cultivar_group@NP_001060322_RRM_1[SRlSC35b][KOG4207]
sc35-Cyanidioschyzon merolae@gnl_CMER_CML202C_RRM_1
SC35
N1
scl-Micromonas pusilla_CCMP1545@jgi_MicpuC2_11901_gw1_8_619_1_RRM_1[KOG4207]
scl-Micromonas pusilla_RCC299@jgi_MicpuN3_74643_gw2_11_587_1_RRM_1[KOG0107]
9g8-Mus musculus@ENSMUSP00000066393_RRM_1[SRl9G8][KOG0107]
9g8-Homo sapiens@ENSP00000368146_RRM_1[SRl9G8][KOG0107]
9g8-Homo sapiens@ENSP00000368134_RRM_1[SRl9G8][KOG0107]
9g8-Homo sapiens@ENSP00000325905_RRM_1[SRl9G8]
9g8-Mus musculus@ENSMUSP00000108040_RRM_1[SRl9G8][KOG0107]
9g8-Mus musculus@ENSMUSP00000070983_RRM_1[SRl9G8][KOG0107]
srp20-Mus musculus@ENSMUSP00000110365_RRM_1[SRlSRp20][KOG0107]
srp20-Homo sapiens@ENSP00000362820_RRM_1[SRlSRp20][KOG0107]
srp20-Homo sapiens@ENSP00000344762_RRM_1[SRlSRp20][KOG0107]
srp20-Mus musculus@ENSMUSP00000049025_RRM_1[SRlSRp20][KOG0107]
srp20-Mus musculus@ENSMUSP00000100538_RRM_1[SRlSRp20][KOG0107]
srp20-Drosophila melanogaster@CG10203_PA_RRM_1[SRl9G8][KOG0107]
SRp20
9G8

Fig. S28

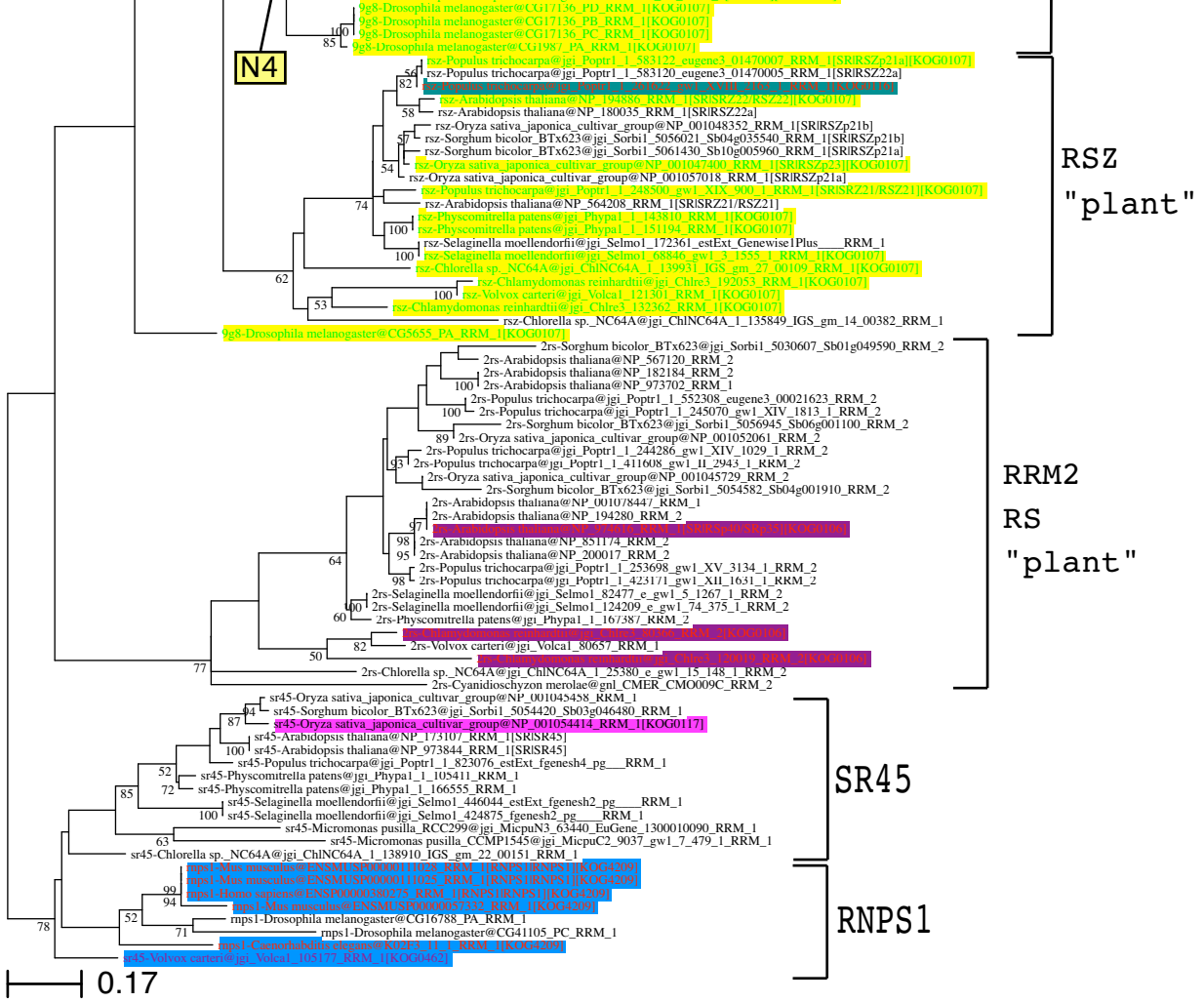


Fig. S29

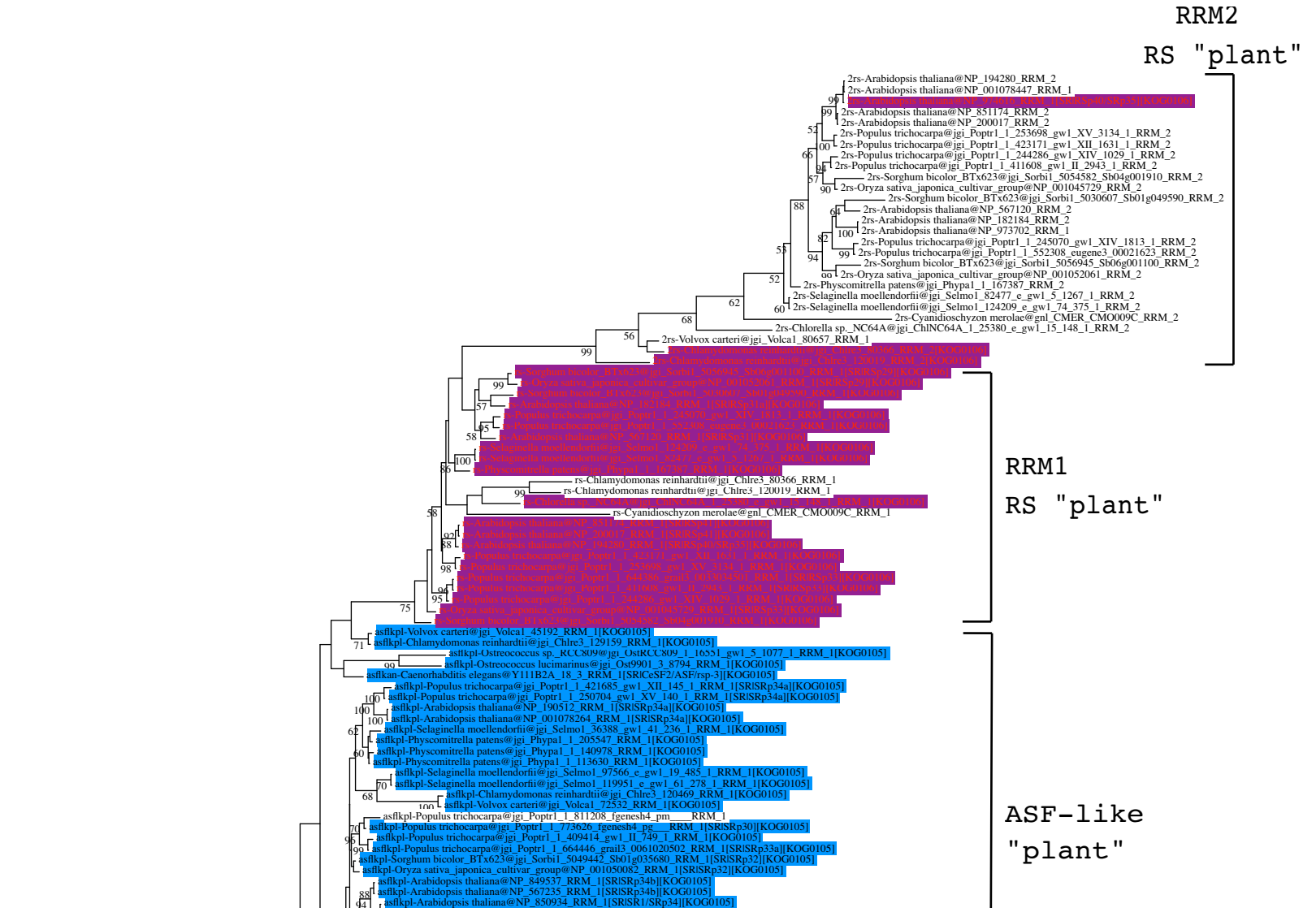


Fig. S29

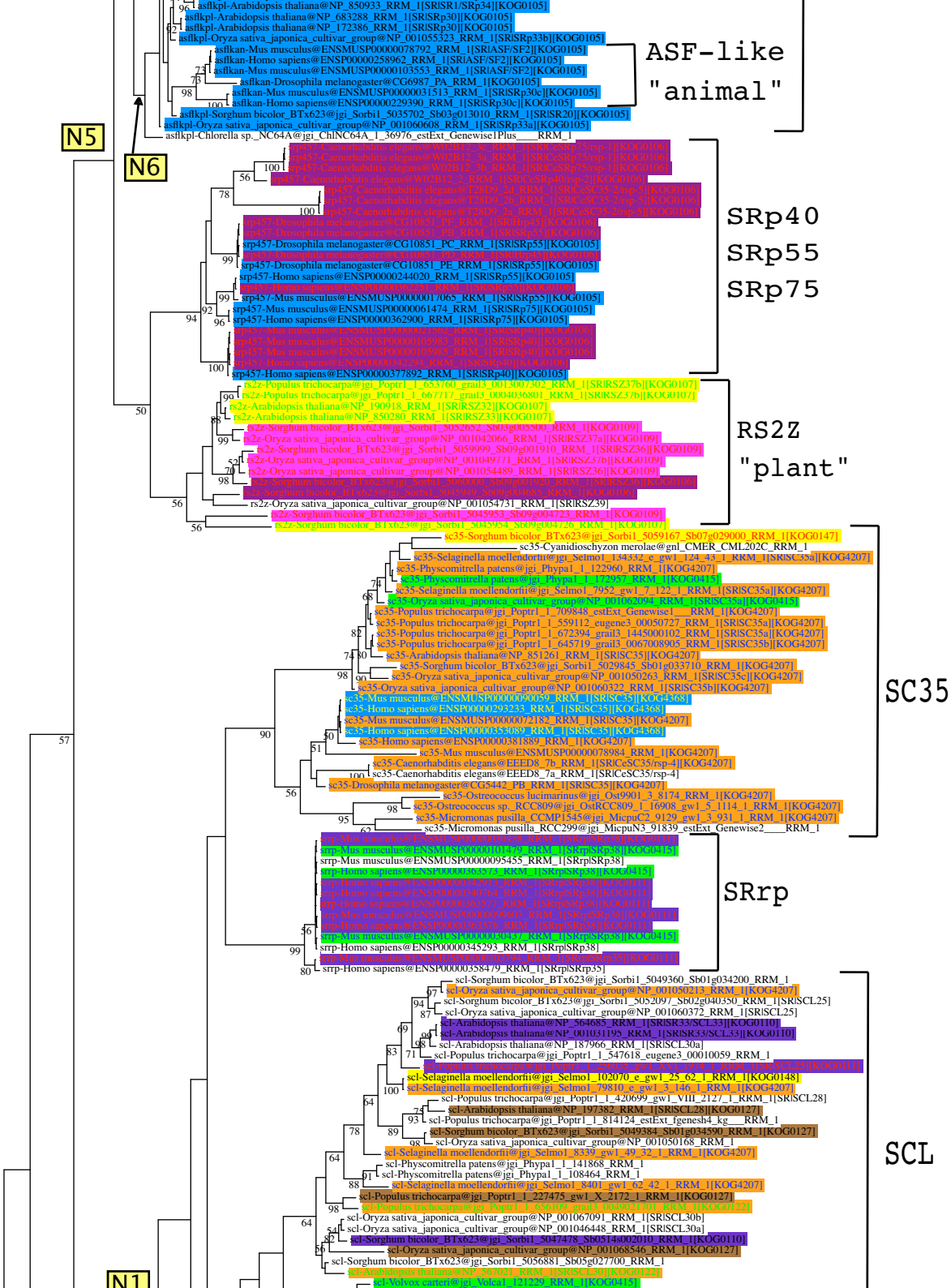
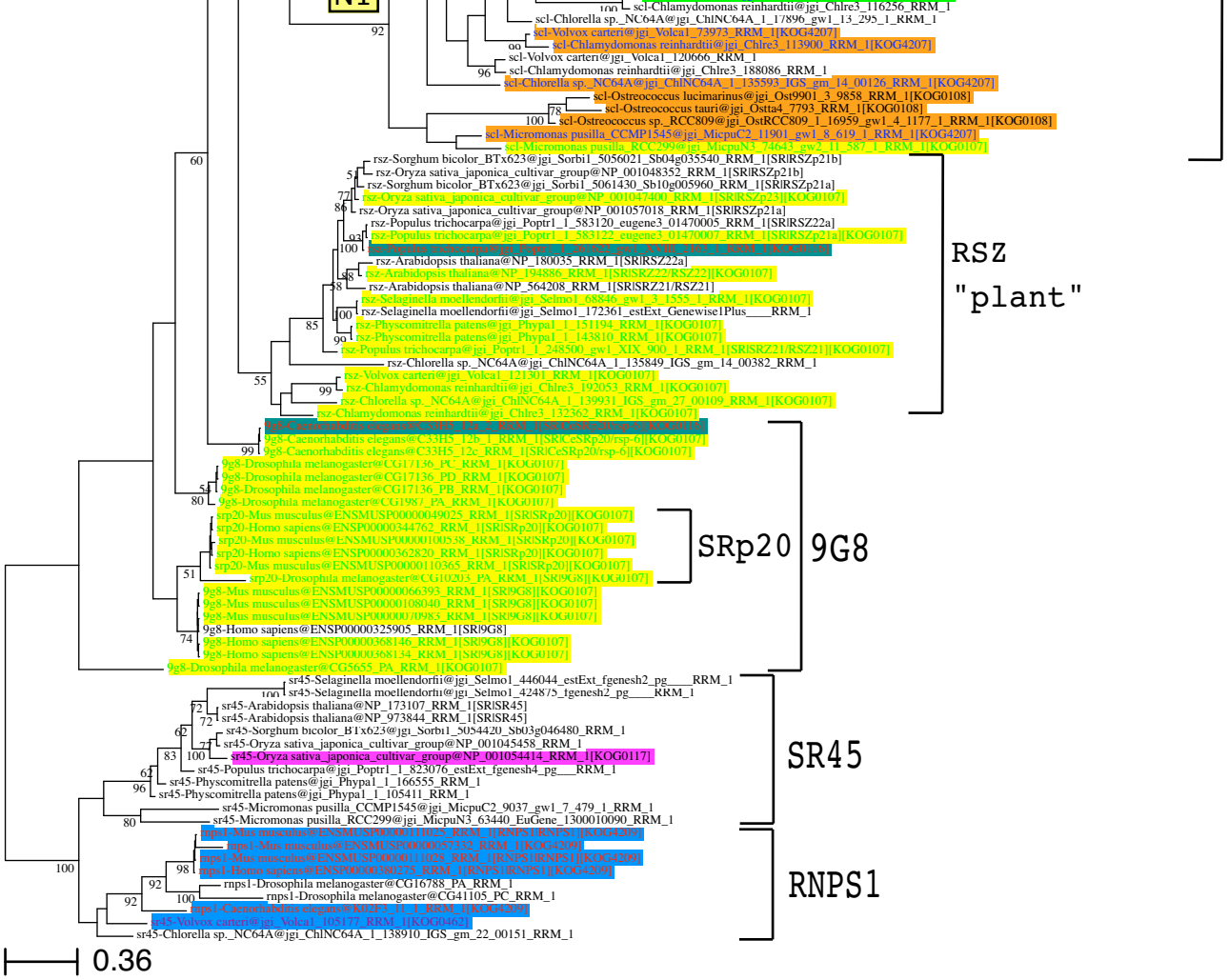

N5
N6
N1
ASF-like
"animal"
SRp40
SRp55
SRp75
RS2Z
"plant"
SC35
SRrp
SCL

Fig. S29


scl-Chlamydomonas reinhardtii@jgi_Chlre3_116256_RRM_1
scl-Chlorella sp._NC64A@jgi_ChlNC64A_1_17896_gw1_13_295_1_RRM_1
scl-Volvox carteri@jgi_Volca1_73973_RRM_1[KOG4207]
scl-Chlamydomonas reinhardtii@jgi_Chlre3_113900_RRM_1[KOG4207]
scl-Volvox carteri@jgi_Volca1_120666_RRM_1
scl-Chlamydomonas reinhardtii@jgi_Chlre3_188086_RRM_1
scl-Chlorella sp._NC64A@jgi_ChlNC64A_1_135593_IGS_gm_14_00126_RRM_1[KOG4207]
scl-Ostreococcus lucimarinus@jgi_Ost9901_3_9858_RRM_1[KOG0108]
scl-Ostreococcus tauri@jgi_Ostta4_7793_RRM_1[KOG0108]
scl-Ostreococcus sp._RCC809@jgi_OstRCC809_1_16959_gw1_4_1177_1_RRM_1[KOG0108]
scl-Micromonas pusilla_CCMP1545@jgi_MicpuC2_11901_gw1_8_619_1_RRM_1[KOG4207]
scl-Micromonas pusilla_RCC299@jgi_MicpuN3_74643_gw2_11_587_1_RRM_1[KOG0107]
rsz-Sorghum bicolor_BTx623@jgi_Sorbi1_5056021_Sb04g035540_RRM_1[SRlRSZp21b]
rsz-Oryza sativa_japonica_cultivar_group@NP_001048352_RRM_1[SRlRSZp21b]
rsz-Sorghum bicolor_BTx623@jgi_Sorbi1_5061430_Sb10g005960_RRM_1[SRlRSZp21a]
rsz-Oryza sativa_japonica_cultivar_group@NP_001047400_RRM_1[SRlRSZp23][KOG0107]
rsz-Oryza sativa_japonica_cultivar_group@NP_001057018_RRM_1[SRlRSZp21a]
rsz-Populus trichocarpa@jgi_Poptr1_1_583120_eugene3_01470005_RRM_1[SRlRSZ22a]
rsz-Populus trichocarpa@jgi_Poptr1_1_583122_eugene3_01470007_RRM_1[SRlRSZp21a][KOG0107]
rsz-Arabidopsis thaliana@NP_180035_RRM_1[SRlRSZ22a]
rsz-Arabidopsis thaliana@NP_194886_RRM_1[SRlSRZ22/RSZ22][KOG0107]
rsz-Arabidopsis thaliana@NP_564208_RRM_1[SRlSRZ21/RSZ21]
rsz-Selaginella moellendorfii@jgi_Selmo1_68846_gw1_3_1555_1_RRM_1[KOG0107]
rsz-Selaginella moellendorfii@jgi_Selmo1_172361_estExt_Genewise1Plus____RRM_1
rsz-Physcomitrella patens@jgi_Phypa1_1_151194_RRM_1[KOG0107]
rsz-Physcomitrella patens@jgi_Phypa1_1_143810_RRM_1[KOG0107]
rsz-Populus trichocarpa@jgi_Poptr1_1_248500_gw1_XIX_900_1_RRM_1[SRlSRZ21/RSZ21][KOG0107]
rsz-Chlorella sp._NC64A@jgi_ChlNC64A_1_135849_IGS_gm_14_00382_RRM_1
rsz-Volvox carteri@jgi_Volca1_121301_RRM_1[KOG0107]
rsz-Chlamydomonas reinhardtii@jgi_Chlre3_192053_RRM_1[KOG0107]
rsz-Chlorella sp._NC64A@jgi_ChlNC64A_1_139931_IGS_gm_27_00109_RRM_1[KOG0107]
rsz-Chlamydomonas reinhardtii@jgi_Chlre3_132362_RRM_1[KOG0107]
RSZ
"plant"
9g8-Caenorhabditis elegans@C33H5_12b_1_RRM_1[SRlCeSRp20/rsp-6][KOG0107]
9g8-Caenorhabditis elegans@C33H5_12c_RRM_1[SRlCeSRp20/rsp-6][KOG0107]
9g8-Drosophila melanogaster@CG17136_PC_RRM_1[KOG0107]
9g8-Drosophila melanogaster@CG17136_PD_RRM_1[KOG0107]
9g8-Drosophila melanogaster@CG17136_PB_RRM_1[KOG0107]
9g8-Drosophila melanogaster@CG1987_PA_RRM_1[KOG0107]
srp20-Mus musculus@ENSMUSP00000049025_RRM_1[SRlSRp20][KOG0107]
srp20-Homo sapiens@ENSP00000344762_RRM_1[SRlSRp20][KOG0107]
srp20-Mus musculus@ENSMUSP00000100538_RRM_1[SRlSRp20][KOG0107]
srp20-Homo sapiens@ENSP00000362820_RRM_1[SRlSRp20][KOG0107]
srp20-Mus musculus@ENSMUSP00000110365_RRM_1[SRlSRp20][KOG0107]
srp20-Drosophila melanogaster@CG10203_PA_RRM_1[SRl9G8][KOG0107]
SRp20
9G8
9g8-Mus musculus@ENSMUSP00000066393_RRM_1[SRl9G8][KOG0107]
9g8-Mus musculus@ENSMUSP00000108040_RRM_1[SRl9G8][KOG0107]
9g8-Mus musculus@ENSMUSP00000070983_RRM_1[SRl9G8][KOG0107]
9g8-Homo sapiens@ENSP00000325905_RRM_1[SRl9G8]
9g8-Homo sapiens@ENSP00000368146_RRM_1[SRl9G8][KOG0107]
9g8-Homo sapiens@ENSP00000368134_RRM_1[SRl9G8][KOG0107]
9g8-Drosophila melanogaster@CG5655_PA_RRM_1[KOG0107]
sr45-Selaginella moellendorfii@jgi_Selmo1_446044_estExt_fgenesh2_pg____RRM_1
sr45-Selaginella moellendorfii@jgi_Selmo1_424875_fgenesh2_pg____RRM_1
sr45-Arabidopsis thaliana@NP_173107_RRM_1[SRlSR45]
sr45-Arabidopsis thaliana@NP_973844_RRM_1[SRlSR45]
sr45-Sorghum bicolor_BTx623@jgi_Sorbi1_5054420_Sb03g046480_RRM_1
sr45-Oryza sativa_japonica_cultivar_group@NP_001045458_RRM_1
sr45-Oryza sativa_japonica_cultivar_group@NP_001054414_RRM_1[KOG0117]
sr45-Populus trichocarpa@jgi_Poptr1_1_823076_estExt_fgenesh4_pg___RRM_1
sr45-Physcomitrella patens@jgi_Phypa1_1_166555_RRM_1
sr45-Physcomitrella patens@jgi_Phypa1_1_105411_RRM_1
sr45-Micromonas pusilla_CCMP1545@jgi_MicpuC2_9037_gw1_7_479_1_RRM_1
sr45-Micromonas pusilla_RCC299@jgi_MicpuN3_63440_EuGene_1300010090_RRM_1
SR45
rnps1-Mus musculus@ENSMUSP00000110026_RRM_1[RNPS1/RNPS1][KOG4209]
rnps1-Mus musculus@ENSMUSP00000057555_RRM_1[KOG4209]
rnps1-Mus musculus@ENSMUSP00000110026_RRM_1[RNPS1/RNPS1][KOG4209]
rnps1-Homo sapiens@ENSP00000380539_RRM_1[RNPS1/RNPS1][KOG4209]
rnps1-Drosophila melanogaster@CG16788_PA_RRM_1
rnps1-Drosophila melanogaster@CG41105_PC_RRM_1
rnps1-Caenorhabditis elegans@K02F3_11_1_RRM_1[KOG4209]
sr45-Volvox carteri@jgi_Volca1_105177_RRM_1[KOG4021]
sr45-Chlorella sp._NC64A@jgi_ChlNC64A_1_138910_IGS_gm_22_00151_RRM_1
RNPS1
0.36