

Supporting Information

Dreborg *et al.* 10.1073/pnas.0805590105

SI Text

Short descriptions of reasons for exclusions of families BAND 4.1 (EPB41). This is a large family with 13 members in humans. Only two of the genes in the family are located inside the investigated chromosomal regions; EPB41 on chromosome 1, 29.09 Mb and XP_932590.1 on chromosome 8, 49.99 Mb. However, these genes are very different and belong to different subfamilies.

SYNAPTOTAGAMIN-LIKE (SYTL). The two genes located within the selected chromosomal regions, SYTL1 on human chromosome 1 and SYTL3 on human chromosome 6, belong to different subfamilies according to the phylogenetic tree.

XP_945180.1/NP_067050.1/NP_660346.1. All proteins in this family are extremely conserved. Spotted green pufferfish, western clawed frog and mouse only have one gene in this family. Chicken has two genes, but they are very similar and cluster together in the tree. In humans and dogs there are three genes in the family but according to the topology in the tree they must have arisen by recent gene duplication events that occurred independently in the human and dog lineages. Therefore, there are no paralogous genes in this family that are the result of R1 or R2.

CONNECTOR ENHANCER OF KINASE SUPPRESSOR OF RAS (CNKSR). The proteins in this family have one or several of the following Pfam

domains: PH, SAM_2, DUF1170, and PDZ. Unfortunately, none of the domains can be found in all (or most) of the proteins and they are therefore difficult to align.

HIGH-MOBILITY GROUP BOX (HMGB). There has been a large expansion of the HMGB family in human and mouse and many of the genes are retrotransposed copies. The numbers of genes are as follows: human, 19; mouse, 39; dog, 4; chicken, 3; frog, 3. It is difficult to determine which genes are orthologues and which are paralogues by only studying the tree topology. Three human genes in this family have four introns, HMGB3 on chromosome X, HMGB2 on chromosome 4 and HMGB1 on chromosome 13. The remaining human genes (including the genes on chromosomes 1 and 20) either lack introns completely or have only one intron, and are most likely retrotransposed copies.

PHOSPHATASE AND ACTIN REGULATOR (PHACTR). The sequences in this family align very poorly, making phylogenetic analysis unreliable. The genes contain a large number of introns and seem to be alternatively spliced. However, the transcripts of orthologous genes in different species are very different and we can not exclude the possibility that we are aligning different transcripts in all species.

Opioid receptors

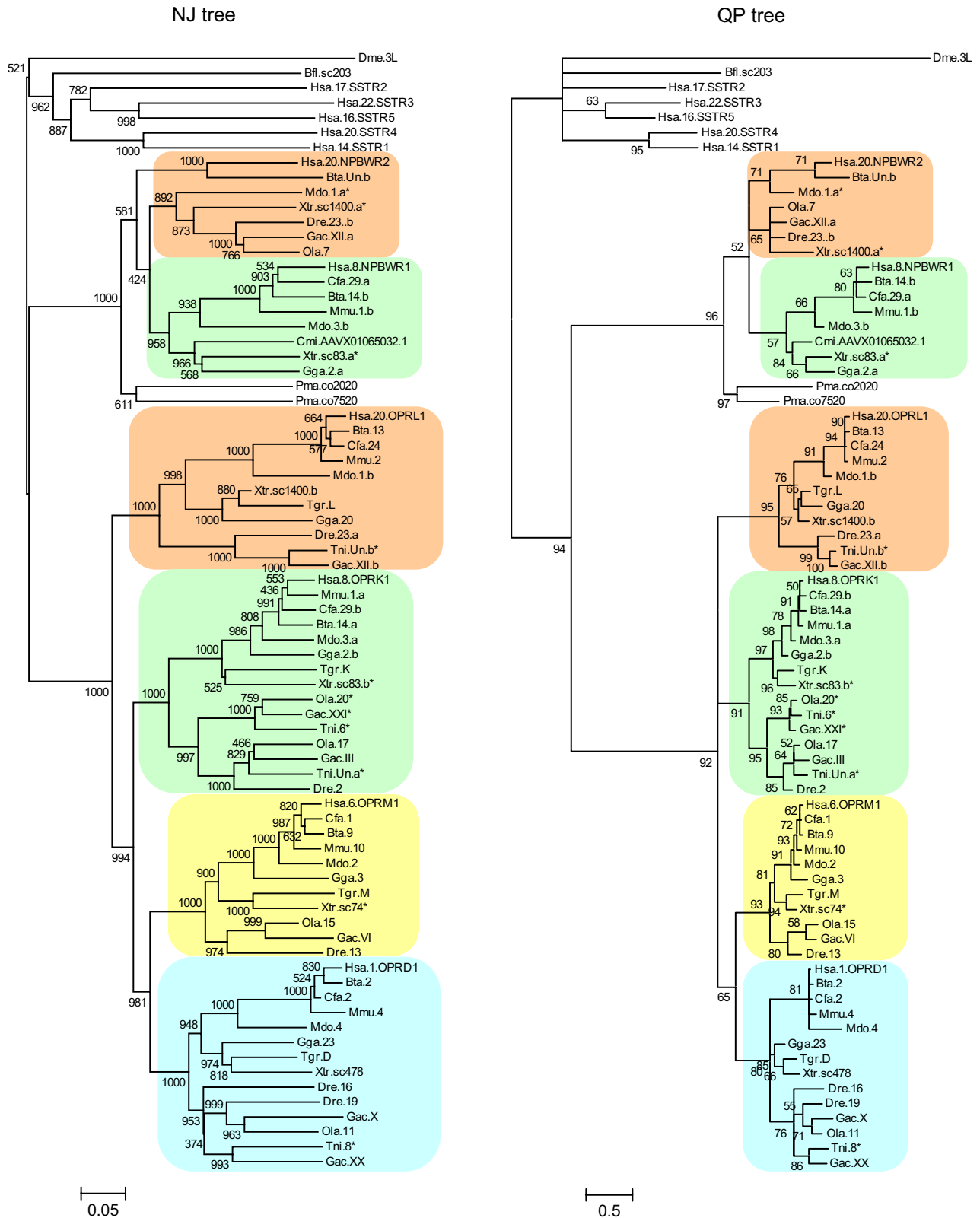


Fig. S1. Neighbor-joining and quartet-puzzling maximum likelihood trees for opioid receptors.

ARID1

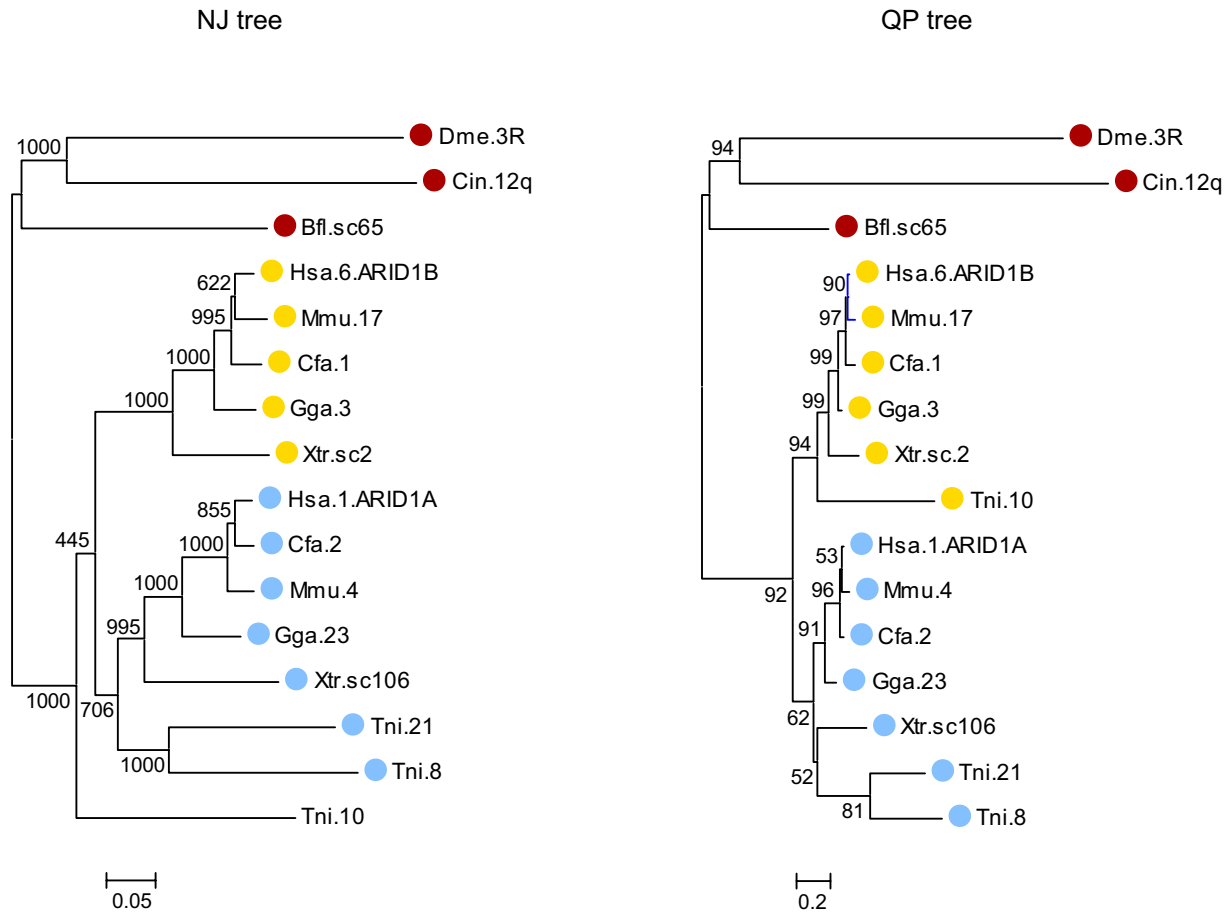


Fig. S2. Neighbor-joining and quartet-puzzling maximum likelihood trees for AT RICH INTERACTIVE DOMAIN 1 (ARID1). The ARID family consist of 15 genes in the human, mouse and dog genome. The proteins all contain the ARID domain, which has DNA-binding properties. The ARID family can be divided into 7 subfamilies., ARID1A and ARID1B belong to the ARID1 subfamily and show 80% sequence identity within the ARID domain and \approx 50% identity across the full length amino acid sequence. ARID1A and ARID1B bind DNA with high affinity but without sequence specificity and they are alternative components of the SWI/SNF-related chromatin remodelling complexes (1, 2). Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Cfa, *Canis familiaris*; Gga, *Gallus gallus*; Xtr, *Xenopus tropicalis*; Tni, *Tetraodon nigroviridis*; Bfl, *Branchiostoma floridae*; Dme, *Drosophila melanogaster*; Cin, *Ciona intestinalis*.

1. Patsialou A, Wilskerand D, Moran E (2005) DNA-binding properties of ARID family proteins. *Nucleic Acids Res* 33:66–80.
2. Wilsker D, Probst L, Wain HM, Maltais L, Tucker PW, Moran E (2005) Nomenclature of the ARID family of DNA-binding proteins. *Genomics* 86:242–251.

GMEB

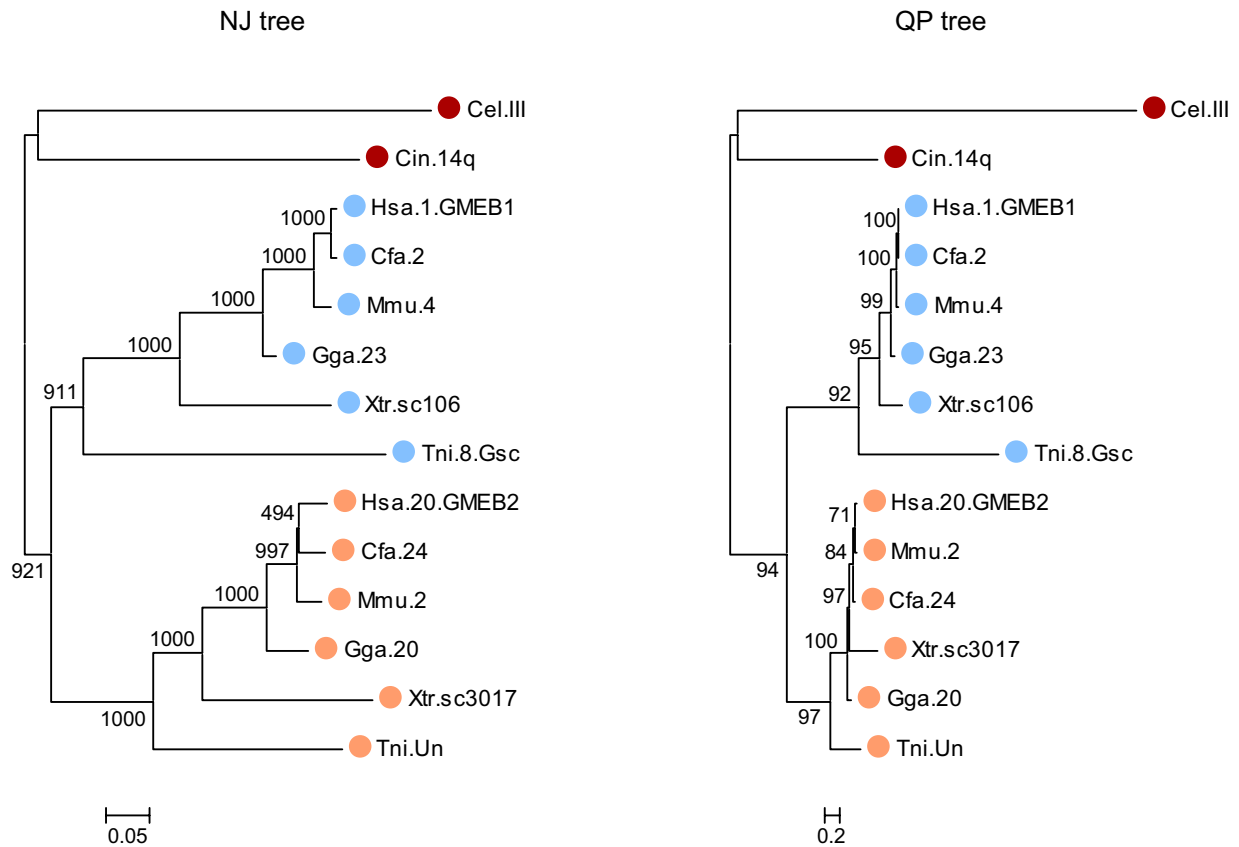


Fig. S3. Neighbor-joining and quartet-puzzling maximum likelihood trees for GLUCOCORTICOID MODULATORY ELEMENT BINDING PROTEIN (GMEB). The two proteins in this family, GMEB1 and GMEB2, contain a Pfam SAND domain, which has DNA binding activity. They are involved in regulation of the glucocorticoid receptors (GRs) sensitivity to steroid hormones. GMEB1 and GMEB2 form a heteromeric complex that binds to a cis-acting element called the glucocorticoid modulatory element (GME) (1, 2). The GMEBs can also bind to the GR-ligand complex and thereby modify the GRs' induction properties (2). Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Cfa, *Canis familiaris*; Gga, *Gallus gallus*; Xtr, *Xenopus tropicalis*; Tni, *Tetraodon nigroviridis*; Cin, *Ciona intestinalis*; Cel, *Caenorhabditis elegans*.

1. Kaul S, Blackford JA, Jr, Chen J, Ogryzko VV, S. Simons SS, Jr (2005) Properties of the glucocorticoid, modulatory element binding proteins GMEB-1 and -2: Potential receptor transactivation and members of the family of KDVK proteins. *Mol Endocrinol* 14:1010–1027.
2. Chen J, He Y, Simons SS, Jr (2004). Structure/activity relationships for GMEB-2: The second member of the glucocorticoid modulatory element-binding complex. *Biochemistry* 43:245–255.

LYPLA

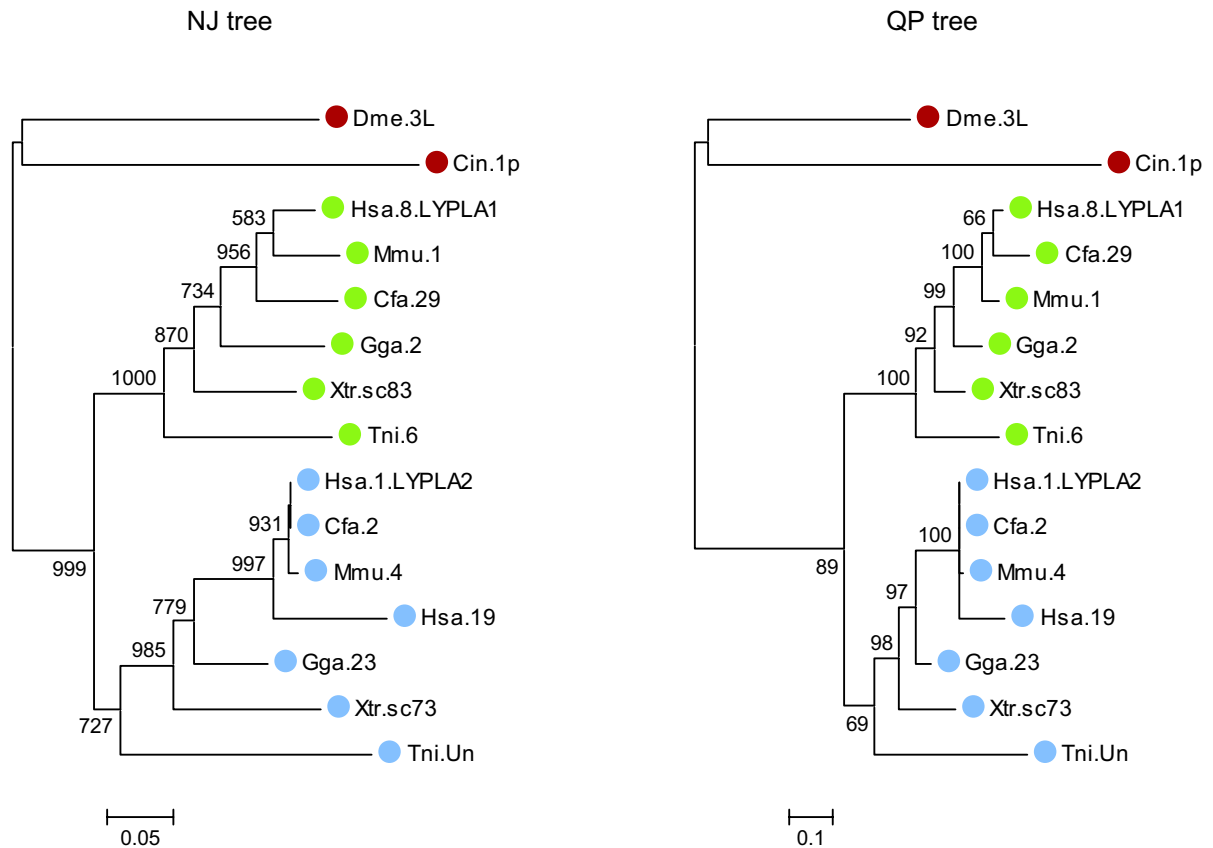


Fig. S4. Neighbor-joining and quartet-puzzling maximum likelihood trees for ACYL THIOESTERASE/LYSOPHOSPHOLIPASE (LYPLA). All proteins in this family contain the Pfam Abhydrolase_2 domain. The LYPLA family proteins are enzymes that hydrolyse the ester bond of lysophospholipids to produce a free fatty acid and a glycerolphosphate derivative. LYPLA1 and LYPLA2 have a catalytic triad composed of Ser-Asp-His and their catalytic mechanism resembles that of serine hydrolases. Lysophospholipids have many biological functions and can act as second messengers by transducing signals from membrane receptors. LYPLA1 and LYPLA2 are two of the many enzymes that control the lysophospholipids by regulating their levels [Wang A, Dennis EA (1999) Mamalian lysophospholipases. *Biochimica et Biophysica Acta* 1439:1–16]. There are four human genes that belong to this Ensembl family and they are located on chromosome 1 (LYPLA2), 6 (LYPLA2P1), 8 (LYPLA1) and 19 (no HGNC name available). LYPLA2P1 and the gene on human chromosome 19 are also found in Chimpanzee but not in any of the other species. Because these two genes are very similar to LYPLA2 but lack introns they are probably retrotransposed copies specific to the human/chimpanzee line. Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Cfa, *Canis familiaris*; Gga, *Gallus gallus*; Xtr, *Xenopus tropicalis*; Tni, *Tetraodon nigroviridis*; Dme, *Drosophila melanogaster*; Cin, *Ciona intestinalis*.

MYT1

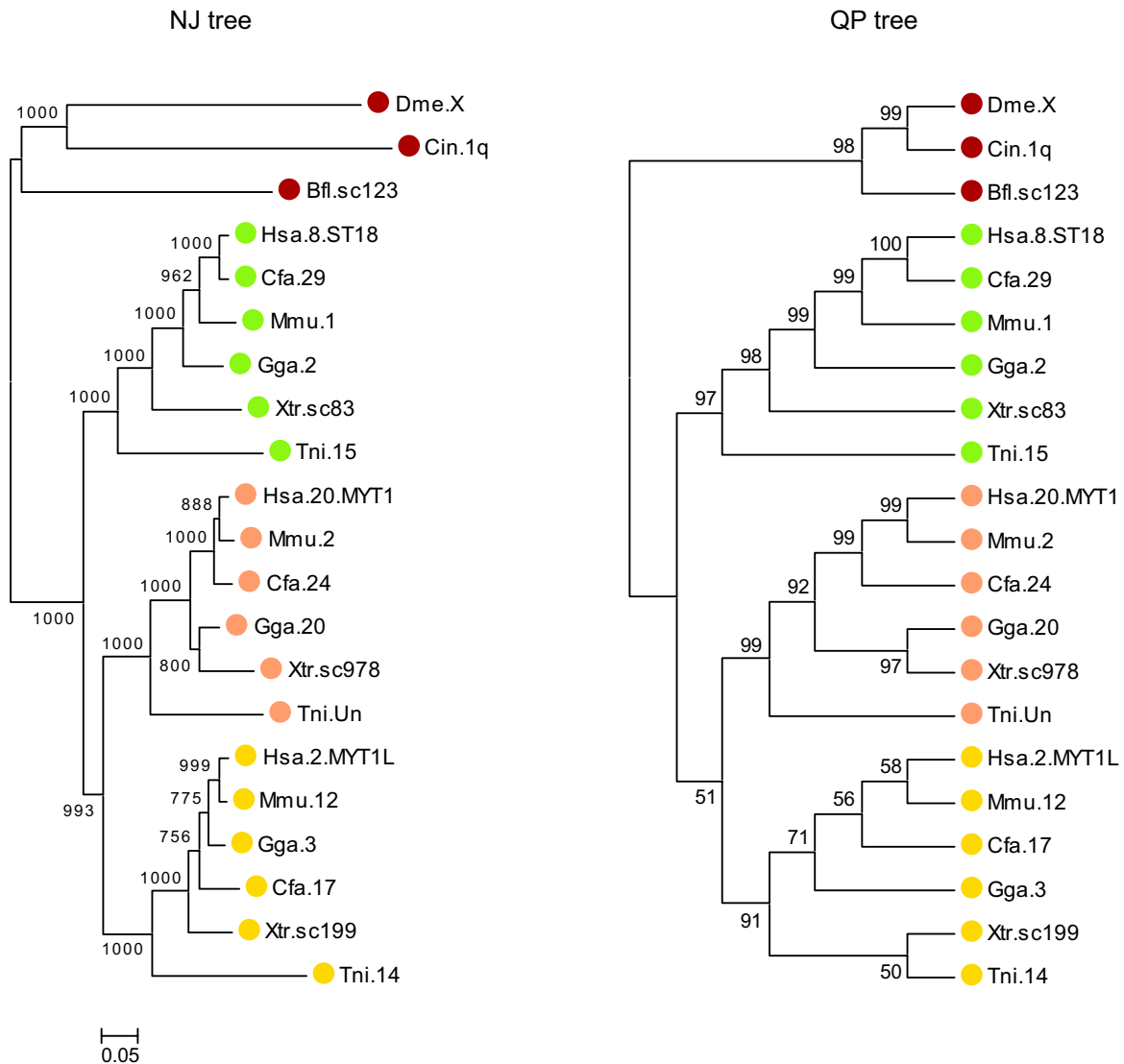


Fig. S5. Neighbor-joining and quartet-puzzling maximum likelihood trees for MYELIN TRANSCRIPTION FACTOR 1 (MYT1). The proteins that belong to this family are neural zinc fingers that act as transcription factors. The MYT1 family proteins bind to the transcriptional corepressor Sin3B while they also bind to promoters. Sin3B in turn binds histone deacetylase (HDAC), which modifies chromatin structure and thereby repress promoter activity [Rom E, Nielsen JA, Kim JG, Hudson LD (2005) Myt1 family recruits histone deacetylase to regulate neural transcription. *J Neurochem* 93:444–1453]. All members of this family contain several zf-C2HC domains and all except the *Ciona intestinalis* and *Drosophila* proteins have a PfamMYT1 domain. However, the MYT1 domain consist of zf-C2HC domains and the *Ciona* and *Drosophila* sequences align quite well with the other sequences. The human members are located on chromosome 2 (MYT1L), 8 (ST18), and 20 (MYT1). Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Cfa, *Canis familiaris*; Gga, *Gallus gallus*; Xtr, *Xenopus tropicalis*; Tni, *Tetraodon nigroviridis*; Bfl, *Branchiostoma floridae*; Dme, *Drosophila melanogaster*; Cin, *Ciona intestinalis*.

NKAIN

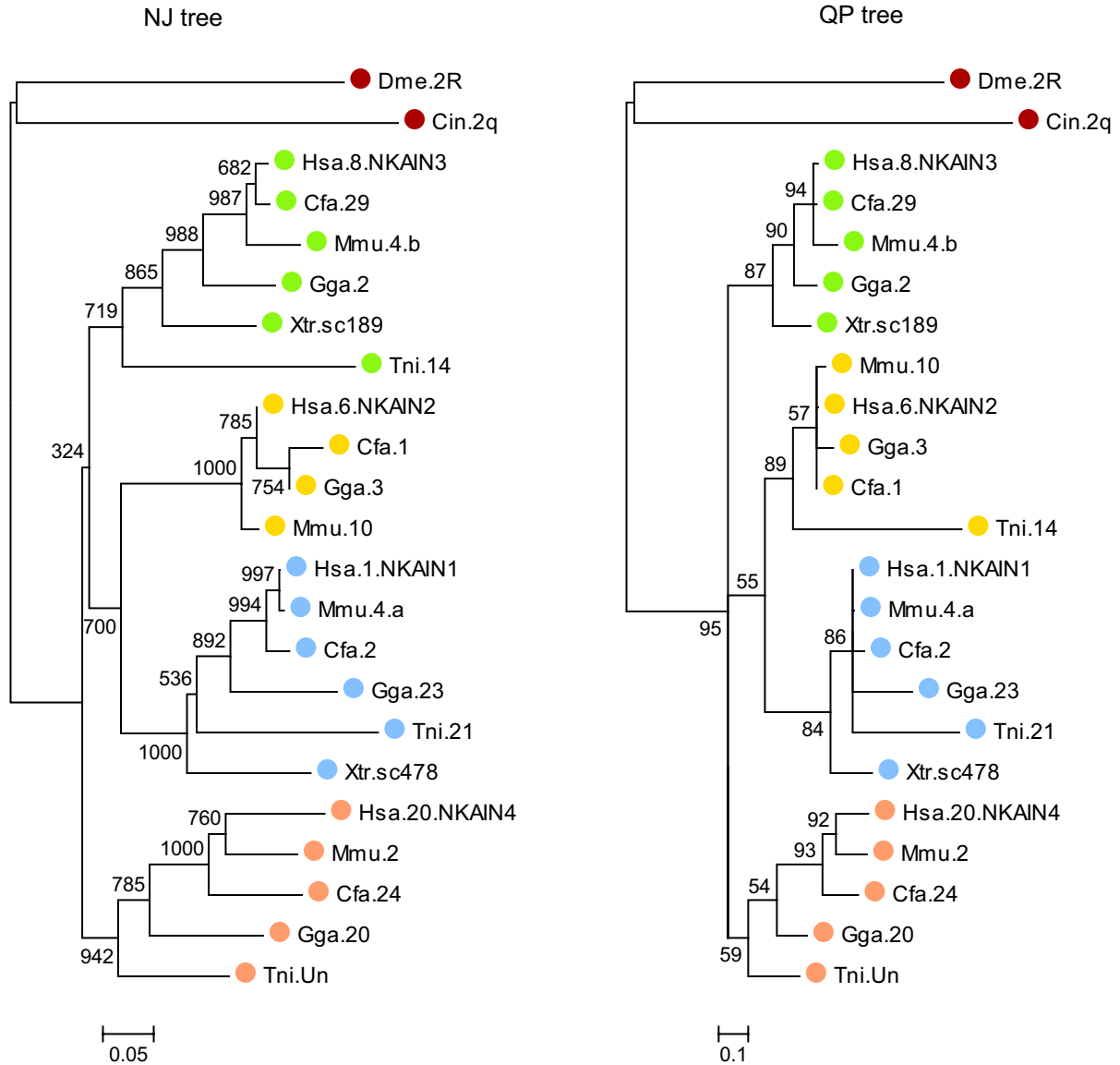


Fig. S6. Neighbor-joining and quartet-puzzling maximum likelihood trees for Na⁺/K⁺ transporting ATPase interacting (NKAIN). The proteins in this family are membrane proteins with three putative transmembrane domains. The NKAIN proteins are localized to neurons and interact with the β 1 subunit of the Na,K-ATPase [Gorokhova S, Bibert S, Geeringand K, Heintz N (2007) A novel family of transmembrane proteins interacting with β subunits of the Na,K-ATPase. *Human Mol Gen* 16:2394–2410]. Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Cfa, *Canis familiaris*; Gga, *Gallus gallus*; Xtr, *Xenopus tropicalis*; Tni, *Tetraodon nigroviridis*; Dme, *Drosophila melanogaster*; Cin, *Ciona intestinalis*.

PCMTD

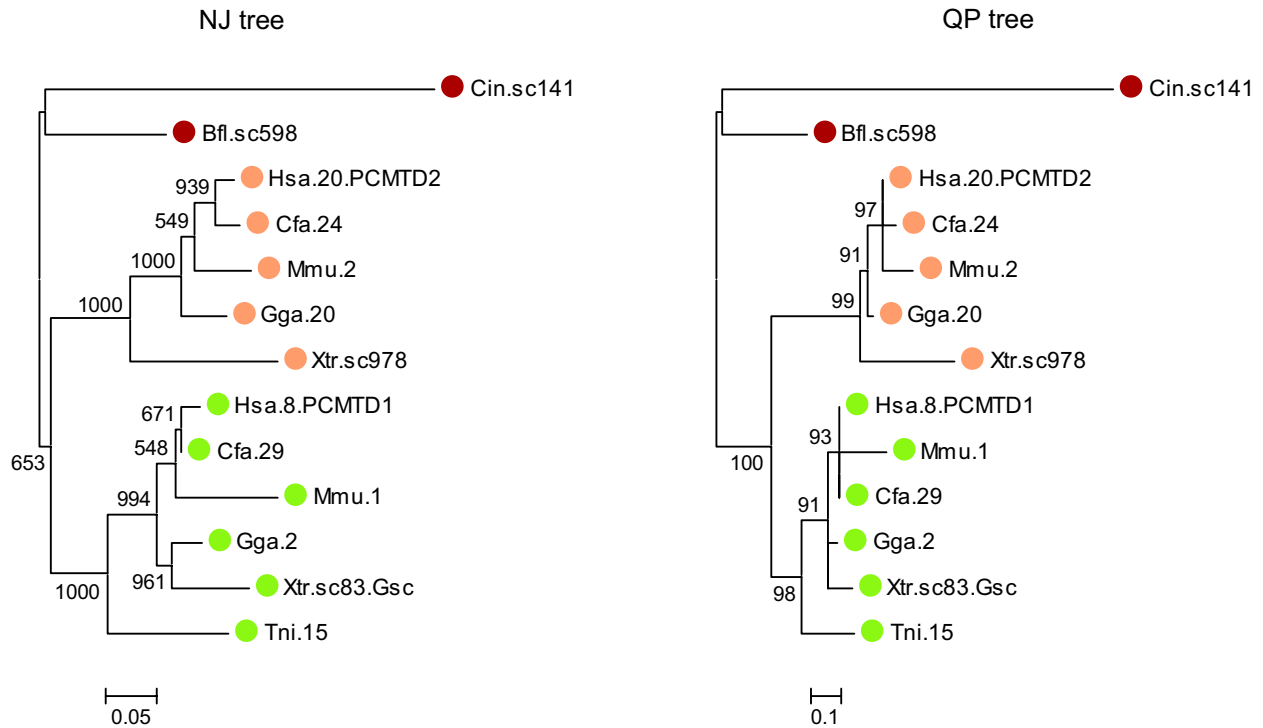


Fig. 57. Neighbor-joining and quartet-puzzling maximum likelihood trees for PROTEIN-L-ISOASPARTATE O-METHYLTRANSFERASE DOMAIN CONTAINING (PCMTD). The proteins in this family, PCMTD1 and PCMTD2, are enzymes and contain the catalytic PCMT domain. PCMTs methylate D-aspartyl and L-isoaspartyl residues in proteins and peptides. D-aspartyl and L-isoaspartyl are produced by spontaneous deamidation or racemization of normal asparagine residues and the PCMTs are thought to play a role in the repair and degradation of proteins with this type of damage (Boivin D, Bilodeau D, Béliveau R (1995) Immunochemical characterization of L-isoaspartyl-protein carboxyl methyltransferase from mammalian tissues. *Biochem J.* 309:993–998). Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Cfa, *Canis familiaris*; Gga, *Gallus gallus*; Xtr, *Xenopus tropicalis*; Tni, *Tetraodon nigroviridis*; Bfl, *Branchiostoma floridae*; Cin, *Ciona intestinalis*.

RGS

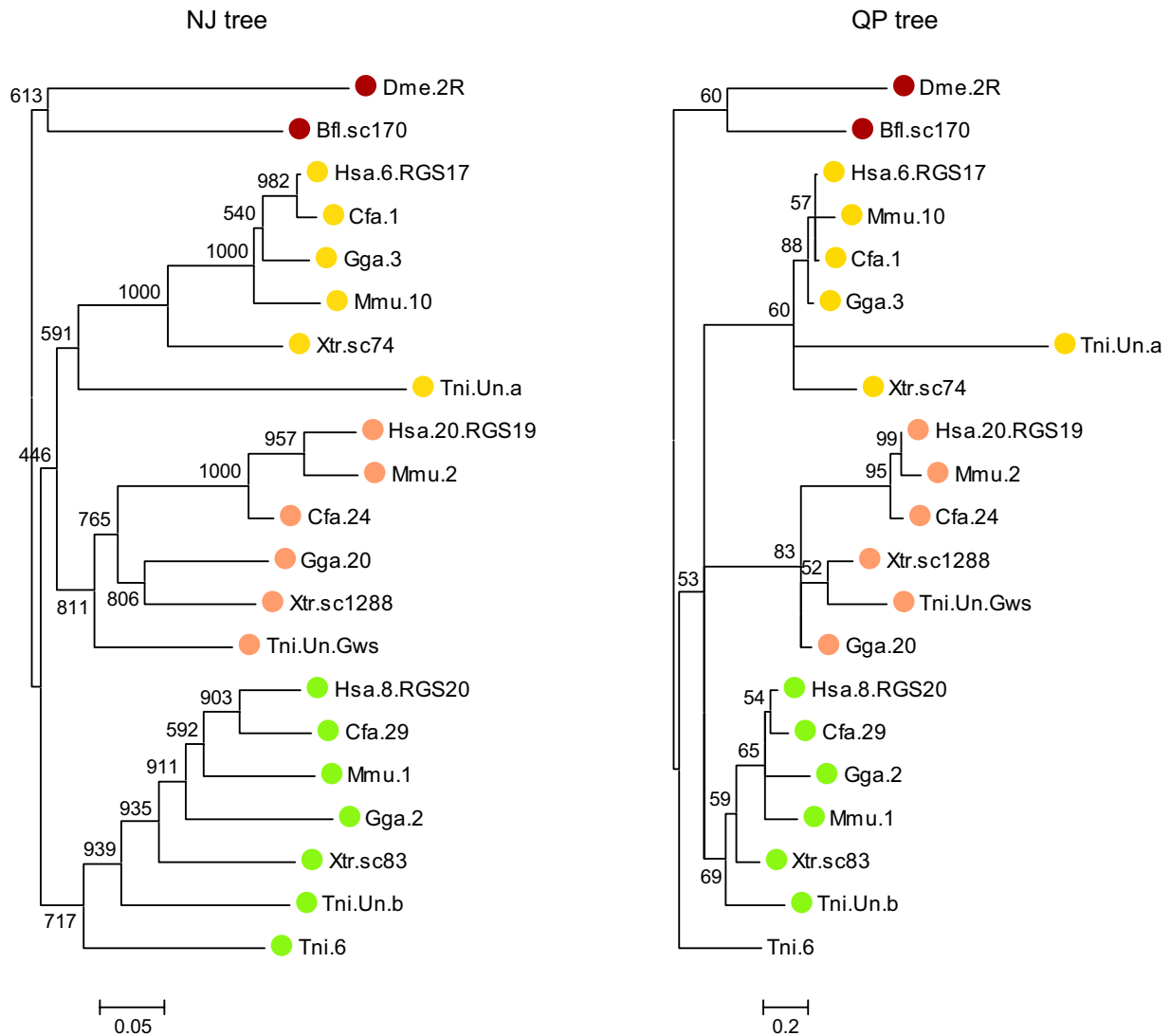


Fig. 58. Neighbor-joining and quartet-puzzling maximum likelihood trees for REGULATOR OF G-PROTEIN SIGNALING (RGS) subfamily RZ/A. RGS proteins activate the intrinsic GTPase activity of the $G\alpha$ subunit of G proteins and thereby increase the hydrolysis rate of the GTP bound to the $G\alpha$. The termination rate of G protein signaling can be increased by a 100-fold or more by the RGS-proteins. There are also indications that RGS proteins can have other functions such as interactions with receptors or intracellular signaling proteins [Nunn C, Mao H, Chidiac P, Albert PR (2006) RGS17/RGS22 and the RZ/A family of regulators of G protein signalling. *Seminars Cell Dev Biol* 17:390–399]. Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Cfa, *Canis familiaris*; Gga, *Gallus gallus*; Xtr, *Xenopus tropicalis*; Tni, *Tetraodon nigroviridis*; Bfl, *Branchiostoma floridae*; Dme, *Drosophila melanogaster*.

SOX

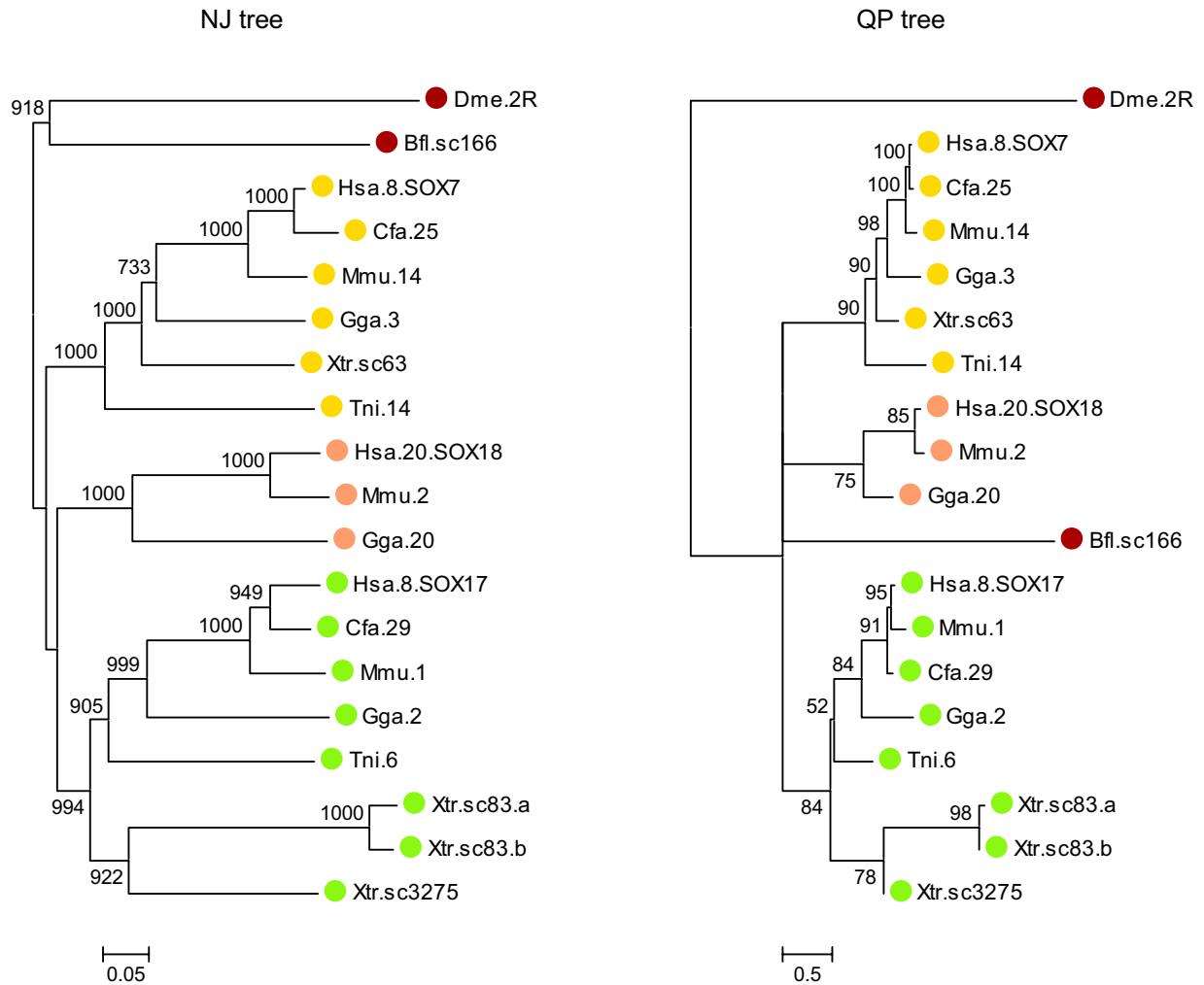


Fig. S9. Neighbor-joining and quartet-puzzling maximum likelihood trees for TRANSCRIPTION FACTOR SOX (SOX). The SOX proteins are transcription factors that contain the DNA binding HMG-box domain. The SOX HMG-domains are well preserved and bind the DNA target sequence, AACAA(A/T)G (1). This Ensembl family of transcription factors has 11 members in humans but can be divided further into subfamilies. The genes located close to OPRK1 and OPRL1, SOX17 and SOX18 respectively, belong to subfamily F together with SOX7 (2). Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Cfa, *Canis familiaris*; Gga, *Gallus gallus*; Xtr, *Xenopus tropicalis*; Tni, *Tetraodon nigroviridis*; Bfl, *Branchiostoma floridae*; Dme, *Drosophila melanogaster*.

1. Koopman P, Schepers G, Brenner S, Venkatesh B (2004). Origin and diversity of the Sox transcription factor gene family: Genome-wide analysis in *Fugu rubripes*. *Gene* 328:177–186.
2. Bowles J, Schepers G, Koopman P (2000) Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev Biol* 227:239–255.

SRC-B

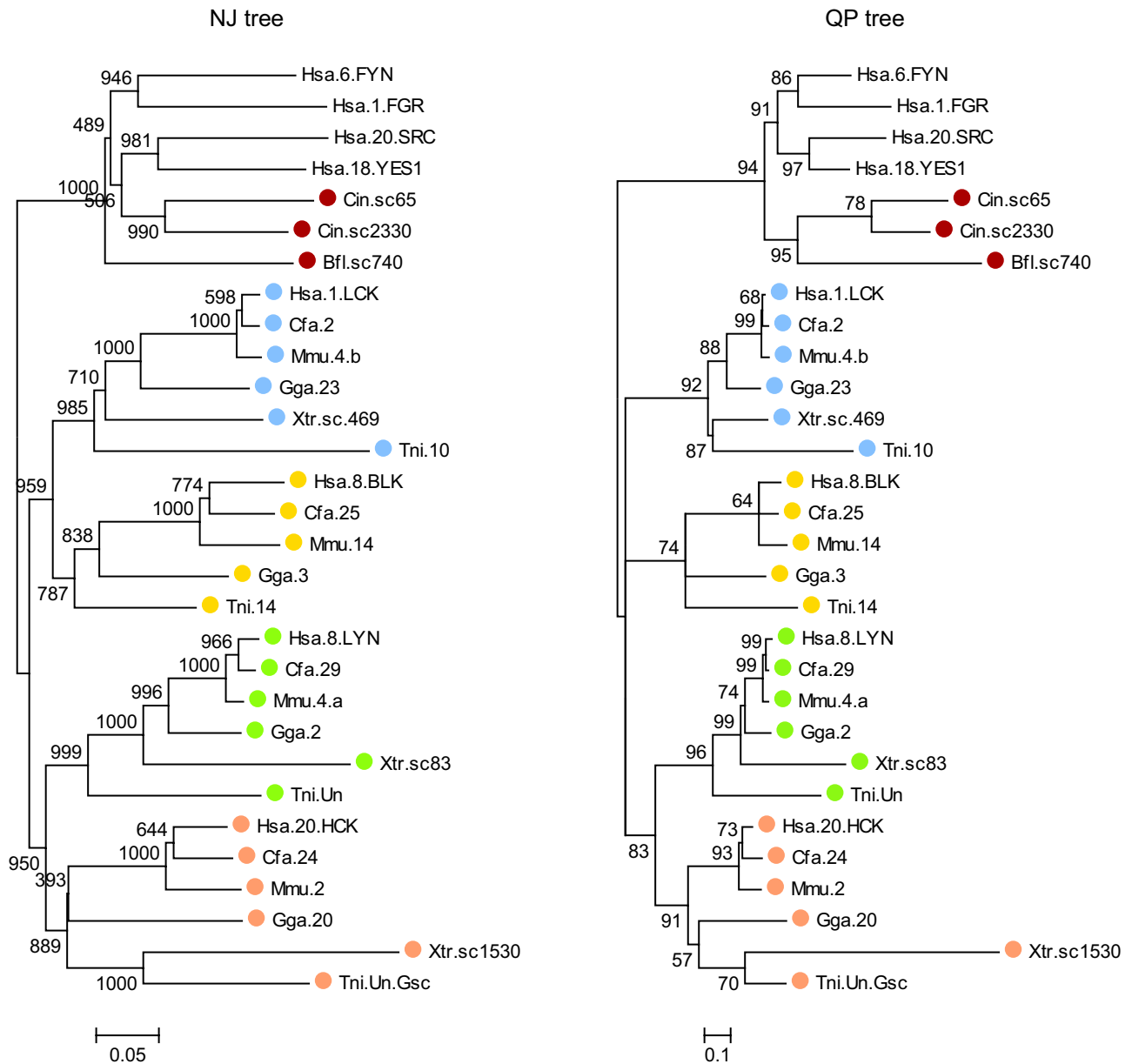


Fig. S10. Neighbor-joining and quartet-puzzling maximum likelihood trees for TYROSINE KINASE. The nonreceptor tyrosine kinases can be divided into 11 sub-families (1) and this Ensembl family consists of five of them. However, only one of these subfamilies, the SRC-B family, has more than one gene in the chromosome location of interest. Therefore, only the SRC-B subfamily was investigated further. *SRC-B family* SRC-B is a subfamily of the nonreceptor protein-tyrosine kinases. They have one Src homology 2 domain (SH2), one Src homology 3 domain (SH3), and a Kinase domain (1). The SRC-B protein-tyrosine kinases are attached to the inside of the cytoplasmic membranes by their N-amino termini and are expressed in specific cells of hematopoietic origin. They are coupled to cytokine receptors and are important for the receptor signaling. (2). The genes included in the SRC-B family are BLK, LYN, HCK and LCK (Robinson *et al.*). Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Cfa, *Canis familiaris*; Gga, *Gallus gallus*; Xtr, *Xenopus tropicalis*; Tni, *Tetraodon nigroviridis*; Bfl, *Branchiostoma floridae*; Cin, *Ciona intestinalis*.

1. Robinson DR, Wu Y-M, Lin S-F (2000) The protein tyrosine kinase family of the human genome. *Oncogene* 19:5548–55572.
2. Neet K, Hunter T (1996). Vertebrate non-receptor protein-tyrosine kinase families. *Genes Cells* 1:147–169.

STMN

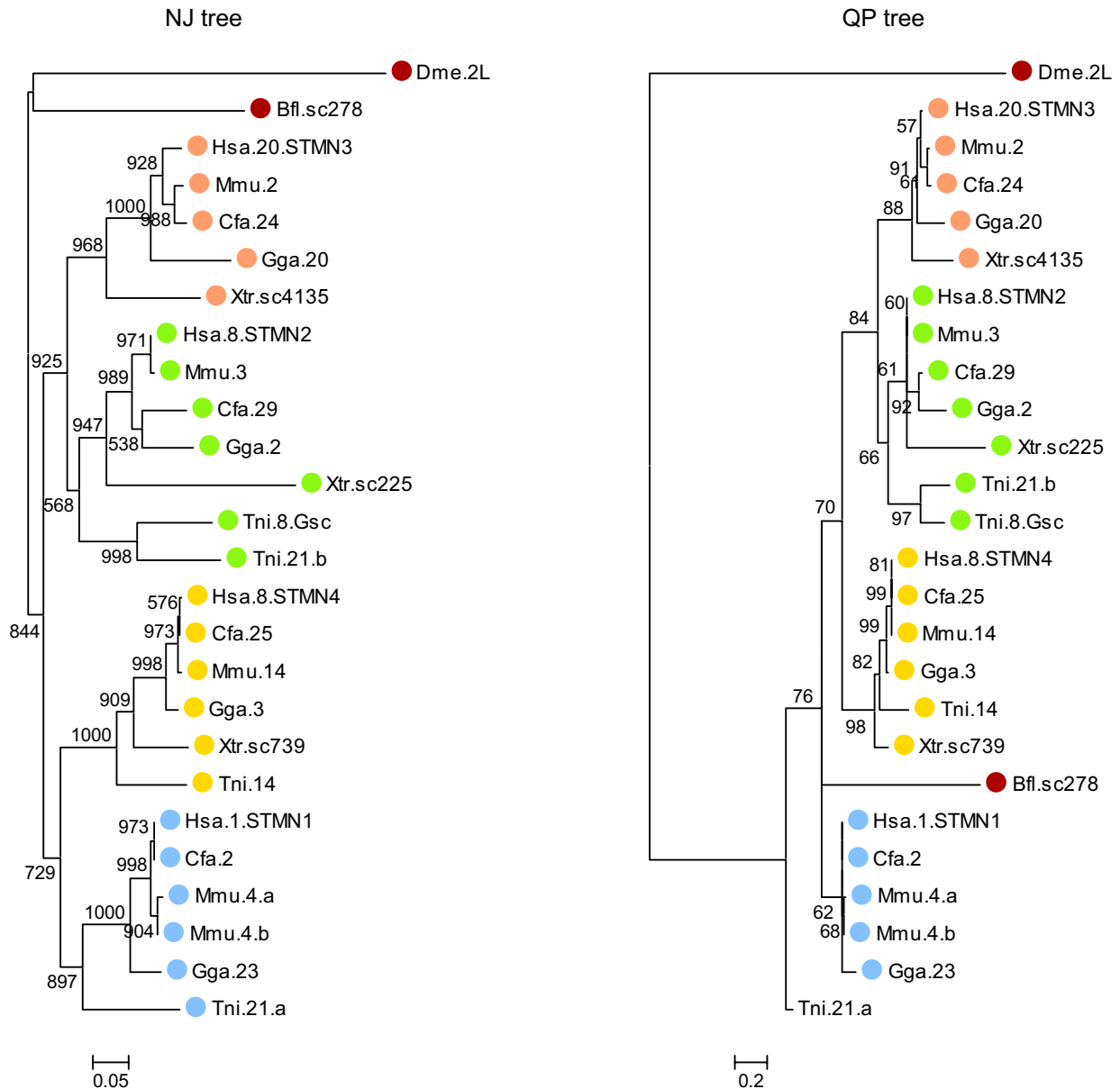


Fig. S11. Neighbor-joining and quartet-puzzling maximum likelihood trees for STATHMIN (STMN). The proteins in this family all have a Pfam Stathmin domain. Four human genes belong to this family. The Stathmin family proteins play important roles in cell proliferation by regulating microtubule growth and shrinkage, important in mitosis. Stathmin can bind free tubulin and thereby prevent microtubule growth but can also bind to the end of microtubules and depolymerize them. Stathmin has four phosphorylation sites and is totally inactivated by phosphorylation of all sites. However, its activity can be regulated and finely tuned by phosphorylating some but not all of the sites to partially inactivated or down regulate its activity (1, 2). Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Cfa, *Canis familiaris*; Gga, *Gallus gallus*; Xtr, *Xenopus tropicalis*; Tni, *Tetraodon nigroviridis*; Bfl, *Branchiostoma floridae*; Dme, *Drosophila melanogaster*.

1. Curmi PA, et al. (1999). Stathmin and its phosphoprotein family: General properties, biochemical and functional interaction with tubulin. *Cell Structure Function* 24:345–357.
2. Rubin CI, Atweh GF (2004) The role of Stathmin in the regulation of the cell cycle. *J Cell Biochem* 93:242–250.

TCEA

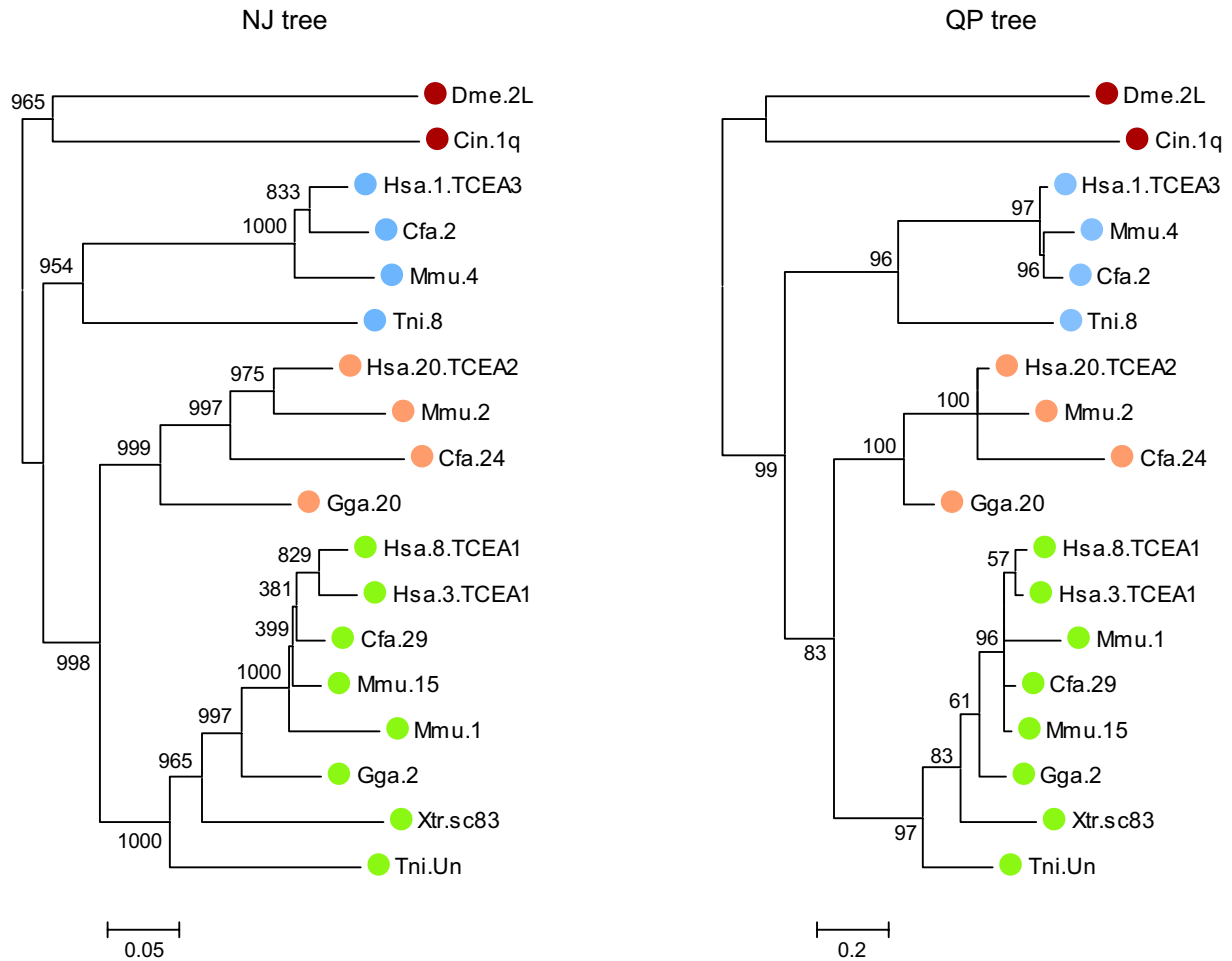


Fig. S12. Neighbor-joining and quartet-puzzling maximum likelihood trees for TRANSCRIPTION ELONGATION FACTOR A (SII) (TCEA). The proteins of this family play an important role in transcription. They bind to polII and when transcription is stalled they restart the arrested polII by stimulating an intrinsic nuclease activity of the polymerase. PolII then cleaves the RNA and thereby makes a fresh 3'-hydroxyl group available to the catalytic site and transcription can continue (Wind M, Reines D (2000) Transcription elongation factor SII. *BioEssays* 22:327–336). Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Cfa, *Canis familiaris*; Gga, *Gallus gallus*; Xtr, *Xenopus tropicalis*; Tni, *Tetraodon nigroviridis*; Dme, *Drosophila melanogaster*; Cin, *Ciona intestinalis*.

XKR

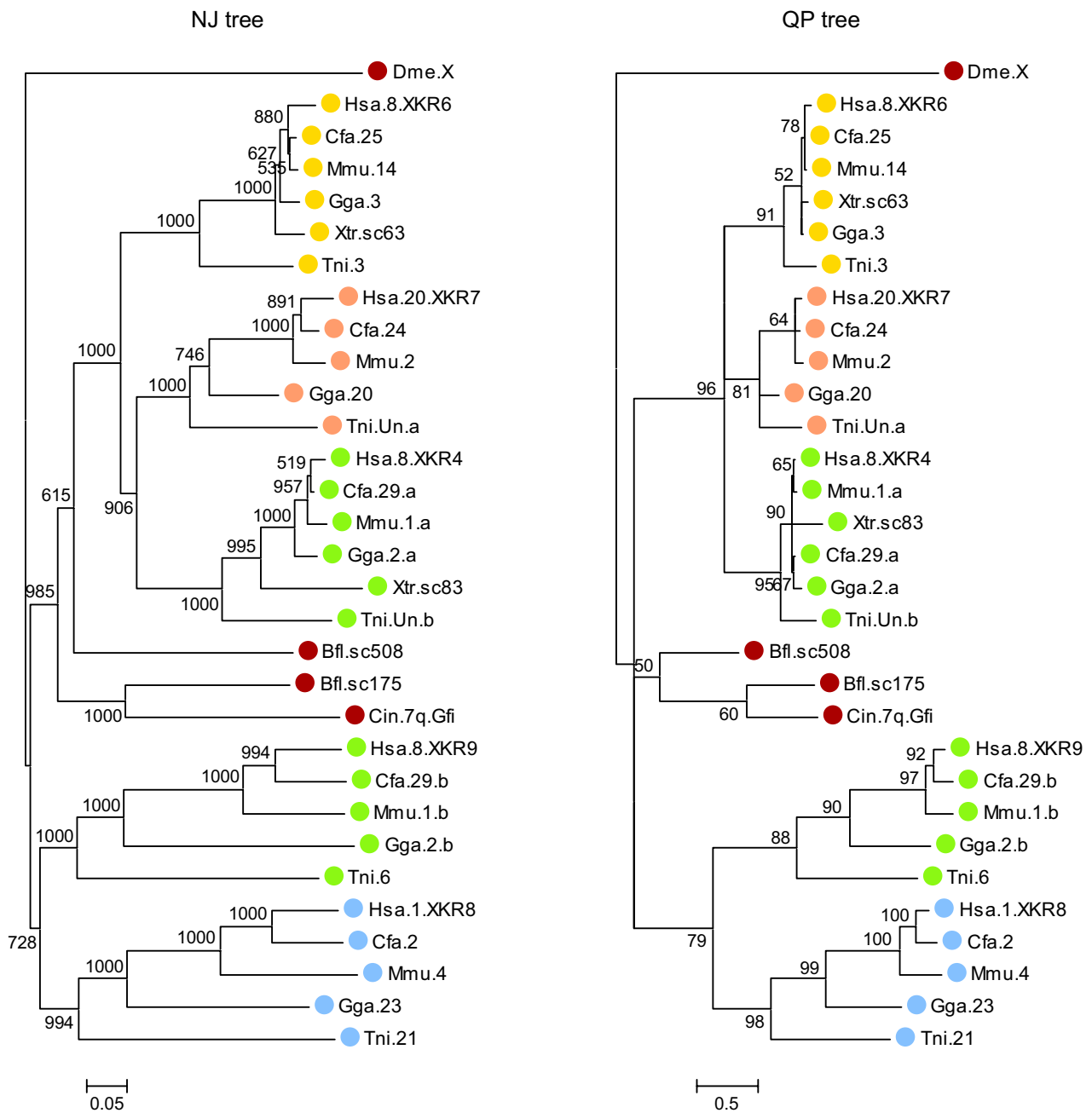


Fig. S13. Neighbor-joining and quartet-puzzling maximum likelihood trees for XK RELATED (XKR). This is a subfamily of the XK (Kell blood group complex) family. Five genes are included in this Ensembl family but our analysis shows that they can be divided into two distinct subfamilies. The proteins have the structural characteristics of membrane transport proteins but the substrate/substrates are not known [Calenda G, *et al.* (2006). Identification of two new members, XPLAC and XTES, of the XK family. *Gene* 370:6–16.]. In humans, the proteins of this family have 6–10 predicted transmembrane regions. However, some of the proteins from other species have as few as two or three transmembrane regions. The varying number of TMs and the different length of the proteins make them difficult to align. Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Cfa, *Canis familiaris*; Gga, *Gallus gallus*; Xtr, *Xenopus tropicalis*; Tni, *Tetraodon nigroviridis*; Bfl, *Branchiostoma floridae*; Dme, *Drosophila melanogaster*; Cin, *Ciona intestinalis*.

YTHDF

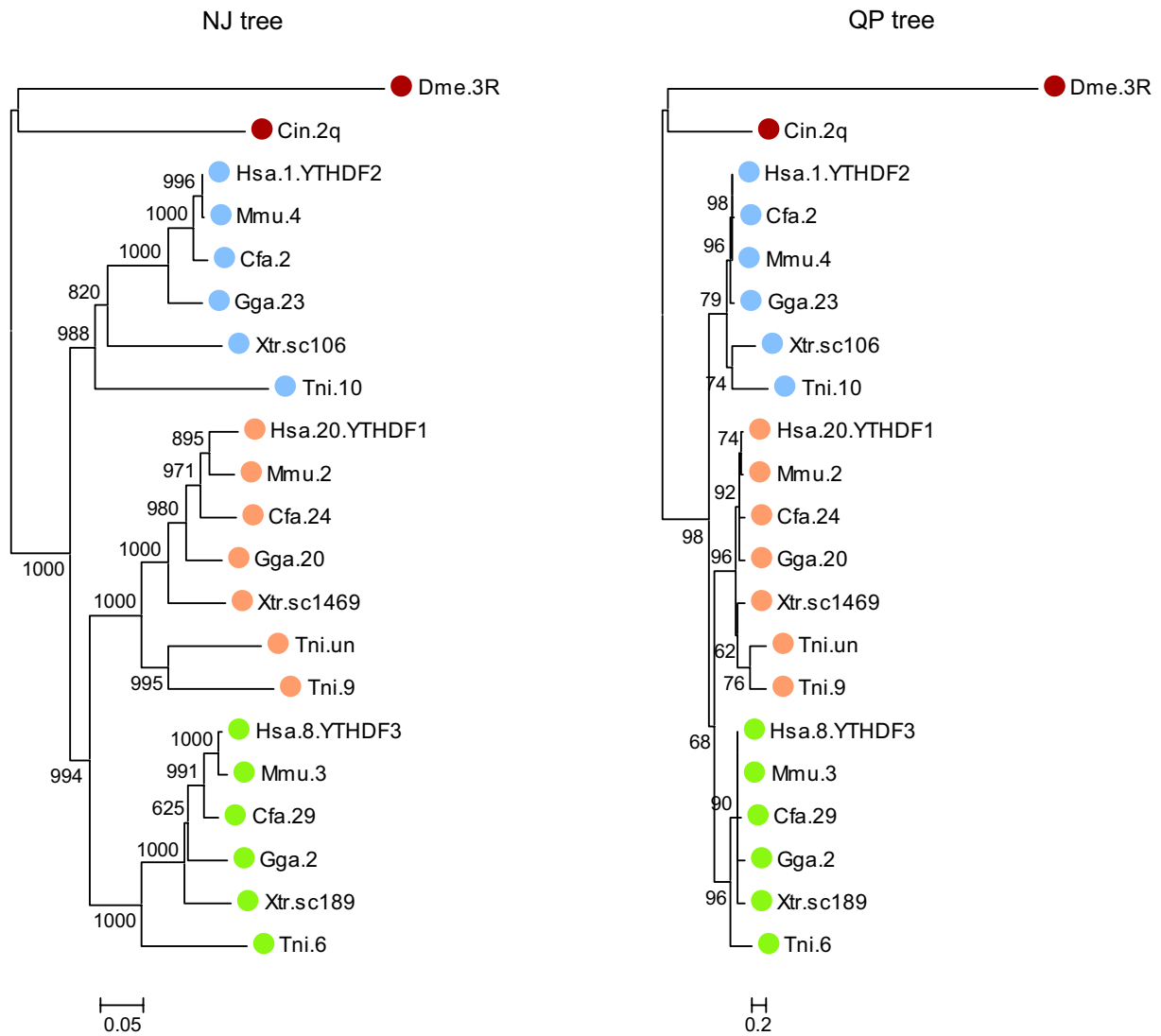


Fig. S14. Neighbor-joining and quartet-puzzling maximum likelihood trees for YTH DOMAIN PROTEIN (YTHDF). The proteins of this family are not characterized but they have a Pfam YTH (YT521-B homology) domain. The YTH domain contains conserved aromatic residues that are similar to aromatic residues in the RNA recognition motif (RRM) domain. The aromatic residues in the RRM domain are crucial for RNA binding it is therefore likely that the YTH domain also binds RNA (1). (The YT521-B protein is not a member of this family but it also contains the YTH domain and can alter splice site selection of the pre-mRNA (1, 2). Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Cfa, *Canis familiaris*; Gga, *Gallus gallus*; Xtr, *Xenopus tropicalis*; Tni, *Tetraodon nigroviridis*; Dme, *Drosophila melanogaster*; Cin, *Ciona intestinalis*.

1. Stoilov P, Rafalska I, Stamm S (2002) YTH: A new domain in nuclear proteins. *Trends Biochem Sci* 27:495–497.
2. Rafalska I, et al. (2004) The intranuclear localization and function of YT521-B is regulated by tyrosine phosphorylation. *Human Mol Gen* 13:1535–1549.

ZDHHC

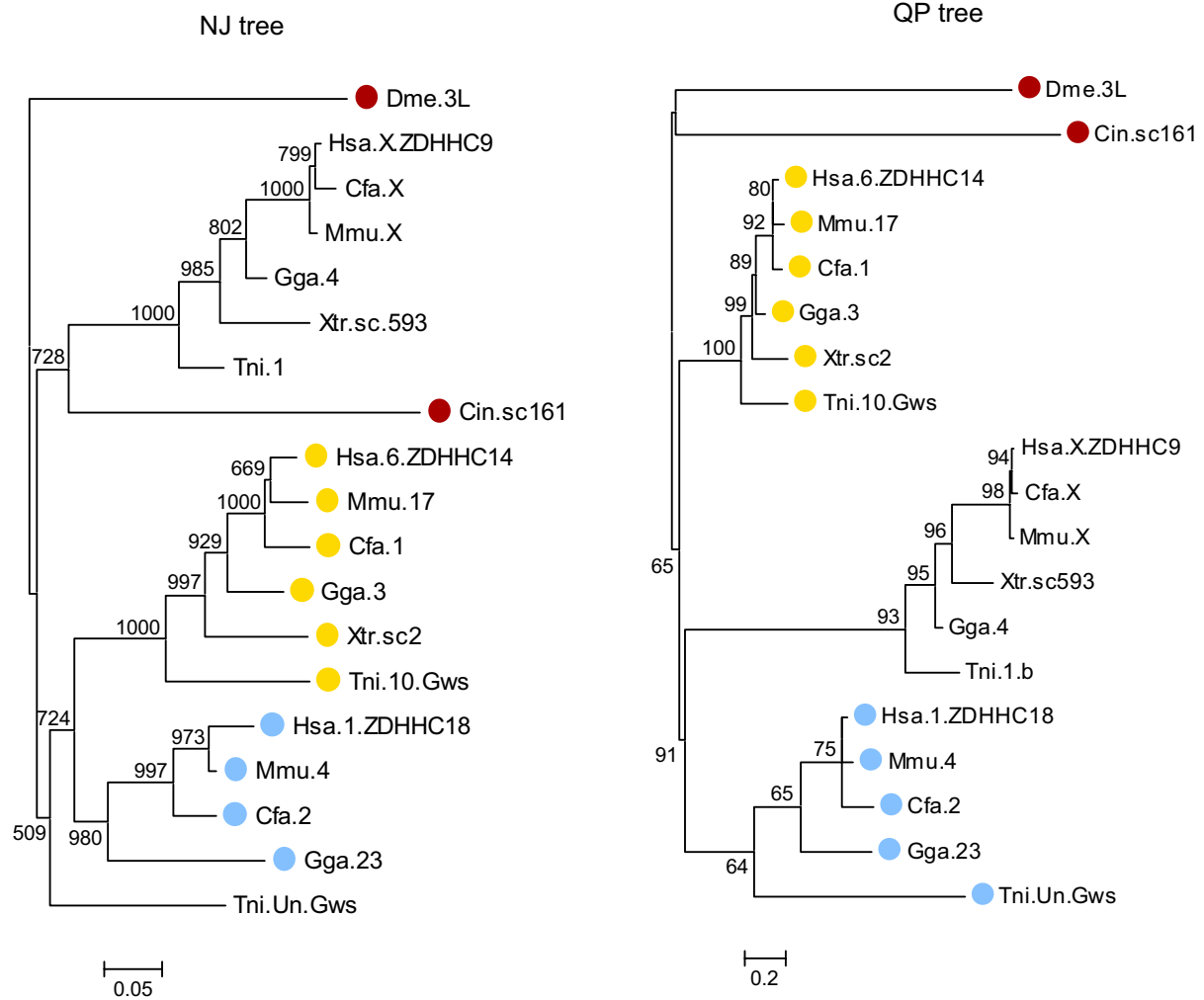


Fig. S15. Neighbor-joining and quartet-puzzling maximum likelihood trees for PALMITOYLTRANSFERASE–ZINC FINGER DHHC DOMAIN CONTAINING (ZDHHC). All proteins in this family have a Pfam zf-DHHC domain, although the domain can also be found in proteins belonging to other families. The proteins of this family are thought to act as palmitoyltransferases, linking long fatty acids (usually palmitoyl) by thioester bonds to the side chains of cysteine (Mitchell DA, Vasudevan A, Linder ME, Deschens RJ (2006) Protein palmitoylation by a family of DHHC protein S-acyltransferases. *J Lipid Res* 47:1118–1127). The Pfam zf-DHHC domain is also predicted to be zinc binding (Pfam). After a phylogenetic analysis, a subfamily to this Ensembl family was found consisting of ZDHHC18, ZDHHC14 and ZDHHC9 in humans. Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Cfa, *Canis familiaris*; Gga, *Gallus gallus*; Xtr, *Xenopus tropicalis*; Tni, *Tetraodon nigroviridis*; Dme, *Drosophila melanogaster*; Cin, *Ciona intestinalis*.


```

      100      110      120      130      140      150      160      170      180
Hsa.8.NPBWR1 PVVYAVICAVGLAGNSAVLVYLLR--APRMKVTNIFILNLAIADLFTLVLPINIADEFLLRQWPFGEGLMCKLIVADQYNTFSSLYFLT
Cfa.29.a PVVYAVICAVGLAGNSAVLVYLLR--APRKTVTNIFILNLAVADELFTLVLPINIADEFLLRQWPFGEGLMCKLIVADQYNTFSSLYFLT
Bta.14.b PVVYAVICAVGLAGNSAVLVYLLR--APRRKVTNIFILNLAVADELFTLVLPVNIADFLRRWPFGEGLLCKLVAVDQYNTFSSLYFLT
Mmu.1.b PVVYGVICAVGLAGNSAVLVYLLR--TPRMKVTNIFILNLAIADLFTLVLPINIADEFLLRWPFGEGLMCKLIVADQYNTFSSLYFLA
Mdo.3.b PVVYGVICAVGLTNTAVLYVLLR--APRMKVTNIFILNLAIADLFTLVLPINIADEFLLRQWPFGEGLMCKLIIISIDQYNTFSSLYFLT
Gga.2.a PVVYIICAVGLTNTAVLYVILK--APMKVTNIFILNLAIADLFTLVLPINIADEFLLRQWPFGEGLMCKLIIISIDQYNTFSSLYFLT
Pti.sc83.a* PVIYSVICAVGLTNTAVLYVILK--APMKVTNIFILNLAIADLFTLVLPINIADEFLLRQWPFGEGLMCKLIIISIDQYNTFSSLYFLT
Cmi.AAVX01065032.1 PVIYSVICAVGLTNTAVLYVILR--APMKVTNIFILNLAIADLFTLVLPINIADEFLLRQWPFGEGLMCKLIIISIDQYNTFSSLYFLT
Hsa.20.NPBWR2 PAVYSGICAVGLTNTAVILVILR--APMKVTNIFILNLAVADGLFTLVLPVNIADHLLQWPFGEGLLCKLVAVDHYNIFSSYFLA
Bta.Un.b PVIYSVICAVGLTNTAVLYVILK--APMKVTNIFILNLAIADLFTLVLPINIADEFLLRQWPFGEGLMCKLIIISIDQYNTFSSLYFLT
Ola.7 PVIYSVICAVGLTNTAVLYVILK--APMKVTNIFILNLAIADLFTLVLPINIADEFLLRQWPFGEGLLCKLVAVDHYNIFSSYFLT
Gac.XII.a PVIYSVICAVGLTNTAVLYVILK--APMKVTNIFILNLAIADLFTLVLPINIADEFLLRQWPFGEGLLCKLVAVDHYNIFSSYFLT
Dre.23..b PVIYSVICAVGLTNTAVLYVILK--APMKVTNIFILNLAIADLFTLVLPINIADEFLLRQWPFGEGLLCKLVAVDHYNIFSSYFLT
Xtr.sc1400.a* PIIYAIICVVGLTNTAVLYVILK--APMKVTNIFILNLAIADLFTLVLPINIADEFLLRQWPFGEGLLCKLVAVDHYNIFSSYFLT
Mdo.1.a* PVMYVICAVGLTNTAVLYVILK--APMKVTNIFILNLAIADLFTLVLPINIADEFLLRQWPFGEGLLCKLVAVDHYNIFSSYFLT
Pma.co7520 PIIYSVICAVGLTNTAVLYVILR--APMKVTNIFILNLAIADLFTLVLPINIADEFLLRQWPFGEGLMCKLIIISIDQYNTFSSLYFLT
Pma.co2020 PIIYSVICAVGLTNTAVLYVILR--APMKVTNIFILNLAIADLFTLVLPINIADEFLLRQWPFGEGLMCKLIIISIDQYNTFSSLYFLT
Cfa.24 VGLYLAVCIGLLGNCLVMYVILR--HTKMKATNIYIFNLALADTLVLLTLPFQGTDLVLLGFWPFGEGLLCKLVAVDHYNIFSSYFLT
Bta.14 LGLYLAVCIGLLGNCLVMYVILR--HTKMKATNIYIFNLALADTLVLLTLPFQGTDLVLLGFWPFGEGLLCKLVAVDHYNIFSSYFLT
Hsa.20.OPRL1 VGLYLAVCIGLLGNCLVMYVILR--HTKMKATNIYIFNLALADTLVLLTLPFQGTDLVLLGFWPFGEGLLCKLVAVDHYNIFSSYFLT
Mmu.2 VGLYLAVCIGLLGNCLVMYVILR--HTKMKATNIYIFNLALADTLVLLTLPFQGTDLVLLGFWPFGEGLLCKLVAVDHYNIFSSYFLT
Mdo.1.b VVYSLVAVCIGLLGNCLVMYVILR--HTKMKATNIYIFNLALADLAVLTLVLPFQGTDFLGFWPFGEGLLCKLVAVDHYNIFSSYFLT
Xtr.sc1400.b -----R--YTKMKATNIYIFNLALADLAVLTLVLPFQGTDFLGFWPFGEGLLCKLVAVDHYNIFSSYFLT
Tgr.L VIVYLIVCIVGLVGNCAVMFVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQGTDFLGFWPFGEGLLCKLVAVDHYNIFSSYFLT
Gga.20 VVYSLVAVCIGLLGNCLVMYVILR--YTKMKATNIYIFNLALADLAVLTLVLPFQGTDFLGFWPFGEGLLCKLVAVDHYNIFSSYFLT
Gac.XII.b VVVYMIYCVIGLVGNFLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQGTDFLGFWPFGEGLLCKLVAVDHYNIFSSYFLT
Tni.Un.b* VVVYMIYCVIGLVGNFLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQGTDFLGFWPFGEGLLCKLVAVDHYNIFSSYFLT
Dre.23.a AVVYMIYCVIGLVGNFLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQGTDFLGFWPFGEGLLCKLVAVDHYNIFSSYFLT
Hsa.8.OPR1 TAVYSVVFVGLVGNLVMFVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLMNSWPFGEGLLCKLVAVDHYNIFSSYFLT
Mmu.1.a TAVYSVVFVGLVGNLVMFVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLMNSWPFGEGLLCKLVAVDHYNIFSSYFLT
Cfa.29.b TAVYSVVFVGLVGNLVMFVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLMNSWPFGEGLLCKLVAVDHYNIFSSYFLT
Bta.14.a TAVYSVVFVGLVGNLVMFVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLMNSWPFGEGLLCKLVAVDHYNIFSSYFLT
Mdo.3.a TAVYSVVFVGLVGNLVMFVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLMNSWPFGEGLLCKLVAVDHYNIFSSYFLT
Gga.2.b TAVYSVVFVGLVGNLVMFVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLMNSWPFGEGLLCKLVAVDHYNIFSSYFLT
Tgr.K TAVYSVVFVGLVGNLVMFVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLMNSWPFGEGLLCKLVAVDHYNIFSSYFLT
Xtr.sc83.b* TAVYSVVFVGLVGNLVMFVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLMNSWPFGEGLLCKLVAVDHYNIFSSYFLT
Ola.17 VAVYSVVFVGLVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Tni.Un.a* VAVYSVVFVGLVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Gac.III VAVYSVVFVGLVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Dre.2 VAVYSVVFVGLVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Ola.20* TAVYSVVFVGLVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Gac.XXI* TAVYSVVFVGLVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Tni.6* TAVYSVVFVGLVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Gac.XX TALYSLICVVGLLGNLVMYGVVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Tni.8* TALYSLICVVGLLGNLVMYGVVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Ola.11 TALYSLICVVGLLGNLVMYGVVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Gac.X TALYSVICVVGLLGNLVMYGVVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Dre.19 TALYSVICVVGLLGNLVMYGVVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Dre.16 TALYSVICVVGLLGNLVMYGVVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Hsa.1.OPRD1 TALYSVAVCAVGLLGNLVMYGVVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Bta.2 TALYSVAVCAVGLLGNLVMYGVVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Cfa.2 TALYSVAVCAVGLLGNLVMYGVVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Mmu.4 TALYSVAVCAVGLLGNLVMYGVVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Mdo.4 TALYSVAVCAVGLLGNLVMYGVVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Xtr.sc478 TALYSVAVCAVGLLGNLVMYGVVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Tgr.D TALYSVAVCAVGLLGNLVMYGVVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Gga.23 TALYSVAVCAVGLLGNLVMYGVVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Hsa.6.OPRM1 MALYSIVCVVGLFVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Cfa.1 MALYSIVCVVGLFVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Bta.9 MALYSIVCVVGLFVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Mmu.10 MALYSIVCVVGLFVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Mdo.2 MALYSIVCVVGLFVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Gga.3 MALYSIVCVVGLFVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Xtr.sc74* TALYSIVCVVGLFVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Tgr.M TALYSIVCVVGLFVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Ola.15 -----YLMGSWPFGEGLLCKLVAVDHYNIFSSYFLT
Gac.VI TALYSIVCVVGLFVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Dre.13 TALYSIVCVVGLFVGNLVMYVIVR--YTKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Hsa.14.SSTR1 SFYYSVCLVGLCGNSLVYVILR--YAKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Hsa.20.SSTR4 QCIYALVCLVGLCGNSLVYVILR--YAKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Hsa.16.SSTR5 PVLVLLVCAALGGNTLVYVILR--FAKMKVTNIFILNLAVADLVLVYMLGLPLLATQNAASFVWPFGEGLLCKLVAVDHYNIFSSYFLT
Hsa.22.SSTR3 PLVYLVVCCVGLLGNLVMYVIVR--HTASPSVTNIFILNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Hsa.17.SSTR2 PFIYFVVCIIIGLGNLTVYVILR--YAKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Bf1.sc203 PFIYFVVCIIIGLGNLTVYVILR--YAKMKATNIYIFNLALADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Dme.3L VVLYGFCVCIIGLGNLTVYVILR--FSKMQVTNIFILNLAVADLAVLTLVLPFQSTVYLLSSWPFGEGLLCKLVAVDHYNIFSSYFLT
Clustal Consensus * * * * *

```

Fig. S17 (continued).

```

      190      200      210      220      230      240      250      260      270
Hsa.8.NPBWR1  VMSADRYLVVLATAESRRVAGRTYSAARAVSLAVNGVITLVVLPFAVFAFLDDEQGR---QCVLVFPQP---EAFWWRASRLYTLV
Cfa.29.a      VMSADRYLVVLATAESRRVAGRTYSAARAVSLAVNGVITLVVLPFAVFAFLDDEQGR---QCVLVFPQP---EAFWWRASRLYTLV
Bta.14.b      VMSADRYLVVLATAESRRVAGRTYGAARAVSLAVNGVATLVVLPFAVFAFLDDEBQGR---QCVLVFPQP---EALWWRASRLYTLV
Mmu.1.b      VMSADRYLVVLATAESRRVSGRTYGAARAVSLAVNGVITLVVLPFAVFAFLDDEBQGR---QCVLVFPQP---EAFWWRASRLYTLV
Mdo.3.b      VMSIDRYLVVLATIESRKMAYRTYRAAKVSLAVWFLVTLVLPFTTIFAQLYEBEGRI---QCVLVFPQP---ENFWWKISRIYTLI
Gga.2.a      VMSVDRYLVAATTKSRKMSYRTYRAAKIVSLCISWFSVTIILPFTVFAKVHKEBQGRS---QCVVFPHP---ESVWVKGSRIYTLI
Xtr.sc83.a*   VMSIDRYLVVATVSKKLSYRTYRAAKIVSLCISWFSVTIILPFTVFAKVHKEBGR---QCLVFPNP---ESLWQMSRIYTLI
Cmi.AAVX01065032.1 VMSIDRYLVVLATVRSKMSYRTYRAAKIVSICVWLFTVIIISPTVFSKIDSEBQGR---QCVVFPHP---ELIWKASRIYTLI
Hsa.20.NPBWR2 VMSVDRYLVLVLTAVRSRHPWRRTYRGAASLVCVWLGVTVLVLPFFSFAQVYVSNELQV---PSCGLSFPWP---EQVWFKASRYVTLV
Bta.Un.b     VMSIDRYLVVLTAVRSRMRPRRTYHRAKVASLVCVWLGVTAVLPLFTFAGVYVNNELQV---TSCGLSFPWP---ERAWFQASRIYTLV
Ola.7        VMSIDRYLVVLATVRSKMRPYRTYRAAKIISFCVWLIVLIVIPFTVFAQVYVNNPFDG-R--KRCGLSLPNP---ETLWFRTRISYITLI
Gac.XII.a    VMSIDRYLVVLATVSSKMRPYRTYRAAKIISLVCVWLIVLIVSPYTVFAQVYERPNDG-R--KSCVLSFPSP---EGSWITTSRIYITLI
Bta.Un.b     VMSVDRYLVLVLTAVRSRMRPRRTYHRAKVASLVCVWLLVIVLMPFVFAQVYVNNELQV---TSCGLSFPWP---ERAWFQASRIYTLV
Xtr.sc1400.a* VMSIDRYLVVLATVRSKMRPYRTYRAAKIVSISLWLLVIIIVLFPFTTIFAQVYVMDMDP---KSCGLNFPKP---EKLWFKASRIYTLI
Mdo.1.a*     VMSVDRYLVLVLTAVQSRRLPYRTYHAARTTSCICWLVWLVIVLFPFFIFASVYTNELQI---KSCGLSFPQP---ERFWFKASRIYTLV
Pma.co7520   VMSADRFVVTAVKSENLEPRRTYRNKIVSLCISWFSVTIILPFTTIFAST-YVDFE-R--NHCGLDFPKP---EPNWLKGSRIYTLV
Pma.co2020   VMSIDRYLVVLATVKSRSFTWRTYAVAKWVCGGVWALVSLVLPFFVFAVSNQIVDSQR-H--TRCGLLELSP---EAEWLRAVRYTTLI
Cfa.24       VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQVE--DE--EIECLVEIPTP---QDYWGPVFAICIFL
Bta.23.a     VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQVE--DE--EIECLVEIPTP---QDYWGPVFAICIFL
Hsa.20.OPRL1 VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQVE--DE--EIECLVEIPTP---QDYWGPVFAICIFL
Mmu.2        VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQVE--DE--EIECLVEIPAP---QDYWGPVFAICIFL
Mdo.1.b      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Xtr.sc1400.b VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Tgr.L        VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Gga.20       VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Gac.XII.b    VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Tni.Un.b*    VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Dre.23.a     VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Hsa.8.OPRK1  VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Mmu.1.a      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Cfa.29.b     VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Bta.14.a     VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Mdo.3.a      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Gga.2.b      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Tgr.K        VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Xtr.sc83.b*  VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Ola.17       VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Tni.Un.a*    VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Gac.III      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Dre.2        VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Ola.20*     VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Gac.XXI*    VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Tni.6*      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Gac.XX      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Tni.8*      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Ola.11      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Gac.X       VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Dre.19      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Dre.16      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Hsa.1.OPRDL VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Bta.2       VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Cfa.2       VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Mmu.4       VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Mdo.4       VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Xtr.sc478   VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Tgr.D       VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Gga.23      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Hsa.6.OPRM1 VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Cfa.1       VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Bta.9       VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Mmu.10      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Mdo.2       VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Gga.3       VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Xtr.sc74*   VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Tgr.M       VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Ola.15      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Gac.VI      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Dre.13      VMSVDRYVAICHPIRA--LDVRTSSKAQAVNVAIWAALASVGVVPAIMGSAQMD--DD--EIECLVEIPAP---QDYWGPVFAICIFL
Hsa.14.SSTR1 VLSVDRYVAIVHPKIA--ARYRRPTVAKVNLGVWVLSLVLVLPVIVVFSRTAAN--SDG-TVACNMLMP---EP-AQRWLGVPLVYTLF
Hsa.20.SSTR4 VLSVDRYVAIVHPKIA--ARYRRPTVAKVNLGVWVLSLVLVLPVIVVFSRTAAN--SDG-TVACNMLMP---EP-AQRWLGVPLVYTLF
Hsa.16.SSTR5 VLSVDRYVAIVHPKIA--ARYRRPTVAKVNLGVWVLSLVLVLPVIVVFSRTAAN--SDG-TVACNMLMP---EP-AQRWLGVPLVYTLF
Hsa.22.SSTR3 VMSVDRYLAIVHPKIA--ARWRPTVAKVNLGVWVLSLVLVLPVIVVFSRTAAN--SDG-TVACNMLMP---EP-AQRWLGVPLVYTLF
Hsa.17.SSTR2 VMSIDRYLAIVHPKIS--AKWRPTVAKVNLGVWVLSLVLVLPVIVVFSRTAAN--SDG-TVACNMLMP---EP-AQRWLGVPLVYTLF
Bfl.sc203    VMSVDRYLAIVHPKIS--AKWRPTVAKVNLGVWVLSLVLVLPVIVVFSRTAAN--SDG-TVACNMLMP---EP-AQRWLGVPLVYTLF
Dme.3L      VMSVDRYLAIVHPKIS--AKWRPTVAKVNLGVWVLSLVLVLPVIVVFSRTAAN--SDG-TVACNMLMP---EP-AQRWLGVPLVYTLF
Clustal Consensus :* **:.: . : * * : * *

```

Fig. S17 (continued).

	370	380	390	400	410	420	430	440	450																		
Hsa.8.NPBWR1	SLSYANSC	LNPF	FLYAF	LDD	SFR	NLR	QLIT	CR	-----	AAAW																	
Cfa.29.a	SLSYANSC	LNPF	FLYAF	LDD	SFR	KSLR	QLL	ACR	-----	AAA																	
Bta.14.b	SLSYANSC	LNPF	FLYAF	LDD	SFR	RS	LR	QLL	ACR	-----	TTS																
Mmu.1.b	SLSYANSC	LNPF	FLYAF	LDD	SFR	RS	LR	QLV	SCR	-----	SA																
Mdo.3.b	SLSYANSC	LNPF	FLYAF	LDD	NFR	RS	FR	KLV	ECR	-----	ASP																
Gga.2.a	SLSYANSC	FNPF	FLYAF	LDD	SFR	RS	FR	KLM	DCR	-----	TTS																
Xtr.sc83.a*	SLSYANSC	LNPF	FLYAF	LDD	SFR	KS	FR	KLL	ECR	-----	SS																
Cmi.AAVX01065032.1	SLSYANSC	LNPF	FLYAF	LDD	SFR	KS	FR	KLL	DCR	-----	A																
Hsa.20.NPBWR2	SLSYANSC	LNPF	FLYAF	LDD	NFR	KNF	FR	SIL	RC	-----	-----																
Bta.Un.b	SLSYTSS	CLNPF	FLYAF	LDD	SFR	KS	LRT	A	CR	CO	-----	GA															
Ola.7	SLSYANSC	LNPF	FLYAF	LDD	SFR	KAF	KM	L	ECR	-----	PA																
Gac.XII.a	SLSYANSC	LNPF	FLYAF	LDD	SFR	KAF	KM	S	ECR	-----	PA																
Dre.23..b	SLSYANSC	LNPF	FLYAF	LDD	SFR	KAF	KM	L	ECR	-----	PA																
Xtr.sc1400.a*	SLSYANSC	LNPF	FLYAF	LDD	SFR	KS	FR	KLL	EC	-----	PA																
Mdo.1.a*	SLSYTNS	CLNPF	FLYAF	LDD	NFR	KS	FR	KM	L	ECR	-----	AT															
Pma.co7520	SLSYANSC	LNPF	FLYAF	LDD	VNFR	NNF	Q	KL	L	ECR	-----	VAS															
Pma.co2020	SLSYTNS	CLNPF	LLYAF	LDD	SFR	RS	FL	KL	L	ECR	-----	AG															
Cfa.24	ALGYVNS	CLNPF	ILYAF	LDD	NFK	CF	RK	F	CC	AP	-----	ALRREMQVSDRVRSTIAK	DVA														
Bta.13	ALGYVNS	GLNPF	ILYAF	LDD	NFK	CF	RK	F	CC	AS	-----	TLRREMQVSDRVRSTIAK	DVA														
Hsa.20.OPRL1	ALGYVNS	CLNPF	ILYAF	LDD	NFK	CF	RK	F	CC	AS	-----	ALRRDQVSDRVRSTIAK	DVA														
Mmu.2	ALGYVNS	CLNPF	ILYAF	LDD	NFK	CF	RK	F	CC	AS	-----	ALHREMQVSDRVRSTIAK	DVG														
Mdo.1.b	VLGYANS	GLNPF	ILYAF	LDD	NFK	CF	RK	F	CC	AS	-----	SLRRELQVSDRVRSTIAK	DVA														
Xtr.sc1400.b	ALGYVNS	SSLNPF	VLYAF	LDD	NFK	CF	RK	F	CF	FP	-----	AFRPELQMSNRMCSTIAK	DVA														
Tgr.L	ALGYVNS	SSLNPF	VLYAF	LDD	NFK	CF	RK	F	CF	FP	-----	AFRSELQMSNRMCSTIAK	DVA														
Gga.20	ALGYANS	SSLNPF	VLYAF	LDD	NFK	CF	RK	F	CF	FP	-----	AFRTELQMSNRMCSTIAK	DVA														
Gac.XII.b	ALGYVNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	PFRLDQQSGRMRSTIAR	EVA													
Tni.Un.b*	ALGYVNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	SFHLDTQQSRRMRSTIAR	EVA													
Dre.23.a	ALGYVNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	RFGIDAQQSGRMRSTIAR	EVA													
Hsa.8.OPRK1	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	KMRMERQS--TSRVRNT	VQDPA													
Mmu.1.a	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	KMRMERQS--TSRVRNT	VQDPA													
Cfa.29.b	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	KMRMERQS--TSRVRNT	VQDPA													
Bta.14.a	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	KMRMERQS--TSRVRNT	VQDPA													
Mdo.3.a	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	R-RMERQS--TSRVRNT	VQDTP													
Gga.2.b	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	KMRMRQS--TSRVRNT	VQDPA													
Tgr.K	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	KIRMERQG--NSRVRNT	IHDPA													
Xtr.sc83.b*	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	KHRLDRQP--NSRVRNT	VQDPA													
Ola.17	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	AVGGQDCQG-VSRVRST	LRDHT													
Tni.Un.a*	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	GQQRECCQG-VSRVRST	LRDHI													
Gac.III	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	AQGHDRSHG-LSRVRST	LRDHS													
Dre.2	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	RTAGDGRG-VSRVRST	LRDHT													
Ola.20*	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	RLKGEKMSG-SKTPST	LQEAA													
Gac.XXI*	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	KLKGEKMSG-SRRRTAS	AREAG													
Tni.6*	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	KLKGEKMSG-SRRRTAS	AREAG													
Gac.XX	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	YRSRLEQSSFSRARN	STKEPL													
Tni.8*	ALGYMNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	RRSRLEQSSFSRARN	TTRBVP													
Ola.11	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	CRTRVEQNSMTSGRARN	VIREPV													
Gac.X	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	CRTHMQQHSITKGRN	NTRBLV													
Dre.19	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	FRTRADQSNLNRARN	ATREPV													
Dre.16	ALGYMNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	FRTRIEQNSFSKARS	VIREPI													
Hsa.1.OPRD1	ALGYANS	SSLNPF	VLYAF	LDD	NFK	CF	R	Q	L	CR	-----	KPCGRPDPSFSRAREA	TARERV														
Bta.2	ALGYANS	SSLNPF	VLYAF	LDD	NFK	CF	R	Q	L	CR	-----	MPCGRPEPSSFSRAREA	TARERV														
Cfa.2	ALGYANS	SSLNPF	VLYAF	LDD	NFK	CF	R	Q	L	CR	-----	SPCGRPEPSSFSRAREA	TARERV														
Mmu.4	ALGYANS	SSLNPF	VLYAF	LDD	NFK	CF	R	Q	L	CR	-----	TPCGRQEPGSLRRPRQA	TTRERV														
Mdo.4	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	RRGPHREPSFSRAREA	TTRERV													
Xtr.sc478	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	FRSHSEQSSFSRARN	TTRDQV													
Tgr.D	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	FRARMEQSSFTRAKN	ATRERV													
Gga.23	ALGYTNS	SSLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	FRARVEQNSFSRARN	TTRERV													
Hsa.6.OPRM1	ALGYTNS	CLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	TSSNIEQQNSTRI	RQNTRDHP													
Cfa.1	ALGYTNS	CLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	TSSTIEQQNSTRI	RQNTRDHP													
Bta.9	ALGYTNS	CLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	TSSTIEQQNSTRI	RQNTRDHP													
Mmu.10	ALGYTNS	CLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	TSSTIEQQNSARI	RQNTREHP													
Mdo.2	ALGYTNS	CLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	TSSTMEQQNSTRI	RHNTRDHP													
Gga.3	ALGYTNS	CLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	TSSTIEQQNSTRV	RQNTRDHA													
Xtr.sc74*	ALGYTNS	CLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	TSSTIEQQNSTRM	RHNTRDHA													
Tgr.M	ALGYTNS	CLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	TSSTAEQQNSIRV	RHNTRDHH													
Ola.15	ALGYTNS	SSLNPF	VLYG	LDD	NFK	CF	R	F	E	CK	PS	-----	SPSVLEIQNSRRTGATSRKVPKRE-H	H													
Gac.VI	ALGYTNS	SSLNPF	VLYG	LDD	NFK	CF	R	F	E	CK	PS	-----	SPSALEMQNSRRTGVTSRKLPQRE-H	H													
Dre.13	ALGYTNS	CLNPF	VLYAF	LDD	NFK	CF	R	F	E	CK	PS	-----	SPSVLDLQNSTRS	RNPQDGG													
Hsa.14.SSTR1	ILGYANS	CANP	ILY	G	F	L	S	D	N	F	R	R	F	F	Q	R	V	L	C	L	R	C	-----	SWMDNAABEFPVYYATALKSR	A		
Hsa.20.SSTR4	ILSYANS	CANP	ILY	G	F	L	S	D	N	F	R	R	F	F	Q	R	V	L	C	L	R	C	-----	LLEGAGGABEELDYATALKSKGGAGCM	H		
Hsa.16.SSTR5	ILSYANS	CANP	ILY	G	F	L	S	D	N	F	R	R	F	F	Q	R	V	L	C	L	R	C	-----	LRKGGAKDAD	ATEP		
Hsa.22.SSTR3	ALPYANS	CANP	ILY	G	F	L	S	D	N	F	R	R	F	F	Q	R	V	L	C	L	R	C	-----	ALPYANS	-----		
Hsa.17.SSTR2	VLTYANS	CANP	ILY	G	F	L	S	D	N	F	R	R	F	F	Q	R	V	L	C	L	R	C	-----	LVKVS	GD		
Bfl.sc203	LSYANS	C	V	N	P	I	L	Y	A	F	L	S	D	N	F	R	R	F	F	Q	R	V	L	C	-----	RRVNSRKAQQNGGTDSTRIEMRSPGGGN	RS
Dme.3L	ALVYSNS	AVNP	I	L	Y	A	F	L	S	D	N	F	R	R	F	F	Q	R	V	L	C	L	R	C	-----	ALVYSNS	-----
Clustal Consensus	* * . * * * . * . *																										

Fig. S17 (continued).

Other Supporting Information Files

[Table S1](#)