



Supplementary Materials

A New Census of Protein Tandem Repeats and their Relationship with Intrinsic Disorder



Figure S1. Summary of the growth of UniProtKB/Swiss-Prot protein knowledgebase. The last protein census dates back to the year 1999 [1]. Since then, the entries in the UniProtKB/Swiss-Prot protein knowledgebase are grown more than seven-fold. Figure from release 2018_09 statistics. https://web.expasy.org/docs/relnotes/relstat.html, retrieved on 2018/10/17.



Figure S2. Proteins with \ge 4 distinct TR regions are sorted by their TR type and shown in different colours for each super kingdom. One can clearly see, that overall super kingdoms small TRs dominate in proteins with many distinct TR regions.



Figure S3. The fraction of proteins containing TRs overall protein entries in UniProtKB/Swiss-Prot is shown for a selection of heavily investigated species and displayed as function of the mean protein length. The tendency of more complex organisms having more and longer TRs can be seen.



Figure S4. The fraction of proteins with homo TRs as a function of sequence length by kingdom resulting in a linear relationship.



Figure S5. The fraction of proteins with micro TRs as a function of sequence length by kingdom resulting in a linear relationship.



Figure S6. The fraction of proteins with small TRs as a function of sequence length by kingdom resulting in a linear relationship.



Figure S7. The fraction of proteins with domain TRs as a function of sequence length by kingdom resulting in a linear relationship.



Figure S8. The amount of TRs (normalized by the amount of protein entries of the species) is displayed separately for each TR-type as a function of the mean length of the proteins. It can be clearly seen, that TRs appear mostly as small TRs. Comparing the fraction of TRs kingdom-wise, some clear tendencies can be seen for micro- and small TRs. For example, chloroplastic proteins with unknown Kingdom tend to have few TRs and short mean protein length. Where in contrast mitochondrial proteins from Viridi plantae and Fungi tend to have many TRs and long mean protein length.



Figure S9. Density plot for the relative positions of TRs within proteins overall super kingdoms. The relative position refers with 0 to the N-terminus and with 1 to the C-terminus of a protein. Colours indicate repeat unit lengths. Interestingly, shorter TRs are biased towards the flanks of the protein.



Figure S10. Density plots for the relative positions of TRs within proteins for four Superkindoms. The relative position referes with 0 to the N-terminus and with 1 to the C-terminus of a protein. Colours indicate repeat unit lengths. Interestingly, short TRs are biased towards the flanks of the protein. In particular for Eukaryotes, there is a clear correlation between TR unit length and location bias to the protein flanks. For Eukaryotes, TRs are particularly prevalent in the N-terminal protein flank. homorepeats in Archaea and, to a lesser degree, in Bacteria show a strong bias to the C-terminal protein flank.



Figure S11. Density plots of position of disorder regions within the protein overall super kingdoms. Both short and long disorder regions tend to cluster towards the flank of proteins, to the N-terminal specifically.



Figure S12. The amino acid ratio was calculated by the number of appearance of each amino acid divided by the overall number of amino acids per category and plotted against the amino acids in increasing disorder promoting potential. The group of all Swiss-Prots represents all protein sequences from Swiss-Prot. Of those, all proteins which have at least one detected TR were subtracted resulting in the group 'all Swissprots w/o TRs'. The group 'only TRs' was calculated by the multiple sequence alignment of the TRs. For the amino acids B, O, U, Z and X was no disorder potential available. One can see that the amino acid ratio of TRs shows a positive linear relationship with increasing disorder propensity. Disorder promoting residues seem to appear more often in TR sequences compared to overall protein sequences and to proteins without TRs.



Figure S13. Overlap regions in proteins with intrinsic disorder and tandem repeats. We distinguish four different overlaps of IDR with TRs: *tail-overlap* where IDR begin within the TR-sequence and finishes after the TR-region. In contrast, we call *head-overlaps* overlap regions when the IDR begins before the TR-sequence and finishes within. If the IDR lies within a TR sequence, we call it. *Disorder-in-TR* and *TR-in-Disorder-overlap* if the TR-region lies within the IDR.



Figure S14. The ten PFAM with the most detected TRs for each superkingdom are plotted according their normalized TR center location (see Methods) and number of site-specific TRs. As expected, it can be seen that the protein family position is consistent overall superkingdoms.

ArchaeaTFIIBTranscription factor TFIIB repeatPF0038235CBSCBS domainPF0057122Fer44Fe-45 binding domainPF1003716Fer4.74Fe-45 dicluster domainPF1283813LAGLIDADG_3LAGLIDADG-like domainPF1452811HexapepBacterial transferase hexapeptide (six repeats)PF001329TF_Zn_RibbonTFIIB zinc-bindingPF002719Ribosomal_L6Ribosomal protein L6PF003477Rad50_zin_hookRad50 zinc hook motifPF042237Fer4_104Fe-4S dicluster domainPF132377BacterialHexapepBacterial transferase hexapeptide (six repeats)PF00412928MraZMraZMraZ protein, putative antitoxin-likePF00381320Ribosomal_L6Ribosomal protein L6PF00347317NTP_transf_3MobA-like NTP transferase domainPF14804244Hexapep_2Hexapetide repeat of succinyl-transferase:PF14602223PD40WD40-like Bta Propeller RepeatPF07676164Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase;PF13720158DNA_gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF04052115DNA_gyraseA_CDNA gyrase C-terminal domainPF04052115DNA_gyraseA_CDNA gyrase C-terminal domainPF13855587EF-hand JTertarticopeptide repeatPF004004449zfC2H2Zinc finger, C2	PFAM Name	PFAM Desc	PFAM Acc	count
TFIIB Transcription factor TFIIB repeat PF00382 35 CBS CBS domain PF00571 22 Fer4 4Fe-4S binding domain PF00037 16 Fer4_7 Afe-4S binding domain PF10037 16 LAGLIDADG_3 LAGLIDAOG-like domain PF12838 13 Hexapep Bacterial transferase hexapeptide (six repeats) PF00132 9 TZ_n_Ribbon TFIIB zinc-binding PF00237 7 Rad50_zn_hook Rad50 zinc hook motif PF04423 7 Fer4_10 4Fe-4S dicluster domain PF13237 7 Bacteria Maz protein, putative antitoxin-like PF00347 317 Hexapep Bacterial transferase hexapeptide (six repeats) PF00132 928 Mra2 Mra2 protein, putative antitoxin-like PF02381 320 Ribosomal_L6 Ribosomal protein L6 PF00347 317 Maz Mra2 protein, putative antitoxin-like PF02381 320 Ribosomal_L6 Ribosomal protein L6 PF00347 317	Archaea			
CBS CBS domain PF00571 22 Fer4 4Fe-4S binding domain PF10037 16 Fer4_7 4Fe-4S dicluster domain PF10528 13 LAGLIDADG_3 LAGLIDADG-like domain PF14528 11 Hexapep Bacterial transferase hexapeptide (six repeats) PF00132 9 TF_Zn_Ribbon TFIIB zinc-binding PF08271 9 Rat50_zin_hook Rad50 zinc hook motif PF04423 7 Fer4_10 4Fe-4S dicluster domain PF13237 7 Bacteria transferase hexapeptide (six repeats) PF00132 928 MraZ MraZ protein, putative antitoxin-like PF02381 320 Ribosomal_L6 Ribosomal protein L6 PF00347 317 NTP_transf_3 MobA-like NTP transferase domain PF12804 244 Hexapep_2 Hexapeptide repeat of succinyl-transferase; PF14520 228 PD40 WD40-like Beta Propeller Repeat PF07676 164 Acetyltransf_11 Udp N-acetylglucosamine 0-acyltransferase; PF13720 158 <td>TFIIB</td> <td>Transcription factor TFIIB repeat</td> <td>PF00382</td> <td>35</td>	TFIIB	Transcription factor TFIIB repeat	PF00382	35
Fer44Fe-4S binding domainPF1003716Fer4_74Fe-4S dicluster domainPF1283813LAGLIDADG_3LAGLIDADC-like domainPF1452811HexapepBacterial transferase hexapeptide (six repeats)PF001329Fr_Zn_RibbonTFIIB zinc-bindingPF082719Ribosomal_L6Ribosomal protein L6PF003477Rad50_zn_hookRad50 zinc hook motifPF044237Fer4_104Fe-4S dicluster domainPF132377Bacteria9HexapepBacterial transferase hexapeptide (six repeats)PF00132928MraZMraZ protein, putative antitoxin-likePF02381320Ribosomal_L6Ribosomal protein L6PF00347317TTP_transf_3MobA-like NFP transferase domainPF12804244Hexapep_2Hexapeptide repeat of succinyl-transferase:PF14602223PD40WD40-like Beta Propeller RepeatPF07676164Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase;PF13720158DNA gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF004001449zt-C2H2Zinc finger, C2H2 typePF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_11RNA recognition motif, (a.k.a. RRM, RBD, or RNP do main) PF00056113PPR_2	CBS	CBS domain	PF00571	22
Fer4_74Fe-4S dicluster domainPF1283813LAGLIDADG_3LAGLIDADG-like domainPF1452811HexapepBacterial transferase hexapeptide (six repeats)PF001329TF_Zn, RibbonTFIIB zinc-bindingPF082719Ribosomal_L6Ribosomal protein L6PF003477Rad50_zn_hookRad50 zinc hook motifPF013237Fer4_104Fe-4S dicluster domainPF128377BacterialTransferase hexapeptide (six repeats)PF00132928MraZMraZ protein, putative antitoxin-likePF00381320MraZMraZ protein, putative antitoxin-likePF00347317NTP_transf_3MobA-like NTP transferase domainPF12804244Hexapep_2Hexapetpide repeat of succinyl-transferasePF14602223PD40WD40-like Beta Propeller RepeatPF07676164Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase; Domain 2PF13720158DMagyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF03986100WD40WD domain, G-beta repeatPF13499520RRM_1RNA recognition motif, a.k.a. RRM, RBD, or RNP do main) PF00056134LIMLIMLiM domainPF00412260PPRPPR repeat familyPF13355587EF-hand_7EF-hand domain pairPF00351184Collagen triple helix repeat (20 copies)PF10331<	Fer4	4Fe-4S binding domain	PF00037	16
LAGLIDADG_3 LAGLIDADG-like domain PF14528 11 Hexapep Bacterial transferase hexapeptide (six repeats) PF00132 9 Ribosomal_L6 Ribosomal protein L6 PF00347 7 Rad50_zn_hook Rad50 zinc hook motif PF00423 7 Fer4_10 4Fe-45 dicluster domain PF13237 7 Bacteria Hexapep Bacterial transferase hexapeptide (six repeats) PF00132 928 Maz MraZ MraZ protein, putative antitoxin-like PF02381 320 Ribosomal_L6 Ribosomal protein L6 PF00132 928 Maz MraZ MraZ protein, putative antitoxin-like PF02381 320 Ribosomal_L6 Ribosomal protein L6 PF00347 317 NTP_transf_3 MobA-like NTP transferase domain PF14602 223 PD40 WD40-like Beta Propeller Repeat PF07676 164 Acetyltransf_11 Udp N-acetylglucosamine O-acyltransferase; PF13720 158 DNA_gyraseA_C DNA gyrase C-terminal domain, beta-propeller PF04052 115 DNA_gyraseA_C DNA gyrase C-terminal domain, beta-propeller	Fer4 7	4Fe-4S dicluster domain	PF12838	13
HexapepBacterial transferase hexapeptide (six repeats)PF001329TF_Zn_RibbonTFIIB zinc-bindingPF082719Ribosomal_L6Ribosomal protein L6PF0083777Rad50_zn_hookRad50 zinc hook motifPF044237Fer4_104Fe-45 dicluster domainPF0132928MraZMraZMraZ protein, putative antitoxin-likePF00331320MraZMraZ protein, putative antitoxin-likePF00347317NTP_transf_3MobA-like NTP transferase domainPF12804244Hexapep_2Hexapeptide repeat of succinyl-transferasePF14602223PD40WD40-like Beta Propeller RepeatPF07676164Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase; Domain 2PF04613127TolB_NTolB amino-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF004001449zf-C2H2Zinc finger, C2H2 typePF004001449zf-C2H2Zinc finger, C2H2 typePF03855587EF-hand domain pairPF13419250226PPRPPR repeatPF01535226PPR_1Tetraticopeptide repeatPF01535226PPR_2PPR repeatPF01535226PPR_2PPR repeatPF01535226PPR_1Tetraticopeptide repeatPF00515184CollagenCollagen triple helix repeatPF00551184CollagenCollagen triple helix rep	LAGLIDADG 3	LAGLIDADG-like domain	PF14528	11
TFZn_RibbonTFIIB zinc-bindingPF082719Ribosomal_L6Ribosomal protein L6PF003477Rad50_zn_hookRad50 zinc hook motifPF004237Fer4_104Fe-4S dicluster domainPF132377BacteriaHexapepBacterial transferase hexapeptide (six repeats)PF00132928MraZMraZ protein, putative antitoxin-likePF02381320Ribosomal_L6Ribosomal protein L6PF00347317NTP_transf_3MobA-like NTP transferase domainPF12804244Hexapep_2Hexapeptide repeat of succinyl-transferasePF07676164Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase;PF13720158DVDN-acyltransferase, LpxDPF04613127TolB_NTolB amino-terminal domainPF04052115DNA gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD domain, G-beta repeatPF004001449zt-C2H2Zinc finger, C2H2 typePF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif (a.k.a. RRM, RBD, or RNP do main) PF00351181VirusesZi-CCHCZinc knucklePF00398Zt-CCHCZinc knucklePF00351184CollagenCollagen triple helix repeatPF00515184CollagenCollagen trotein (nob domain)PF0054037RV	Hexapep	Bacterial transferase hexapeptide (six repeats)	PF00132	9
Ribosomal_L6Ribosomal protein L6PF003477Rad50_zn_hookRad50 zinc hook motifPF044237Fer4_104Fe-4S dicluster domainPF132377BacteriaHexapepBacteria transferase hexapeptide (six repeats)PF00132928MraZMraZMraZ protein, putative antitoxin-likePF02381320Ribosomal_L6Ribosomal protein L6PF00347317NTP_transf_3MobA-like NTP transferase domainPF12804244Hexapep_2Hexapeptide repeat of succinyl-transferasePF14602223PD40WD40-like Beta Propeller RepeatPF07676164Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase; Domain 2PF13720158UDP-3-O-[3-hydroxymyristoyl] glucosamineLpxDN-acyltransferase, LpxDPF04613127TolB_NTolB amino-terminal domainPF04052115116DNA gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF004001449zf-C2H2Zinc finger, C2H2 typePF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13855587PFR_1Tertarticopeptide repeatPF00412260PPR_2PPR repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391PR1Tertarticopeptide repeatPF0054037RVPRetroviral aspartyl proteasePF00023111Adeno_knobAdenoviral fib	TF Zn Ribbon	TFIIB zinc-binding	PF08271	9
Rad50_zn_hookRad50 zinc hook motifPF044237Fer4_104Fe-4S dicluster domainPF132377BacteriaFerd 104Fe-4S dicluster domainPF132377BacteriaMraZMraZ protein, putative antitoxin-likePF03132928MraZMraZ maz protein, putative antitoxin-likePF03381320Ribosomal L6Ribosomal protein L6PF00347317NTP_transf_3MobA-like NTP transferase domainPF12804244Hexapep_2Hexapeptide repeat of succinyl-transferasePF14602223PD40WD40-like Beta Propeller RepeatPF07676164Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase; Domain 2PF13720158UDP-3O-[3-hydroxymyrisoyl] glucosamineUDP-3O-[3-hydroxymyrisoyl] glucosaminePF04052115DNA_gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF00515184CollagenCollagen triple helix repeatPF00535226PPR_12PPR repeatPF13041225TPR_1Tetratroopetide repeatPF00515184CollagenCollagen triple helix repeatPF0005437RVPRetroviral a	Ribosomal L6	Ribosomal protein L6	PF00347	7
Fer4_104Fe-4S dicluster domainPF132377BacteriaHexapepBacterial transferase hexapeptide (six repeats)PF00132928MraZMraZ protein, putative antitoxin-likePF02381320Ribosomal_L6Ribosomal protein L6PF00347317NTP_transf_3MobA-like NTP transferase domainPF12804244Hexapep_2Hexapeptide repeat of succinyl-transferasePF14602223PD40WD40-like Beta Propeller RepeatPF07676164Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase; Domain 2PF13720158UDP-3-O-[3-hydroxymyristoyl] glucosamineN-acyltransferase, LpxDPF04613127TolB_NTolB amino-terminal domainPF04052115DNA gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF000096828LRR 8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIMLiM domainPF0355226PPR_2PPR repeatPF01535226PRR_2PPR repeat familyPF13041225TFR_1Tetratricopeptide repeatPF00515184Collagen triple helix repeat (20 copies)PF01391181VirusesZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein) <t< td=""><td>Rad50 zn hook</td><td>Rad50 zinc hook motif</td><td>PF04423</td><td>7</td></t<>	Rad50 zn hook	Rad50 zinc hook motif	PF04423	7
Bacteria Hexapep Bacterial transferase hexapeptide (six repeats) PF00132 928 MraZ MraZ protein, putative antitoxin-like PF02381 320 Ribosomal_L6 Ribosomal protein L6 PF00347 317 NTP_transf_3 MobA-like NTP transferase domain PF12804 244 Hexapep_2 Hexapeptide repeat of succinyl-transferase PF14602 223 PD40 WD40-like Beta Propeller Repeat PF07676 164 Acetyltransf_11 Udp N-acetylglucosamine O-acyltransferase; Domain 2 PF13720 158 UDP-3-O-[3-hydroxymyristoyl] glucosamine LpxD N-acyltransferase, LpxD PF04613 127 ToIB_N ToIB amino-terminal domain PF04052 115 115 DNA gyraseA_C DNA gyrase C-terminal domain, beta-propeller PF04052 115 DNA gyraseA_C DN domain, G-beta repeat PF04052 125 DNA gyraseA_C DN domain pair PF13855 587 EF-hand_7 EF-hand domain pair PF13499 520 RRM_1 RNA recognition motif. (a.k.a. RRM, RBD, or RNP do	Fer4 10	4Fe-4S dicluster domain	PF13237	7
Hexapep HexapepBacterial transferase hexapeptide (six repeats)PF00132928MraZMraZ protein, putative antitoxin-likePF02381320Ribosomal_L6Ribosomal protein L6PF00347317NTP_transf_3MobA-like NTP transferase domainPF12804244Hexapep_2Hexapeptide repeat of succinyl-transferasePF14602223PD40WD40-like Beta Propeller RepeatPF07676164Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase; Domain 2PF13720158UDP-3-O-[3-hydroxymyristoyl] glucosamineUDP-3-O-[3-hydroxymyristoyl] glucosamine127ToIB_NToIB amino-terminal domainPF04613127ToIB_NToIB amino-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF04052115DNA_gyraseA_CDNA gyrase C-terminal domain pairPF13855587Er-hand_7EF-hand domain pairPF13855520RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF00412260PPR_2PPR repeatPF01355226PPR_2PPR repeat familyPF13911225TFR_1Tetratricopeptide repeatPF00515184Collagen triple helix repeat (20 copies)PF01391181VirusesCollagen triple helix repeat (20 copies)PF01391181AnkAnkyrin repeatPF00555184Collagen triple helix repeat (20 copies)	Bacteria			
MraZMraZ protein, putative antitoxin-likePF02381320Ribosomal_L6Ribosomal protein L6PF00347317NTP_transf_3MobA-like NTP transferase domainPF12804244Hexapep_2Hexapeptide repeat of succinyl-transferasePF14602223PD40WD40-like Beta Propeller RepeatPF07676164Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase; Domain 2PF13720158UDP-3-O-[3-hydroxymyristoyl] glucosamineLpxDN-acyltransferase, LpxDPF04613127TolB_NTolB amino-terminal domainPF04052115DNA_gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF004001449zf-C2H2Zinc finger, C2H2 typePF00906828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif (a.k.a. RRM, RBD, or RNP do main) PF000176413LIMLIM domainPF00155226PPR_2PPR repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181VirusesZirc knucklePF00023111Adeno_knobAdenoviral fibre protein (knob domain)PF00023111Adeno_knobAdenoviral fibre protein (knob domain)PF00023111Adeno_knobAdenoviral fibre protein (knob domain)PF000665111BTBBTB/	Hexapep	Bacterial transferase hexapeptide (six repeats)	PF00132	928
Ribosomal_L6Ribosomal protein L6PF00347317NTP_transf_3MobA-like NTP transferase domainPF12804244Hexapep_2Hexapeptide repeat of succinyl-transferasePF14602223PD40WD40-like Beta Propeller RepeatPF07676164Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase; Domain 2PF13720158UDP-3-O-[3-hydroxymyristoyl] glucosamineUDP-3-O-[3-hydroxymyristoyl] glucosamine127ToIB_NToIB amino-terminal domainPF04052115DNA_gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF040001449zf-C2H2Zinc finger, C2H2 typePF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF00153226PPR_2PPR repeatPF0135226PPR_2PPR repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181VirusesZf-CCHCZinc knucklePF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_knobAdenoviral fibre protein (knob domain)PF0066511AnkAnkyrin repeatPF000651111Adeno_knobAdenoviral fibre protein (knob domain) <t< td=""><td>MraZ</td><td>MraZ protein, putative antitoxin-like</td><td>PF02381</td><td>320</td></t<>	MraZ	MraZ protein, putative antitoxin-like	PF02381	320
NTP_transf_3MobA-like NTP transferase domainPF12804244Hexapep_2Hexapeptide repeat of succinyl-transferasePF14602223PD40WD40-like Beta Propeller RepeatPF07676164Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase; Domain 2PF13720158UDP-3-O-[3-hydroxymyristoyl] glucosamineFF13720158LpxDN-acyltransferase, LpxDPF04613127TolB_NTolB amino-terminal domainPF04052115DNA_gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF00076413LIMLIM domainPF0012260PPR_2PPR repeatPF01535226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181VirusesZf-CCHCZinc knucklePF0007713AnkAnkyrin repeatPF0007713AnkAdenoviral fibre protein (knob domain)PF00550111Adeno_knobAdenoviral fibre protein (knob domain)PF00668111revIntegrase core domainPF0066511RNase_HBTB/POZ domainPF000759SushiSushiSushi repeat (SCR repeat)PF0007	Ribosomal L6	Ribosomal protein L6	PF00347	317
Hexapep_2 PD40Hexapeptide repeat of succinyl-transferase WD40-like Beta Propeller RepeatPF14602 PE13720223 PD36Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase; Domain 2PF13720158 UDP-3-O-[3-hydroxymyristoyl] glucosamineLpxDN-acyltransferase, LpxDPF04613127TolB_NTolB amino-terminal domainPF04052115DNA gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF0389100EukaryotaWD40WD domain, G-beta repeatPF004001449zf-C2H2Zinc finger, C2H2 typePF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif, (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF045135226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF01535226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF00515184Collagen triple helix repeat (20 copies)PF01391181VirusesZf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037AnkAnkyrin repeatPF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0056111AnkAnkyrin repeatPF0066511RNABTB/POZ domainPF00651 </td <td>NTP transf 3</td> <td>MobA-like NTP transferase domain</td> <td>PF12804</td> <td>244</td>	NTP transf 3	MobA-like NTP transferase domain	PF12804	244
PD40WD40-like Beta Propeller RepeatPF07676164Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase; Domain 2PF13720158LpxDN-acetylglucosamine O-acyltransferase; DUP-3-O-[3-hydroxymyristoyl] glucosaminePF04613127TolB_NTolB amino-terminal domainPF04052115DNA_gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF004001449zf-C2H2Zinc finger, C2H2 typePF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181VirusesZf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAnkyrin repeatPF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0006511Adeno_shaftAdenoviral fibre protein (knob domain)PF0066511BTBBTB/POZ domainPF0005110RNase_HPF000759Suishi geneat (SCR reneat)PF000759SuishiSuishi reneatScreenationPF0005110	Hexapep 2	Hexapeptide repeat of succinvl-transferase	PF14602	223
Acetyltransf_11Udp N-acetylglucosamine O-acyltransferase; Domain 2PF13720158LpxDUDP-3-O-[3-hydroxymyristoyl] glucosamineUDP-3-O-[3-hydroxymyristoyl] glucosamineLpxDN-acyltransferase, LpxDPF04613127TolB_NTolB amino-terminal domainPF04052115DNA_gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF004001449xf-C2H2Zinc finger, C2H2 typePF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13855587EF-hand_7EF-hand domain pairPF00412260PPRPPR repeatPF00412260PPRPPR repeatPF01535226PPR_2PPR repeatPF051535226PPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181VirusesZi-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAdenoviral fibre protein (knob domain)PF0066511Adeno_knobAdenoviral fibre protein (repeat/shaft region)PF0066511BTBBTB/POZ domainPF0006511RNase_HPF000759Suishi repeatPF000759SuishiSuishi repeat <td>PD40</td> <td>WD40-like Beta Propeller Repeat</td> <td>PF07676</td> <td>164</td>	PD40	WD40-like Beta Propeller Repeat	PF07676	164
Acetyltranst_11Dep H 1000 (g) and 2 Domain 2PF 13/20158UDP-3-O-[3-hydroxymyristoyl] glucosamineLpxDN-acyltransferase, LpxDPF04613127TolB_NTolB amino-terminal domainPF04052115DNA gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF004001449zf-C2H2Zinc finger, C2H2 typePF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF01535226PPRPPR repeatPF01535226PPR_2PPR repeatPF01535226PPR_1Tetratricopeptide repeatPF01535226PPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181Viruseszf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAnkyrin repeatPF00023111Adeno_knobAdenoviral fibre protein (repeat/shaft region)PF0066511BTBBTB/POZ domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR reneat)PF000759		Udp N-acetylglucosamine Q-acyltransferase:		
UDP-3-O-[3-hydroxymyristoyl] glucosamineLpxDN-acyltransferase, LpxDPF04613127TolB_NTolB amino-terminal domainPF04052115DNA_gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF004001449zf-C2H2Zinc finger, C2H2 typePF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainLM domainPF01535226PPRPPR repeatPF01535226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeat (20 copies)PF01391181VirusesZf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAdenoviral fibre protein (knob domain)PF00654111Adeno_knobAdenoviral fibre protein (knob domain)PF0065511BTBBTB/POZ domainPF0065511BTBBTB/POZ domainPF000759SushiSushi repeat (SCR repeat)PF000759	Acetyltransf_11	Domain 2	PF13720	158
LpxDN-acyltransferase, LpxDPF04613127TolB_NTolB amino-terminal domainPF04052115DNA_gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF004001449xdf-C2H2Zinc finger, C2H2 typePF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF00412260PPRPPR repeatPF01535226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181VirusesZr-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (knob domain)PF0066511BTBBTB/POZ domainPF0065110RNase_HRNase HPF000759SushiSushi proteat (SCR repeat)PF000849		UDP-3-O-[3-hydroxymyristoy]] glucosamine		
TolB_NTolB amino-terminal domainPF04052115DNA_gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF04052115DNA_gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF004001449zf-C2H2Zinc finger, C2H2 typePF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF00412260PPRPPR repeatPF01535226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181Viruseszf-CCHCZinc knucklePF0007713AnkAnkyrin repeatPF0007713AnkAdenoviral fibre protein (knob domain)PF0054111Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0060811rveIntegrase core domainPF0065511BTBBTB/POZ domainPF0065511RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	L pxD	N-acyltransferase I pxD	PF04613	127
DNA_gyraseA_CDNA gyrase C-terminal domain, beta-propellerPF03989100EukaryotaWD40WD domain, G-beta repeatPF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF001535226PPRPPR repeatPF01535226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181Viruseszf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0066511BTBBTB/POZ domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000249	TolB N	ToIB amino-terminal domain	PF04052	115
EukaryotaWD40WD domain, G-beta repeatPF004001449zf-C2H2Zinc finger, C2H2 typePF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF00412260PPRPPR repeatPF001535226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181VirusesZf-CCHCZinc knucklePF0007713AnkAnkyrin repeatPF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0066511BTBBTB/POZ domainPF00065110RNase_HPF000759SushiSushi repeatSUshi repeatPF0007599	DNA gyraseA C	DNA gyrase C-terminal domain, beta-propeller	PF03989	100
WD40WD domain, G-beta repeatPF004001449zf-C2H2Zinc finger, C2H2 typePF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF00412260PPRPPR repeatPF01535226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181Viruseszf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0066511BTBBTB/POZ domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	Fukarvota			
zf-C2H2Zinc finger, C2H2 typePF00096828LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13499520RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF00412260PPRPPR repeatPF01535226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181Viruseszf-CCHCZinc knucklePF009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0066511BTBBTB/POZ domainPF0065110RNase_HRNase HPF00759SushiSushi repeat (SCR repeat)PE000849	WD40	WD domain G-beta repeat	PF00400	1449
LRR_8Leucine rich repeatPF13855587EF-hand_7EF-hand domain pairPF13855587RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF00412260PPRPPR repeatPF01535226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181VirusesZf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAnkyrin repeatPF002311Adeno_knobAdenoviral fibre protein (repeat/shaft region)PF0066811rveIntegrase core domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	7f-C2H2	Zinc finger C2H2 type	PF00096	828
EF-hand_7EF-hand domain pairPF13049520RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF00412260PPRPPR repeatPF01535226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181Viruseszf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAnkyrin repeatPF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (knob domain)PF0060811rveIntegrase core domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849		Leucine rich reneat	PE13855	587
RRM_1RNA recognition motif. (a.k.a. RRM, RBD, or RNP do main) PF00076413LIMLIM domainPF00412260PPRPPR repeatPF01535226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181Viruseszf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAnkyrin repeatPF0002311Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0066511rveIntegrase core domainPF0065110RNase_HRNase HPF0007599SushiSushi repeat (SCR repeat)PF000849	FE-hand 7	EF-hand domain pair	PF13499	520
LIMLIM domainPF00412260PPRPPR repeatPF01535226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181Viruseszf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAnkyrin repeatPF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0066511BTBBTB/POZ domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	RRM 1	RNA recognition motif (a k a RRM RBD or RNP do main)	PF00076	413
PPRPPR repeatPF01535226PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181Viruseszf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAnkyrin repeatPF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0060811rveIntegrase core domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849		I IM domain	PF00412	260
PPR_2PPR repeat familyPF13041225TPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181Viruseszf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAnkyrin repeatPF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0060811rveIntegrase core domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	PPR	PPR reneat	PE01535	226
TPR_1Tetratricopeptide repeatPF00515184CollagenCollagen triple helix repeat (20 copies)PF01391181Viruseszf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAnkyrin repeatPF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0060811rveIntegrase core domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	PPR 2	PPR repeat family	PF13041	225
CollagenCollagen triple helix repeat (20 copies)PF01391181Viruseszf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAnkyrin repeatPF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0060811rveIntegrase core domainPF0065511BTBBTB/POZ domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	TPR 1	Tetratricopentide repeat	PE00515	184
VirusesZf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAnkyrin repeatPF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0060811rveIntegrase core domainPF0066511BTBBTB/POZ domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	Collagen	Collagen triple helix repeat (20 copies)	PF01391	181
zf-CCHCZinc knucklePF0009856Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAnkyrin repeatPF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0060811rveIntegrase core domainPF0066511BTBBTB/POZ domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	Viruses		1101001	101
Gag_p17gag gene protein p17 (matrix protein)PF0054037RVPRetroviral aspartyl proteasePF0007713AnkAnkyrin repeatPF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0060811rveIntegrase core domainPF0066511BTBBTB/POZ domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	zf-CCHC	Zinc knuckle	PE00098	56
RVPRetroviral aspartyl proteasePF0007713AnkAnkyrin repeatPF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0060811rveIntegrase core domainPF0065511BTBBTB/POZ domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	Gag n17	gag gene protein p17 (matrix protein)	PE00540	37
AnkAnkyrin repeatPF0002311Adeno_knobAdenoviral fibre protein (knob domain)PF0054111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0060811rveIntegrase core domainPF0066511BTBBTB/POZ domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	RVP	Retroviral aspartyl protease	PF00077	13
AdenoAdenoviral fibre protein (knob domain)PF0052111Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0060811rveIntegrase core domainPF0066511BTBBTB/POZ domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	Ank	Ankvrin reneat	PE00023	11
Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0060811Adeno_shaftAdenoviral fibre protein (repeat/shaft region)PF0060511rveIntegrase core domainPF0065511BTBBTB/POZ domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	Adeno knob	Adenoviral fibre protein (knob domain)	PE00541	11
riskri	Adeno_shaft	Adenoviral fibre protein (repeat/shaft region)	PE00608	11
BTBBTB/POZ domainPF0065110RNase_HRNase HPF000759SushiSushi repeat (SCR repeat)PF000849	rve	Integrase core domain	PE00665	11
RNase_H RNase H PF00075 9 Sushi Sushi repeat (SCR repeat) PF00084 9	BTR	BTB/POZ domain	PF00651	10
Sushi Sushi repeat (SCR repeat) PF00084 9	RNase H	RNase H	PE00075	9
	Sushi	Sushi repeat (SCR repeat)	PE00084	g

Table S1. For each superkingdom are the ten most frequent PFAMs listed toghether with their PFAM Description and Accession number. 'Count' represents the number of appearances of the PFAM model in our data.

References

1. Marcotte, E.; Pellegrini, M.; Yeates, T.; Eisenberg, D. A census of protein repeats. J. Mol.Biol. 1999, 293, 151–160.



© 2020 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).