Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure

DOI: 10.1006/jmbi.2001.5080
Journal of Molecular Biology, 2001, 313, 903-19.

#	Paper	IF	Citations
1041	The Spin/Ssty repeat: a new motif identified in proteins involved in vertebrate development from gamete to embryo. **2002**, 3, RESEARCH0003		20
1040	References. **2001**, 328-342		
1039	Structural characterization of the human proteome. **2002**, 12, 1625-41		57
1038	The COBRA family of putative GPI-anchored proteins in Arabidopsis. A new fellowship in expansion. **2002**, 130, 538-48		107
1037	A comparison of profile hidden Markov model procedures for remote homology detection. **2002**, 30, 4321-8		113
1036	SCOP database in 2002: refinements accommodate structural genomics. **2002**, 30, 264-7		348
1035	Gene3D: structural assignment for whole genes and genomes using the CATH domain structure database. **2002**, 12, 503-14		49
1034	SUPERFAMILY: HMMs representing all proteins of known structure. SCOP sequence searches, alignments and genome assignments. **2002**, 30, 268-72		216
1033	Comparison of the small molecule metabolic enzymes of Escherichia coli and Saccharomyces cerevisiae. **2002**, 12, 916-29		20
1032	The use of structure information to increase alignment accuracy does not aid homologue detection with profile HMMs. **2002**, 18, 1243-9		19
1031	The repertoire of protein kinases encoded in the draft version of the human genome: atypical variations and uncommon domain combinations. **2002**, 3, RESEARCH0066		36
1030	The dominance of the population by a selected few: power-law behaviour applies to a wide variety of genomic properties. **2002**, 3, RESEARCH0040		71
1029	Protein domain analysis in the era of complete genomes. **2002**, 513, 129-34		46
1028	Proteins with class alpha/beta fold have high-level participation in fusion events. *Journal of Molecular Biology*, **2002**, 320, 713-9	6.5	8
1027	Sequence conservation in families whose members have little or no sequence similarity: the four-helical cytokines and cytochromes. *Journal of Molecular Biology*, **2002**, 322, 205-33	6.5	39
1026	The constraints protein-protein interactions place on sequence divergence. *Journal of Molecular Biology*, **2002**, 324, 399-407	6.5	92
1025	Taking the complexity out of protein sequence analysis. **2002**, 7, 172-5		0

1024	Structural genomics of proteins from conserved biochemical pathways and processes. **2002**, 12, 383-91	34
1023	Automated detection of remote homology. **2002**, 12, 362-7	26
1022	Did evolution leap to create the protein universe?. **2002**, 12, 409-16	41
1021	OB-fold domains: a snapshot of the evolution of sequence, structure and function. **2002**, 12, 794-801	168
1020	A regulatory cytoplasmic poly(A) polymerase in Caenorhabditis elegans. **2002**, 419, 312-6	241
1019	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. **2002**, 420, 563-73	1350
1018	Multi-domain protein families and domain pairs: comparison with known structures and a random model of domain recombination. **2003**, 4, 67-78	74
1017	A structural perspective on genome evolution. **2003**, 13, 359-69	19
1016	The STIR-domain superfamily in signal transduction, development and immunity. **2003**, 28, 226-9	191
1015	Visualisation and graph-theoretic analysis of a large-scale protein structural interactome. **2003**, 4, 45	18
1014	PCAS--a precomputed proteome annotation database resource. **2003**, 4, 42	6
1013	Specialized hidden markov model databases for microbial genomics. **2003**, 4, 250-4	3
1012	Prediction of protein structural classes by a new measure of information discrepancy. **2003**, 27, 373-80	34
1011	ParG, a protein required for active partition of bacterial plasmids, has a dimeric ribbon-helix-helix structure. **2003**, 50, 1141-53	70
1010	ATP N-glycosidase - a novel ATP-converting activity from a marine sponge Axinella polypoides. **2003**, 270, 4122-32	11
1009	Hidden Markov models that use predicted local structure for fold recognition: alphabets of backbone geometry. **2003**, 51, 504-14	148
1008	Using evolutionary information for the query and target improves fold recognition. **2004**, 54, 342-50	23
1007	Prediction of protein function from protein sequence and structure. **2003**, 36, 307-40	300

1006	Modeling Protein Structure from its Sequence. **2003**, 3, 5.1.1		4
1005	On the role of structural information in remote homology detection and sequence alignment: new methods using hybrid sequence profiles. *Journal of Molecular Biology*, **2003**, 334, 1043-62	6.5	77
1004	The crystal structure of AF1521 a protein from Archaeoglobus fulgidus with homology to the non-histone domain of macroH2A. *Journal of Molecular Biology*, **2003**, 330, 503-11	6.5	97
1003	Expression patterns of the RGS9-1 anchoring protein R9AP in the chicken and mouse suggest multiple roles in the nervous system. **2003**, 24, 687-95		26
1002	Evolution of the protein repertoire. **2003**, 300, 1701-3		419
1001	META-PP: single interface to crucial prediction servers. **2003**, 31, 3308-10		10
1000	Proteomics in the postgenomic age. **2003**, 65, 1-23		11
999	Efficient estimation of emission probabilities in profile hidden Markov models. **2003**, 19, 2359-68		6
998	Development and evaluation of an automated annotation pipeline and cDNA annotation system. **2003**, 13, 1542-51		24
997	Comprehensive analysis of the mouse metabolome based on the transcriptome. **2003**, 13, 1345-9		23
996	Evolution of transcription factors and the gene regulatory network in Escherichia coli. **2003**, 31, 1234-44		234
995	The structure of full-length LysR-type transcriptional regulators. Modeling of the full-length OxyR transcription factor dimer. **2003**, 31, 1444-54		48
994	Conservation of structure and function among tyrosine recombinases: homology-based modeling of the lambda integrase core-binding domain. **2003**, 31, 805-18		23
993	The esterase and PHD domains in CR1-like non-LTR retrotransposons. **2003**, 20, 38-46		65
992	Vascular endothelial growth factor receptor-1 is deposited in the extracellular matrix by endothelial cells and is a ligand for the alpha 5 beta 1 integrin. **2003**, 116, 3479-89		86
991	Making decisions for structural genomics. **2003**, 4, 150-67		12
990	The immunoglobulin superfamily in Drosophila melanogaster and Caenorhabditis elegans and the evolution of complexity. **2003**, 130, 6317-28		89
989	Functional map and domain structure of MET, the product of the c-met protooncogene and receptor for hepatocyte growth factor/scatter factor. **2003**, 100, 12039-44		149

988	The mouse secretome: functional classification of the proteins secreted into the extracellular environment. **2003**, 13, 1350-9	65
987	Mouse Genome Encyclopedia Project. **2003**, 68, 195-204	6
986	A multi-agent system for supporting the prediction of protein structures. **2004**, 11, 259-280	4
985	. **2004**,	2
984	Using property based sequence motifs and 3D modeling to determine structure and functional regions of proteins. **2004**, 11, 583-93	43
983	transition priors for protein hidden Markov models: an empirical study towards maximum discrimination. **2004**, 11, 181-93	3
982	C-terminal nsP1a protein of human astrovirus colocalizes with the endoplasmic reticulum and viral RNA. **2004**, 78, 13627-36	36
981	Sequence-based prediction of protein domains. **2004**, 32, 3522-30	62
980	Characterization of the UDP-N-acetylgalactosamine binding domain of bovine polypeptide alphaN-acetylgalactosaminyltransferase T1. **2004**, 17, 635-46	4
979	Evolution of sialic acid-binding proteins: molecular cloning and expression of fish siglec-4. **2004**, 14, 959-68	39
978	Tutorial section: domains and motifs - proteins in bite-sized chunks. **2004**, 5, 71-4	
977	3D-GENOMICS: a database to compare structural and functional annotations of proteins between sequenced genomes. **2004**, 32, D245-50	13
976	Comparative structural modeling and inference of conserved protein classes in Drosophila seminal fluid. **2004**, 101, 13542-7	108
975	Prokaryotic diversity of the Saccharomyces cerevisiae Atx1p-mediated copper pathway. **2004**, 20, 2644-55	11
974	The immunoglobulin superfamily in Caenorhabditis elegans and Drosophila melanogaster. **2004**, 131, 2237-8; author reply 2238-40	5
973	Looking at the bigger picture. **2004**, 131, 2238-2240	
972	SPrCY: comparison of structural predictions in the Saccharomyces cerevisiae genome. **2004**, 20, 2312-4	6
971	Conservation of adjacency as evidence of paralogous operons. **2004**, 32, 5392-7	33

970	The CATH Domain Structure Database and related resources Gene3D and DHS provide comprehensive domain family information for genome analysis. **2005**, 33, D247-51		204
969	The UCSC Proteome Browser. **2005**, 33, D454-8		36
968	A complementary bioinformatics approach to identify potential plant cell wall glycosyltransferase-encoding genes. **2004**, 136, 2609-20		61
967	Essential role of the NH2-terminal WD/EPF motif in the phosphorylation-activated protective function of mammalian Hsp27. **2004**, 279, 23463-71		89
966	HIV-1 Tat directly binds to NFkappaB enhancer sequence: role in viral and cellular gene expression. **2004**, 32, 1270-8		36
965	Tissue-specific modification of gld-2 mRNA in C. elegans: likely C-to-U editing. **2004**, 10, 1444-8		7
964	CooPPS: a system for the cooperative prediction of protein structures. **2004**, 2, 471-95		4
963	The Universal Protein Resource (UniProt). **2005**, 33, D154-9		1231
962	Delineation of modular proteins: domain boundary prediction from sequence information. **2004**, 5, 179-92		24
961	Sensing and responding to diverse extracellular signals? Analysis of the sensor kinases and response regulators of Streptomyces coelicolor A3(2). **2004**, 150, 2795-2806		125
960	Components of coated vesicles and nuclear pore complexes share a common molecular architecture. **2004**, 2, e380		318
959	Novel cell wall architecture of isoxaben-habituated Arabidopsis suspension-cultured cells: global transcript profiling and cellular analysis. *Plant Journal*, **2004**, 40, 260-75	6.9	116
958	Structural analysis of the chromosome segregation protein Spo0J from Thermus thermophilus. **2004**, 53, 419-32		84
957	Identification of a gene required for the formation of lyso-ornithine lipid, an intermediate in the biosynthesis of ornithine-containing lipids. **2004**, 53, 1757-70		68
956	Gene regulatory network growth by duplication. **2004**, 36, 492-6		406
955	NERD: a DNA processing-related domain present in the anthrax virulence plasmid, pXO1. **2004**, 29, 106-10		10
954	Dynamic transcriptome of mice. **2004**, 22, 161-7		15
953	Structure, function and evolution of multidomain proteins. **2004**, 14, 208-16		299

952	The interplay of fold recognition and experimental structure determination in structural genomics. **2004**, 14, 307-12	38
951	Structure and evolution of transcriptional regulatory networks. **2004**, 14, 283-91	556
950	4SCOPmap: automated assignment of protein structures to evolutionary superfamilies. **2004**, 5, 197	30
949	PASS2: an automated database of protein alignments organised as structural superfamilies. **2004**, 5, 35	32
948	Structural characterization of genomes by large scale sequence-structure threading. **2004**, 5, 37	5
947	Genome-wide analysis of the globin gene family of C. elegans. **2004**, 56, 697-702	25
946	Analysis of the S-locus structure in Prunus armeniaca L. Identification of S-haplotype specific S-RNase and F-box genes. **2004**, 56, 145-57	86
945	A structure-based method for identifying DNA-binding proteins and their sites of DNA-interaction. **2004**, 5, 255-265	1
944	The anthracnose resistance locus Co-4 of common bean is located on chromosome 3 and contains putative disease resistance-related genes. **2004**, 109, 690-9	47
943	CHOP proteins into structural domain-like fragments. **2004**, 55, 678-88	53
942	Fold recognition by combining sequence profiles derived from evolution and from depth-dependent structural alignment of fragments. **2005**, 58, 321-8	195
941	Neurotrophic activities of trk receptors conserved over 600 million years of evolution. **2004**, 60, 12-20	28
940	Protein sequence databases. **2004**, 8, 76-80	173
939	Phylogenetic distribution of DNA-binding transcription factors in bacteria and archaea. **2004**, 28, 341-50	75
938	UniProt: the Universal Protein knowledgebase. **2004**, 32, D115-9	2195
937	Identification and characterization of a novel RNA binding protein that associates with the 5'-untranslated region of the chloroplast psbA mRNA. **2004**, 43, 8541-50	31
936	SCOP database in 2004: refinements integrate structure and sequence family data. **2004**, 32, D226-9	733
935	Enzymatic characterization of the streptococcal endopeptidase, IdeS, reveals that it is a cysteine protease with strict specificity for IgG cleavage due to exosite binding. **2004**, 43, 15540-9	111

934	GeneDB: a resource for prokaryotic and eukaryotic organisms. **2004**, 32, D339-43		190
933	Structure and evolution of transcriptional regulatory networks. **2004**, 14, 283-283		1
932	The SUPERFAMILY database in 2004: additions and improvements. **2004**, 32, D235-9		198
931	The zebrafish metaxin 3 gene (mtx3): cDNA and protein structure, and comparison to zebrafish metaxins 1 and 2. **2004**, 330, 67-73		3
930	The active site of the Escherichia coli glycogen synthase is similar to the active site of retaining GT-B glycosyltransferases. **2004**, 316, 960-6		25
929	Mammalian N-acetylglutamate synthase. **2004**, 81 Suppl 1, S4-11		32
928	Improving profile HMM discrimination by adapting transition probabilities. *Journal of Molecular Biology*, **2004**, 338, 847-54	6.5	24
927	Toward consistent assignment of structural domains in proteins. *Journal of Molecular Biology*, **2004**, 339, 647-78	6.5	59
926	Estimating the prevalence of protein sequences adopting functional enzyme folds. *Journal of Molecular Biology*, **2004**, 341, 1295-315	6.5	36
925	Fold recognition for drug discovery. **2004**, 3, 18-23		8
924	Classification of Protein Function. 167-183		
923	Feature extraction for improved profile HMM based biological sequence analysis. **2004**,		3
922	Significance of conservative asparagine residues in the thermal hysteresis activity of carrot antifreeze protein. **2004**, 377, 589-95		28
921	Starch-acting enzymes. **2004**, 128-155		2
920	Congenital thrombotic thrombocytopenic purpura in association with a mutation in the second CUB domain of ADAMTS13. **2004**, 103, 627-9		78
919	Structure comparison and protein structure classifications. **2005**,		
918	Structural similarity to bridge sequence space: finding new families on the bridges. **2005**, 14, 1305-14		11
917	Improved membrane protein topology prediction by domain assignments. **2005**, 14, 1723-8		40

916	Comparative Protein Structure Modeling. **2005**, 831-860	11
915	A novel DNA modification by sulphur. **2005**, 57, 1428-38	92
914	Detection of parallel functional modules by comparative analysis of genome sequences. **2005**, 23, 253-60	27
913	Evolutionary aspects of whole-genome biology. **2005**, 15, 248-53	51
912	Dynamic covariation between gene expression and proteome characteristics. **2005**, 6, 215	9
911	Sequence variation in ligand binding sites in proteins. **2005**, 6, 240	59
910	Prediction of a common structural scaffold for proteasome lid, COP9-signalosome and eIF3 complexes. **2005**, 6, 71	66
909	Improved profile HMM performance by assessment of critical algorithmic features in SAM and HMMER. **2005**, 6, 99	40
908	Evolution of a microbial nitrilase gene family: a comparative and environmental genomics study. **2005**, 5, 42	56
907	Gene fusions and gene duplications: relevance to genomic annotation and functional analysis. **2005**, 6, 33	10
906	COG4849 is a novel family of nucleotidyltransferases. **2005**, 18, 422-5	2
905	The cadherin superfamily in Anopheles gambiae: a comparative study with Drosophila melanogaster. **2005**, 6, 204-16	3
904	Molecular analysis of the CRINKLY4 gene family in Arabidopsis thaliana. **2005**, 220, 645-57	42
903	A structure-based method for identifying DNA-binding proteins and their sites of DNA-interaction. **2004**, 5, 255-65	2
902	The structure at 1.6 Angstroms resolution of the protein product of the At4g34215 gene from Arabidopsis thaliana. **2005**, 61, 1655-61	12
901	The structure at 2.4 A resolution of the protein from gene locus At3g21360, a putative Fe(II)/2-oxoglutarate-dependent enzyme from Arabidopsis thaliana. **2005**, 61, 469-72	4
900	The structure at 2.5 A resolution of human basophilic leukemia-expressed protein BLES03. **2005**, 61, 812-7	1
899	Natural history of S-adenosylmethionine-binding proteins. **2005**, 5, 19	196

898	Protein classification based on text document classification techniques. **2005**, 58, 955-70	68
897	Large-scale protein annotation. **2005**,	
896	The genome sequence of Rickettsia felis identifies the first putative conjugative plasmid in an obligate intracellular parasite. **2005**, 3, e248	224
895	Statistical evaluation and comparison of a pairwise alignment algorithm that a priori assigns the number of gaps rather than employing gap penalties. **2005**, 21, 1421-8	11
894	Functional coverage of the human genome by existing structures, structural genomics targets, and homology models. **2005**, 1, e31	51
893	Constitutive signal transduction by mutant Ssy5p and Ptr3p components of the SPS amino acid sensor system in Saccharomyces cerevisiae. **2005**, 4, 1116-24	12
892	New genes from old: redeployment of dUTPase by herpesviruses. **2005**, 79, 12880-92	58
891	Overview of Structural Bioinformatics. **2005**, 15-44	
890	Meta-DP: domain prediction meta-server. **2005**, 21, 2917-20	43
889	Statistical analysis of domains in interacting protein pairs. **2005**, 21, 993-1001	73
888	The incorporation of glucosamine into enterobacterial core lipopolysaccharide: two enzymatic steps are required. **2005**, 280, 36648-56	13
887	The MicrobesOnline Web site for comparative genomics. **2005**, 15, 1015-22	160
886	The properties of protein family space depend on experimental design. **2005**, 21, 2618-22	10
885	Computational characterization of proteins. **2005**, 2, 129-38	0
884	Evolution in simple systems and the emergence of complexity. **2005**,	1
883	hCLCA1 and mCLCA3 are secreted non-integral membrane proteins and therefore are not ion channels. **2005**, 280, 27205-12	107
882	Cyclophilin A-deficient mice are resistant to immunosuppression by cyclosporine. **2005**, 174, 6030-8	85
881	Predicted hexameric structure of the Agrobacterium VirB4 C terminus suggests VirB4 acts as a docking site during type IV secretion. **2005**, 102, 1685-90	47

880	Phylogeny determined by protein domain content. **2005**, 102, 373-8		174
879	Identification of the rctA gene, which is required for repression of conjugative transfer of rhizobial symbiotic megaplasmids. **2005**, 187, 7341-50		62
878	Novel activator of mannose-specific phosphotransferase system permease expression in Listeria innocua, identified by screening for pediocin AcH resistance. **2005**, 71, 1283-90		22
877	Global profiling of Shewanella oneidensis MR-1: expression of hypothetical genes and improved functional annotations. **2005**, 102, 2099-104		106
876	Three globin lineages belonging to two structural classes in genomes from the three kingdoms of life. **2005**, 102, 11385-9		131
875	Prokaryotic phylogenies inferred from protein structural domains. **2005**, 15, 393-402		30
874	Caspases. Regulating death since the origin of life. **2005**, 137, 841-7		117
873	A sequence sub-sampling algorithm increases the power to detect distant homologues. **2005**, 33, 3772-8		1
872	Channel formation by CarO, the carbapenem resistance-associated outer membrane protein of Acinetobacter baumannii. **2005**, 49, 4876-83		93
871	The Proteomics Protocols Handbook. **2005**,		204
870	A new approach for HMM based protein sequence family modeling and its application to remote homology classification. **2005**,		2
869	Evolution in simple systems and the emergence of complexity. **2005**,		
868	Structural insights into the beta-xylosidase from Trichoderma reesei obtained by synchrotron small-angle X-ray scattering and circular dichroism spectroscopy. **2005**, 44, 15578-84		8
867	Practical lessons from protein structure prediction. **2005**, 33, 1874-91		97
866	Convergent evolution of domain architectures (is rare). **2005**, 21, 1464-71		120
865	InterProScan: protein domains identifier. **2005**, 33, W116-20		1881
864	Conservation of orientation and sequence in protein domain--domain interactions. *Journal of Molecular Biology*, **2005**, 345, 1265-79	6.5	33
863	A ligand-centric analysis of the diversity and evolution of protein-ligand relationships in E.coli. *Journal of Molecular Biology*, **2005**, 347, 415-36	6.5	23

862	Multi-domain proteins in the three kingdoms of life: orphan domains and other unassigned regions. *Journal of Molecular Biology*, **2005**, 348, 231-43	6.5	179
861	Searching for hypothetical proteins: theory and practice based upon original data and literature. **2005**, 77, 90-127		141
860	De novo RNA synthesis and homology modeling of the classical swine fever virus RNA polymerase. **2005**, 112, 9-23		11
859	The PIN-domain toxin-antitoxin array in mycobacteria. **2005**, 13, 360-5		104
858	Protein Sequence Databases. **2005**, 609-618		
857	Protein families and their evolution-a structural perspective. **2005**, 74, 867-900		253
856	Bioinformatics Technologies. **2005**,		10
855	Mining the Arabidopsis thaliana genome for highly-divergent seven transmembrane receptors. **2006**, 7, R96		67
854	PhyloFacts: an online structural phylogenomic encyclopedia for protein functional and structural classification. **2006**, 7, R83		47
853	A gold standard set of mechanistically diverse enzyme superfamilies. **2006**, 7, R8		56
852	Cytochrome c: occurrence and functions. **2006**, 106, 90-115		208
851	E2D: A Novel Tool for Annotating Protein Domains in Expressed Sequence Tags. **2006**,		
850	A machine learning information retrieval approach to protein fold recognition. **2006**, 22, 1456-63		166
849	Protein family expansions and biological complexity. **2006**, 2, e48		167
848	Comparative protein structure modeling using Modeller. **2006**, Chapter 5, Unit-5.6		1717
847	MaGe: a microbial genome annotation system supported by synteny results. **2006**, 34, 53-65		325
846	Structure and mechanism of an ADP-glucose phosphorylase from Arabidopsis thaliana. **2006**, 45, 3154-62		25
845	Structure of Arabidopsis thaliana At1g77540 protein, a minimal acetyltransferase from the COG2388 family. **2006**, 45, 14325-36		9

844	Leveraging enzyme structure-function relationships for functional inference and experimental design: the structure-function linkage database. **2006**, 45, 2545-55		142
843	Conservation of intrinsic disorder in protein domains and families: I. A database of conserved predicted disordered regions. **2006**, 5, 879-87		99
842	The ADP-glucose binding site of the Escherichia coli glycogen synthase. **2006**, 453, 188-96		18
841	Location and characterization of the O-GlcNAcase active site. **2006**, 1760, 829-39		20
840	Identification of protein domains by shotgun proteolysis. *Journal of Molecular Biology*, **2006**, 358, 364-71	6.5	20
839	Protein Repertoire, Evolution of. **2006**,		
838	Structure of an ETHE1-like protein from Arabidopsis thaliana. **2006**, 62, 964-70		37
837	Identification of similar regions of protein structures using integrated sequence and structure analysis tools. **2006**, 6, 4		16
836	Structural proteomics of minimal organisms: conservation of protein fold usage and evolutionary implications. **2006**, 6, 7		9
835	A tale of two ferredoxins: sequence similarity and structural differences. **2006**, 6, 8		21
834	Structure of phage protein BC1872 from Bacillus cereus, a singleton with new fold. **2006**, 64, 280-3		3
833	Strategies for high-throughput comparative modeling: applications to leverage analysis in structural genomics and protein family organization. **2007**, 66, 766-77		20
832	Role of the carboxy-terminal region of the outer membrane protein AatA in the export of dispersin from enteroaggregative Escherichia coli. **2006**, 256, 266-72		6
831	Isolation and characterization of functional insertion sequences of rhizobia. **2006**, 261, 25-31		9
830	Specificity of Streptococcus pyogenes NAD(+) glycohydrolase in cytolysin-mediated translocation. **2006**, 62, 1203-14		29
829	Creating hierarchical models of protein families based on Expressed Sequence Tags: the "Sprockets" analysis pipeline. **2006**, 564, 123-32		
828	Unveiling hidden features of orphan nuclear receptors: the case of the small heterodimer partner (SHP). **2006**, 24, 362-72		25
827	Pattern recognition methods for advanced stochastic protein sequence analysis using HMMs. **2006**, 39, 2267-2280		11

826	Protein interaction networks in plants. **2006**, 224, 771-81	38
825	Some operations research methods for analyzing protein sequences and structures. **2006**, 4, 91-123	5
824	Evolution of +1 programmed frameshifting signals and frameshift-regulating tRNAs in the order Saccharomycetales. **2006**, 63, 545-61	23
823	The Nudix hydrolase superfamily. **2006**, 63, 123-43	436
822	Duplication and combination of P-loop containing nucleotide triphosphate hydrolases superfamily. **2006**, 11, 577-580	1
821	Two families of extracellular phospholipase C genes are present in aspergilli. **2006**, 110, 1140-51	8
820	Automated functional classification of experimental and predicted protein structures. **2006**, 7, 278	5
819	High throughput profile-profile based fold recognition for the entire human proteome. **2006**, 7, 288	17
818	MANET: tracing evolution of protein architecture in metabolic networks. **2006**, 7, 351	42
817	Application of protein structure alignments to iterated hidden Markov model protocols for structure prediction. **2006**, 7, 410	11
816	A framework for protein structure classification and identification of novel protein structures. **2006**, 7, 456	11
815	On single and multiple models of protein families for the detection of remote sequence relationships. **2006**, 7, 48	6
814	A phylogenomic profile of globins. **2006**, 6, 31	173
813	Lactobacillus plantarum gene clusters encoding putative cell-surface protein complexes for carbohydrate utilization are conserved in specific gram-positive bacteria. **2006**, 7, 126	81
812	Coding limits on the number of transcription factors. **2006**, 7, 239	63
811	The Epc-N domain: a predicted protein-protein interaction domain found in select chromatin associated proteins. **2006**, 7, 6	24
810	Genomic patterns of pathogen evolution revealed by comparison of Burkholderia pseudomallei, the causative agent of melioidosis, to avirulent Burkholderia thailandensis. **2006**, 6, 46	121
809	A phylogenomic reconstruction of the protein world based on a genomic census of protein fold architecture. **2006**, 12, 27-40	30

808	Phytome: a platform for plant comparative genomics. **2006**, 34, D724-30	30
807	AGMIAL: implementing an annotation strategy for prokaryote genomes as a distributed system. **2006**, 34, 3533-45	79
806	Genomic scale sub-family assignment of protein domains. **2006**, 34, 3625-33	19
805	A TFIIB-like protein is indispensable for spliced leader RNA gene transcription in Trypanosoma brucei. **2006**, 34, 1676-84	47
804	Chromosome-wide identification of novel imprinted genes using microarrays and uniparental disomies. **2006**, 34, e88	58
803	Web-ware bioinformatical analysis and structure modelling of N-terminus of human multisynthetase complex auxiliary component protein p43. **2006**, 13, 687-91	1
802	The liganding of glycolipid transfer protein is controlled by glycolipid acyl structure. **2006**, 4, e362	48
801	Distinguishing protein-coding from non-coding RNAs through support vector machines. **2006**, 2, e29	111
800	Exploiting protein structure data to explore the evolution of protein function and biological complexity. **2006**, 361, 425-40	21
799	Comparative analysis of proteins with a mucus-binding domain found exclusively in lactic acid bacteria. **2006**, 152, 273-280	135
798	Structure of pyrimidine 5’-nucleotidase type 1. Insight into mechanism of action and inhibition during lead poisoning. **2006**, 281, 20521-9	23
797	Global phylogeny determined by the combination of protein domains in proteomes. **2006**, 23, 2444-54	77
796	Research Article: Biochemical detection, pharmacological inhibition, and phylogenetic analysis of Caenorhabditis elegans metalloproteases. **2006**, 77, 113-126	1
795	A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. **2006**, 188, 2761-73	225
794	Statistical limits to the identification of ion channel domains by sequence similarity. **2006**, 127, 755-66	18
793	Functional bioinformatics for Arabidopsis thaliana. **2006**, 22, 1130-6	20
792	A global topology map of the Saccharomyces cerevisiae membrane proteome. **2006**, 103, 11142-7	113
791	Modern proteomes contain putative imprints of ancient shifts in trace metal geochemistry. **2006**, 103, 17822-7	182

790	Identification of a gene negatively affecting antibiotic production and morphological differentiation in Streptomyces coelicolor A3(2). **2006**, 188, 8368-75	38
789	Discrimination of non-protein-coding transcripts from protein-coding mRNA. **2006**, 3, 40-8	90
788	The predicted secretome of Lactobacillus plantarum WCFS1 sheds light on interactions with its environment. **2006**, 152, 3175-3183	88
787	Sequence analysis and organization of the Neodiprion abietis nucleopolyhedrovirus genome. **2006**, 80, 6952-63	32
786	Structure and mechanism of mouse cysteine dioxygenase. **2006**, 103, 3084-9	157
785	Genome of crocodilepox virus. **2006**, 80, 4978-91	52
784	Identification of regions critically affecting kinetics and allosteric regulation of the Escherichia coli ADP-glucose pyrophosphorylase by modeling and pentapeptide-scanning mutagenesis. **2007**, 189, 5325-33	35
783	Gene3D: comprehensive structural and functional annotation of genomes. **2008**, 36, D414-8	64
782	Using likelihood-free inference to compare evolutionary dynamics of the protein networks of H. pylori and P. falciparum. **2007**, 3, e230	56
781	Superfamily assignments for the yeast proteome through integration of structure prediction with the gene ontology. **2007**, 5, e76	42
780	Following the Viterbi Path to Deduce Flagellar Actin-Interacting Proteins of Leishmania spp.: Report on Cofilins and Twinfilins. **2007**,	1
779	Using phylogeny to improve genome-wide distant homology recognition. **2007**, 3, e3	4
778	Methods of remote homology detection can be combined to increase coverage by 10% in the midnight zone. **2007**, 23, 2353-60	31
777	POGs/PlantRBP: a resource for comparative genomics in plants. **2007**, 35, D852-6	29
776	TOPOFIT-DB, a database of protein structural alignments based on the TOPOFIT method. **2007**, 35, D317-21	13
775	New developments in the InterPro database. **2007**, 35, D224-8	397
774	ForestTreeDB: a database dedicated to the mining of tree transcriptomes. **2007**, 35, D888-94	21
773	The SUPERFAMILY database in 2007: families and functions. **2007**, 35, D308-13	187

772	Gene duplication models for directed networks with limits on growth. **2007**, 2007, P11007-P11007		14
771	Reductive evolution of architectural repertoires in proteomes and the birth of the tripartite world. **2007**, 17, 1572-85		104
770	fastSCOP: a fast web server for recognizing protein structural domains and SCOP superfamilies. **2007**, 35, W438-43		15
769	Plasmids and rickettsial evolution: insight from Rickettsia felis. **2007**, 2, e266		176
768	Hierarchy and feedback in the evolution of the Escherichia coli transcription network. **2007**, 104, 5516-20		75
767	Molecular analysis of duck hepatitis virus type 1 indicates that it should be assigned to a new genus. **2007**, 123, 190-203		150
766	Simple alignment-free methods for protein classification: a case study from G-protein-coupled receptors. **2007**, 89, 602-12		26
765	Glycolipid transfer proteins. **2007**, 1771, 746-60		67
764	Positive selection in phytotoxic protein-encoding genes of Botrytis species. **2007**, 44, 52-63		79
763	Identification of the Salmonella phage epsilon 34 tailspike gene. **2007**, 386, 211-7		6
762	A model of globin evolution. **2007**, 398, 132-42		90
761	A domain of the thyroid adenoma associated gene (THADA) conserved in vertebrates becomes destroyed by chromosomal rearrangements observed in thyroid adenomas. **2007**, 403, 110-7		24
760	Molecular characterisation of a candidate gut sucrase in the pea aphid, Acyrthosiphon pisum. **2007**, 37, 307-17		46
759	Functional expression and characterisation of a gut facilitative glucose transporter, NlHT1, from the phloem-feeding insect Nilaparvata lugens (rice brown planthopper). **2007**, 37, 1138-48		28
758	The transcriptional regulatory network of the amino acid producer Corynebacterium glutamicum. *Journal of Biotechnology*, **2007**, 129, 191-211	3.7	57
757	Determinants of protein function revealed by combinatorial entropy optimization. **2007**, 8, R232		202
756	The innate immune repertoire in cnidaria--ancestral complexity and stochastic gene loss. **2007**, 8, R59		275
755	A network perspective on the evolution of metabolism by gene duplication. **2007**, 8, R26		42

754	Integration of rotation and piston motions in coiled-coil signal transduction. **2007**, 189, 6048-56		27
753	Cysteine dioxygenase: structure and mechanism. **2007**, 3338-49		125
752	Classification of response regulators based on their surface properties. **2007**, 422, 141-69		2
751	Comparative Modeling of Drug Target Proteins. **2007**, 215-236		1
750	Annotating the human proteome: beyond establishing a parts list. **2007**, 1774, 175-91		26
749	Plant progesterone 5beta-reductase is not homologous to the animal enzyme. Molecular evolutionary characterization of P5betaR from Digitalis purpurea. **2007**, 68, 853-64		41
748	Evolution of mitochondrial-type cytochrome c domains and of the protein machinery for their assembly. **2007**, 101, 1798-811		31
747	Matrix Metalloproteinase-9 (MMP-9) polymorphisms in patients with cutaneous malignant melanoma. **2007**, 8, 10		34
746	Crystal structure of an acetyltransferase protein from Vibrio cholerae strain N16961. **2007**, 69, 422-7		4
745	Classification and functional annotation of eukaryotic protein kinases. **2007**, 68, 893-914		121
744	Distant homology detection using a LEngth and STructure-based sequence Alignment Tool (LESTAT). **2008**, 71, 1409-19		5
743	Piecing together the structure-function puzzle: experiences in structure-based functional annotation of hypothetical proteins. **2007**, 7, 2920-32		26
742	The folding and evolution of multidomain proteins. **2007**, 8, 319-30		257
741	ERp57 is essential for efficient folding of glycoproteins sharing common structural domains. **2007**, 26, 28-40		156
740	Arabidopsis extra-large G proteins (XLGs) regulate root morphogenesis. *Plant Journal*, **2008**, 53, 248-63	6.9	83
739	Modeling protein network evolution under genome duplication and domain shuffling. **2007**, 1, 49		36
738	HMM-ModE--improved classification using profile hidden Markov models by optimising the discrimination threshold and modifying emission probabilities with negative training sequences. **2007**, 8, 104		25
737	Improving model construction of profile HMMs for remote homology detection through structural alignment. **2007**, 8, 435		16

736	Molecular evolution of the MAGUK family in metazoan genomes. **2007**, 7, 129		59
735	cTFbase: a database for comparative genomics of transcription factors in cyanobacteria. **2007**, 8, 104		28
734	Analysis of 13000 unique Citrus clusters associated with fruit quality, production and salinity tolerance. **2007**, 8, 31		60
733	Wide diversity in structure and expression profiles among members of the Caenorhabditis elegans globin protein family. **2007**, 8, 356		39
732	Characterization of Entamoeba histolytica alpha-actinin2. **2007**, 154, 82-9		14
731	Comparative protein structure modeling using MODELLER. **2007**, Chapter 2, Unit 2.9		790
730	Yeast two-hybrid map of Arabidopsis TFIID. **2007**, 64, 73-87		24
729	The DNA-binding domain as a functional indicator: the case of the AraC/XylS family of transcription factors. **2008**, 133, 65-76		34
728	Sulfoacetate released during the assimilation of taurine-nitrogen by Neptuniibacter caesariensis: purification of sulfoacetaldehyde dehydrogenase. **2008**, 190, 159-68		27
727	Rethinking proteasome evolution: two novel bacterial proteasomes. **2008**, 66, 494-504		26
726	Evolution of exceptionally large genes in prokaryotes. **2008**, 66, 333-49		2
725	Expression and in silico structural analysis of a rice (Oryza sativa) hemoglobin 5. **2008**, 46, 855-9		13
724	The maize (Zea mays L.) roothairless3 gene encodes a putative GPI-anchored, monocot-specific, COBRA-like protein that significantly affects grain yield. *Plant Journal*, **2008**, 54, 888-98	6.9	118
723	Patterns of evolutionary constraints on genes in humans. **2008**, 8, 275		13
722	Comparative analysis of function and interaction of transcription factors in nematodes: extensive conservation of orthology coupled to rapid sequence evolution. **2008**, 9, 399		41
721	Comparative secretome analysis suggests low plant cell wall degrading capacity in Frankia symbionts. **2008**, 9, 47		40
720	Genome-wide survey of prokaryotic serine proteases: analysis of distribution and domain architectures of five serine protease families in prokaryotes. **2008**, 9, 549		49
719	Protein subfamily assignment using the Conserved Domain Database. **2008**, 1, 114		22

718	Harnessing knowledge from structural genomics. **2008**, 16, 16-8		5
717	The Basics of Protein Sequence Analysis. **2008**, 1-38		
716	Template Based Prediction of Three-Dimensional Protein Structures: Fold Recognition and Comparative Modeling. **2008**, 87-116		1
715	NetGrep: fast network schema searches in interactomes. **2008**, 9, R138		29
714	The hidden universal distribution of amino acid biosynthetic networks: a genomic perspective on their origins and evolution. **2008**, 9, R95		27
713	A structural annotation resource for the selection of putative target proteins in the malaria parasite. **2008**, 7, 90		1
712	Understanding protein evolution: from protein physics to Darwinian selection. **2008**, 59, 105-27		59
711	Architectures and functional coverage of protein-protein interfaces. *Journal of Molecular Biology*, **2008**, 381, 785-802	6.5	95
710	A cyanase is transcriptionally regulated by arginine and involved in cyanate decomposition in Sordaria macrospora. **2008**, 45, 1458-69		32
709	Application of hidden Markov model to products shelf lives. **2008**, 19, 156-161		8
708	A protein domain-based interactome network for C. elegans early embryogenesis. **2008**, 134, 534-45		161
707	Genome-based analysis of heme biosynthesis and uptake in prokaryotic systems. **2008**, 7, 4946-54		43
706	Differentiating protein-coding and noncoding RNA: challenges and ambiguities. **2008**, 4, e1000176		387
705	Simple is beautiful: a straightforward approach to improve the delineation of true and false positives in PSI-BLAST searches. **2008**, 24, 1339-43		18
704	Molecular cloning and characterization of a moss (Ceratodon purpureus) nonsymbiotic hemoglobin provides insight into the early evolution of plant nonsymbiotic hemoglobins. **2008**, 25, 1482-7		21
703	Yeast Pgc1p (YPL206c) controls the amount of phosphatidylglycerol via a phospholipase C-type degradation mechanism. **2008**, 283, 17107-15		33
702	FlagelLink. **2008**,		1
701	Message-passing algorithms for the prediction of protein domain interactions from protein-protein interaction data. **2008**, 24, 2064-70		13

700	Profile Comparer: a program for scoring and aligning profile hidden Markov models. **2008**, 24, 2630-1	78
699	Pfam 10 years on: 10,000 families and still growing. **2008**, 9, 210-9	101
698	The effectiveness of position- and composition-specific gap costs for protein similarity searches. **2008**, 24, i15-23	5
697	Chapter 1. Target selection in structural genomics projects to increase knowledge of protein structure and function space. **2008**, 75, 1-52	1
696	Distribution of Protein Superfamilies in the Three Superkingdoms of Life. **2008**,	
695	An efficient parallel approach for identifying protein families in large-scale metagenomic data sets. **2008**,	7
694	Validating annotations for uncharacterized proteins in Shewanella oneidensis. **2008**, 12, 211-5	5
693	Building a biological space based on protein sequence similarities and biological ontologies. **2008**, 11, 653-60	1
692	. **2008**,	6
691	Probabilistic Models for Long-Range Features in Biosequences. 241-261	
690	. **2008**,	14
689	Recognition of double strand breaks by a mutator protein (MU2) in Drosophila melanogaster. **2009**, 5, e1000473	26
688	Evolutionary descent of prion genes from a ZIP metal ion transport ancestor. **2009**,	0
687	AS-ALPS: a database for analyzing the effects of alternative splicing on protein structure, interaction and network in human and mouse. **2009**, 37, D305-9	27
686	Significant speedup of database searches with HMMs by search space reduction with PSSM family models. **2009**, 25, 3251-8	8
685	Role of the transcriptional regulator RamB (Rv0465c) in the control of the glyoxylate cycle in Mycobacterium tuberculosis. **2009**, 191, 7260-9	42
684	A new taxonomy-based protein fold recognition approach based on autocross-covariance transformation. **2009**, 25, 2655-62	146
683	Benchmarking homology detection procedures with low complexity filters. **2009**, 25, 2500-5	11

682	webPRC: the Profile Comparer for alignment-based searching of public domain databases. **2009**, 37, W48-52	17
681	The evolutionary mechanics of domain organization in proteomes and the rise of modularity in the protein world. **2009**, 17, 66-78	97
680	Protein domain organisation: adding order. **2009**, 10, 39	41
679	Does negative auto-regulation increase gene duplicability?. **2009**, 9, 193	3
678	Comprehensive in silico prediction and analysis of chlamydial outer membrane proteins reflects evolution and life style of the Chlamydiae. **2009**, 10, 634	26
677	A novel virus of the late blight pathogen, Phytophthora infestans, with two RNA segments and a supergroup 1 RNA-dependent RNA polymerase. **2009**, 392, 52-61	31
676	Target selection and annotation for the structural genomics of the amidohydrolase and enolase superfamilies. **2009**, 10, 107-25	24
675	Crystal structure of YfeU protein from Haemophilus influenzae: a predicted etherase involved in peptidoglycan recycling. **2009**, 10, 151-6	2
674	Accurate domain identification with structure-anchored hidden Markov models, saHMMs. **2009**, 76, 343-52	
673	Computational protein design as a tool for fold recognition. **2009**, 77, 139-58	24
672	Prediction of surface exposed proteins in Streptococcus pyogenes, with a potential application to other Gram-positive bacteria. **2009**, 9, 61-73	92
671	Development of an accurate classification system of proteins into structured and unstructured regions that uncovers novel structural domains: its application to human transcription factors. **2009**, 9, 26	32
670	Component-based discriminative classification for hidden Markov models. **2009**, 42, 2637-2648	24
669	Assembly and transport mechanism of tripartite drug efflux systems. **2009**, 1794, 817-25	65
668	Prediction and integration of regulatory and protein-protein interactions. **2009**, 541, 101-43	9
667	The origin, evolution and structure of the protein world. **2009**, 417, 621-37	170
666	Combination of several bioinformatics approaches for the identification of new putative glycosyltransferases in Arabidopsis. **2009**, 8, 743-53	26
665	Detecting sequence and structure homology via an integrative kernel: A case-study in recognizing enzymes. **2009**,	

664	The origin and evolution of modern metabolism. **2009**, 41, 285-97		84
663	Modeling of the full-length Escherichia coli SeqA protein, in complex with DNA. **2009**, 57, e61-6		1
662	Crystal structure of histidine phosphotransfer protein ShpA, an essential regulator of stalk biogenesis in Caulobacter crescentus. *Journal of Molecular Biology*, **2009**, 390, 686-98	6.5	12
661	Histone deacetylases regulate multicellular development in the social amoeba Dictyostelium discoideum. *Journal of Molecular Biology*, **2009**, 391, 833-48	6.5	12
660	Online tools for predicting integral membrane proteins. **2009**, 528, 3-23		8
659	Improved homology-driven computational validation of protein-protein interactions motivated by the evolutionary gene duplication and divergence hypothesis. **2009**, 10, 21		5
658	Genome and proteome annotation: organization, interpretation and integration. **2009**, 6, 129-47		36
657	Computational Systems Biology. **2009**,		3
656	Biomedical informatics. Preface. **2009**, 569, v		
655	Bioinformatics: A Concept-Based Introduction. **2009**,		1
654	Tandem and cryptic amino acid repeats accumulate in disordered regions of proteins. **2009**, 10, R59		87
653	DNA-Prot: identification of DNA binding proteins from protein sequence information using random forest. **2009**, 26, 679-86		101
652	A comparative evolutionary study of transcription networks. The global role of feedback and hierachical structures. **2009**, 5, 170-9		12
651	Chapter 11 Unexpected Turns and Twists in Structure/Function of PR-Proteins that Connect Energy Metabolism and Immunity. **2009**, 51, 439-489		13
650	Genetic makeup of the Corynebacterium glutamicum LexA regulon deduced from comparative transcriptomics and in vitro DNA band shift assays. **2009**, 155, 1459-1477		38
649	Zebrafish granulocyte colony-stimulating factor receptor signaling promotes myelopoiesis and myeloid cell migration. **2009**, 113, 2535-46		92
648	Hidden Markov models and the Viterbi algorithm applied to integrated bioinformatics analyses of putative flagellar actin-interacting proteins in Leishmania spp.. **2009**, 1, 420		2
647	Basic protein structure prediction for the biologist: A review. **2010**, 62, 857-871		18

646	Expression of usherin in the anthozoan Nematostella vectensis. **2010**, 218, 105-12	7
645	A comparative study of multi-classification methods for protein fold recognition. **2010**, 1, 332	5
644	Part II: defining and quantifying individual and co-cultured intracellular proteomes of two thermophilic microorganisms by GeLC-MS2 and spectral counting. **2010**, 398, 391-404	10
643	Part I: characterization of the extracellular proteome of the extreme thermophile Caldicellulosiruptor saccharolyticus by GeLC-MS2. **2010**, 398, 377-89	16
642	TAT-pathway-dependent lipoproteins as a niche-based adaptation in prokaryotes. **2010**, 70, 359-70	26
641	Homologous recombination in the antibiotic producer Penicillium chrysogenum: strain DeltaPcku70 shows up-regulation of genes from the HOG pathway. **2010**, 85, 1081-94	62
640	The tetratricopeptide repeats (TPR)-like superfamily of proteins in Leishmania spp., as revealed by multi-relational data mining. **2010**, 31, 2178-2189	3
639	Some operations research methods for analyzing protein sequences and structures. **2010**, 175, 9-35	8
638	Understanding the Horizontal Dimension of Molecular Evolution to Annotate, Classify, and Discover Proteins with Functional Domains. **2010**, 25, 82-94	
637	EfeO-cupredoxins: major new members of the cupredoxin superfamily with roles in bacterial iron transport. **2010**, 23, 1-17	47
636	Generation and gene ontology based analysis of expressed sequence tags (EST) from a Panax ginseng C. A. Meyer roots. **2010**, 37, 3465-72	50
635	AlgK is a TPR-containing protein and the periplasmic component of a novel exopolysaccharide secretin. **2010**, 18, 265-73	82
634	Predicting gene function using hierarchical multi-label decision tree ensembles. **2010**, 11, 2	105
633	PhenoFam-gene set enrichment analysis through protein structural information. **2010**, 11, 254	4
632	The Zur regulon of Corynebacterium glutamicum ATCC 13032. **2010**, 11, 12	50
631	Analyses of genome architecture and gene expression reveal novel candidate virulence factors in the secretome of Phytophthora infestans. **2010**, 11, 637	141
630	The complete genome sequence of Corynebacterium pseudotuberculosis FRC41 isolated from a 12-year-old girl with necrotizing lymphadenitis reveals insights into gene-regulatory networks contributing to virulence. **2010**, 11, 728	81
629	Transposon identification using profile HMMs. **2010**, 11 Suppl 1, S10	3

628	SoyDB: a knowledge database of soybean transcription factors. **2010**, 10, 14	89
627	A detailed genome-wide reconstruction of mouse metabolism based on human Recon 1. **2010**, 4, 140	114
626	Structural identity of telomeric complexes. **2010**, 584, 3785-99	38
625	Genome Evolution Studied Through Protein Structure. **2010**, 153-164	1
624	Pacidamycin biosynthesis: identification and heterologous expression of the first uridyl peptide antibiotic gene cluster. **2010**, 11, 1700-9	62
623	Computation of mutual information from Hidden Markov Models. **2010**, 34, 328-33	2
622	Cell surface display of chimeric glycoproteins via the S-layer of Paenibacillus alvei. **2010**, 345, 1422-31	18
621	Identification of recurring protein structure microenvironments and discovery of novel functional sites around CYS residues. **2010**, 10, 4	10
620	Comparison of DNA binding across protein superfamilies. **2010**, 78, 52-62	11
619	Is Asp-His-Ser/Thr-Trp tetrad hydrogen-bond network important to WD40-repeat proteins: a statistical and theoretical study. **2010**, 78, 1186-94	15
618	The unique set of putative membrane-associated anti-sigma factors in Clostridium thermocellum suggests a novel extracellular carbohydrate-sensing mechanism involved in gene regulation. **2010**, 308, 84-93	58
617	Renalase, a new secretory enzyme responsible for selective degradation of catecholamines: achievements and unsolved problems. **2010**, 75, 951-8	25
616	. **2010**,	7
615	Biodiversity of mannose-specific adhesion in Lactobacillus plantarum revisited: strain-specific domain composition of the mannose-adhesin. **2010**, 1, 61-6	37
614	TreeVector: scalable, interactive, phylogenetic trees for the web. **2010**, 5, e8934	25
613	Quality measures for protein alignment benchmarks. **2010**, 38, 2145-53	80
612	Integrative Protein Fold Recognition by Alignments and Machine Learning. **2010**, 195-218	
611	Role of vimA in cell surface biogenesis in Porphyromonas gingivalis. **2010**, 156, 2180-2193	13

610	Widespread head-to-head hydrocarbon biosynthesis in bacteria and role of OleA. **2010**, 76, 3850-62		96
609	Heterogeneous patterns of gene-expression diversification in mammalian gene duplicates. **2010**, 27, 325-35		32
608	Paraoxonase 1, quorum sensing, and P. aeruginosa infection: a novel model. **2010**, 660, 183-93		21
607	Genomic repertoires of DNA-binding transcription factors across the tree of life. **2010**, 38, 7364-77		102
606	Paraoxonases in Inflammation, Infection, and Toxicology. **2010**,		1
605	ANALYZING GENE COEXPRESSION DATA BY AN EVOLUTIONARY MODEL. **2010**,		
604	Two components of a velvet-like complex control hyphal morphogenesis, conidiophore development, and penicillin biosynthesis in Penicillium chrysogenum. **2010**, 9, 1236-50		130
603	Cholera- and anthrax-like toxins are among several new ADP-ribosyltransferases. **2010**, 6, e1001029		43
602	Prototypes of elementary functional loops unravel evolutionary connections between protein functions. **2010**, 26, i497-503		28
601	Emergence and evolution of modern molecular functions inferred from phylogenomic analysis of ontological data. **2010**, 27, 1710-33		38
600	Marsupial Sequencing Projects and Bioinformatics Challenges. **2010**, 121-132		1
599	A structure-centric view of protein evolution, design, and adaptation. **2007**, 75, 133-91, xi-xii		16
598	Modeling sequence and function similarity between proteins for protein functional annotation. **2010**, 2010, 499-502		4
597	The evolution and structure prediction of coiled coils across all genomes. *Journal of Molecular Biology*, **2010**, 403, 480-93	6.5	67
596	Amaranth lunasin-like peptide internalizes into the cell nucleus and inhibits chemical carcinogen-induced transformation of NIH-3T3 cells. **2010**, 31, 1635-42		42
595	Amino acid-containing membrane lipids in bacteria. **2010**, 49, 46-60		111
594	A systematic screen for protein-lipid interactions in Saccharomyces cerevisiae. **2010**, 6, 430		132
593	Marsupial Genetics and Genomics. **2010**,		5

592	Identity and divergence of protein domain architectures after the yeast whole-genome duplication event. **2010**, 6, 2305-15	16
591	History of biological metal utilization inferred through phylogenomic analysis of protein structures. **2010**, 107, 10567-72	210
590	The functional make up of proteomes is remarkably conserved. **2011**,	
589	Reductive evolution of proteomes and protein structures. **2011**, 108, 11954-8	58
588	A universal molecular clock of protein folds and its power in tracing the early history of aerobic metabolism and planet oxygenation. **2011**, 28, 567-82	104
587	Principles and patterns in the interaction between mono-heme cytochrome c and its partners in electron transfer processes. **2011**, 3, 354-62	8
586	Lafora disease E3-ubiquitin ligase malin is related to TRIM32 at both the phylogenetic and functional level. **2011**, 11, 225	16
585	Mnl2, a novel component of the ER associated protein degradation pathway. **2011**, 414, 528-32	5
584	The crystal structure of the leptospiral hypothetical protein LIC12922 reveals homology with the periplasmic chaperone SurA. **2011**, 173, 312-22	8
583	Computational interaction analysis of organophosphorus pesticides with different metabolic proteins in humans. **2011**, 25, 335-47	22
582	Annotation of Protein Domains Reveals Remarkable Conservation in the Functional Make up of Proteomes Across Superkingdoms. **2011**, 2, 869-911	22
581	siRNA knockdown of ribosomal protein gene RPL19 abrogates the aggressive phenotype of human prostate cancer. **2011**, 6, e22672	37
580	HMMerThread: detecting remote, functional conserved domains in entire genomes by combining relaxed sequence-database searches with fold recognition. **2011**, 6, e17568	2
579	Assembling a protein-protein interaction map of the SSU processome from existing datasets. **2011**, 6, e17701	23
578	A bioinformatics classifier and database for heme-copper oxygen reductases. **2011**, 6, e19117	40
577	Origins and evolution of the HET-s prion-forming protein: searching for other amyloid-forming solenoids. **2011**, 6, e27342	8
576	Besnoitia besnoiti protein disulfide isomerase (BbPDI): molecular characterization, expression and in silico modelling. **2011**, 129, 164-74	12
575	The retroviral proteinase active site and the N-terminus of Ddi1 are required for repression of protein secretion. **2011**, 585, 139-42	18

574	Screening of Aspergillus-derived FAD-glucose dehydrogenases from fungal genome database. **2011** , 33, 2255-63	30
573	OsBC1L4 encodes a COBRA-like protein that affects cellulose synthesis in rice. **2011**, 75, 333-45	50
572	In silico gene expression analysis in Codonopsis lanceolata root. **2011**, 38, 3541-9	3
571	Proteome evolution and the metabolic origins of translation and cellular life. **2011**, 72, 14-33	47
570	A model of genetic search for beneficial mutations: estimating the constructive capacities of mutagenesis. **2011**, 73, 337-54	
569	KF-1 ubiquitin ligase: anxiety suppressor model. **2011**, 60, 69-75	1
568	Binary classification of protein molecules into intrinsically disordered and ordered segments. **2011**, 11, 29	56
567	Probable presence of an ubiquitous cryptic mitochondrial gene on the antisense strand of the cytochrome oxidase I gene. **2011**, 6, 56	60
566	Crystallization and preliminary X-ray analysis of the Entamoeba histolytica □actinin-2 rod domain. **2011**, 67, 1214-7	1
565	A network of SCOP hidden Markov models and its analysis. **2011**, 12, 191	2
564	A discriminative method for family-based protein remote homology detection that combines inductive logic programming and propositional models. **2011**, 12, 83	7
563	The proteomic complexity and rise of the primordial ancestor of diversified life. **2011**, 11, 140	72
562	Are viruses a source of new protein folds for organisms? - Virosphere structure space and evolution. **2011**, 33, 626-35	28
561	SCORER 2.0: an algorithm for distinguishing parallel dimeric and trimeric coiled-coil sequences. **2011**, 27, 1908-14	40
560	Evidences for increased expression variation of duplicate genes in budding yeast: from cis- to trans-regulation effects. **2011**, 39, 837-47	21
559	A performance enhanced PSI-BLAST based on hybrid alignment. **2011**, 27, 31-7	10
558	A novel Lawsonia intracellularis autotransporter protein is a prominent antigen. **2011**, 18, 1282-7	10
557	Phylogenetic relationships of 3/3 and 2/2 hemoglobins in Archaeplastida genomes to bacterial and other eukaryote hemoglobins. **2011**, 4, 42-58	33

556	Positive transcriptional control of the pyridoxal phosphate biosynthesis genes pdxST by the MocR-type regulator PdxR of Corynebacterium glutamicum ATCC 13032. **2011**, 157, 77-88	33
555	A novel FK-506-binding-like protein that lacks peptidyl-prolyl isomerase activity is involved in intracellular infection and in vivo virulence of Burkholderia pseudomallei. **2011**, 157, 2629-2638	20
554	A systems biology approach for the investigation of the heparin/heparan sulfate interactome. **2011** , 286, 19892-904	164
553	Molecular evolution and selection patterns of plant F-box proteins with C-terminal kelch repeats. **2011**, 155, 835-50	55
552	SUPERFAMILY 1.75 including a domain-centric gene ontology method. **2011**, 39, D427-34	124
551	The quantitative proteome of a human cell line. **2011**, 7, 549	586
550	Androglobin: a chimeric globin in metazoans that is preferentially expressed in Mammalian testes. **2012**, 29, 1105-14	77
549	Viral evolution: Primordial cellular origins and late adaptation to parasitism. **2012**, 2, 247-252	28
548	IDEAL: Intrinsically Disordered proteins with Extensive Annotations and Literature. **2012**, 40, D507-11	68
547	Identification of positive selection in disease response genes within members of the Poaceae. **2012** , 7, 1667-75	6
546	D□P□: database of disordered protein predictions. **2013**, 41, D508-16	398
545	ASGARD: an open-access database of annotated transcriptomes for emerging model arthropod species. **2012**, 2012, bas048	19
544	Genome3D: a UK collaborative project to annotate genomic sequences with predicted 3D structures based on SCOP and CATH domains. **2013**, 41, D499-507	48
543	Structure of the Archaeoglobus fulgidus orphan ORF AF1382 determined by sulfur SAD from a moderately diffracting crystal. **2012**, 68, 1242-52	5
542	Do miRNAs have a deep evolutionary history?. **2012**, 34, 857-66	77
541	Structural prediction of a novel chitinase from the psychrophilic Glaciozyma antarctica PI12 and an analysis of its structural properties and function. **2012**, 26, 947-61	36
540	The purification of the Chlamydomonas reinhardtii chloroplast ClpP complex: additional subunits and structural features. **2012**, 80, 189-202	15
539	The transcriptional regulatory network of Corynebacterium jeikeium K411 and its interaction with metabolic routes contributing to human body odor formation. *Journal of Biotechnology*, **2012**, 159, 235-48 37	12

538	TrAnsFuSE refines the search for protein function: oxidoreductases. **2012**, 4, 765-77	9
537	pGraph: Efficient Parallel Construction of Large-Scale Protein Sequence Homology Graphs. **2012**, 23, 1923-1933	22
536	Functional analyses of regulators of G protein signaling in Gibberella zeae. **2012**, 49, 511-20	34
535	Tools to kill: genome of one of the most destructive plant pathogenic fungi Macrophomina phaseolina. **2012**, 13, 493	122
534	Functional characterization of a tomato COBRA-like gene functioning in fruit development and ripening. **2012**, 12, 211	36
533	Distribution and prediction of catalytic domains in 2-oxoglutarate dependent dioxygenases. **2012**, 5, 410	11
532	Structure of RdxA--an oxygen-insensitive nitroreductase essential for metronidazole activation in Helicobacter pylori. **2012**, 279, 4306-17	34
531	Giant viruses coexisted with the cellular ancestors and represent a distinct supergroup along with superkingdoms Archaea, Bacteria and Eukarya. **2012**, 12, 156	92
530	Pleiotropic roles of a ribosomal protein in Dictyostelium discoideum. **2012**, 7, e30644	3
529	A phylogenetic analysis of the globins in fungi. **2012**, 7, e31856	23
528	Ribosomal history reveals origins of modern protein synthesis. **2012**, 7, e32776	103
527	3D profile-based approach to proteome-wide discovery of novel human chemokines. **2012**, 7, e36151	6
526	The transcriptome profile of the mosquito Culex quinquefasciatus following permethrin selection. **2012**, 7, e47163	47
525	Horizontal gene transfers as metagenomic gene duplications. **2012**, 8, 790-5	9
524	Characterization of tannase protein sequences of bacteria and fungi: an in silico study. **2012**, 31, 306-27	25
523	Target selection for structural genomics based on combining fold recognition and crystallisation prediction methods: application to the human proteome. **2012**, 13, 37-46	7
522	The phylogenomic roots of modern biochemistry: origins of proteins, cofactors and protein biosynthesis. **2012**, 74, 1-34	64
521	Evolution and functional diversification of the small heat shock protein/□crystallin family in higher plants. **2012**, 235, 1299-313	53

520	The superfamily of heme-copper oxygen reductases: types and evolutionary considerations. **2012**, 1817, 629-37	121
519	Structure and function of the AAA+ nucleotide binding pocket. **2012**, 1823, 2-14	194
518	Protein domain structure uncovers the origin of aerobic metabolism and the rise of planetary oxygen. **2012**, 20, 67-76	45
517	The NtrY/X two-component system of Brucella spp. acts as a redox sensor and regulates the expression of nitrogen respiration enzymes. **2012**, 85, 39-50	44
516	The evolutionary history of protein fold families and proteomes confirms that the archaeal ancestor is more ancient than the ancestors of other superkingdoms. **2012**, 12, 13	45
515	In silico characterization of alkaline proteases from different species of Aspergillus. **2012**, 166, 243-57	25
514	Molecular cloning, tissue distribution and ontogenetic expression of Xiang pig chemerin and its involvement in regulating energy metabolism through Akt and ERK1/2 signaling pathways. **2012**, 39, 1887-94	16
513	IGSF9 family proteins. **2013**, 38, 1236-51	15
512	Expression of a conifer COBRA-like gene ClCOBL1 from Chinese fir (Cunninghamia lanceolata) alters the leaf architecture in tobacco. **2013**, 70, 483-91	14
511	A domain-centric solution to functional genomics via dcGO Predictor. **2013**, 14 Suppl 3, S9	32
510	Reconstituting protein interaction networks using parameter-dependent domain-domain interactions. **2013**, 14, 154	15
509	Variant surface antigens of malaria parasites: functional and evolutionary insights from comparative gene family classification and analysis. **2013**, 14, 427	52
508	Bacterial and archaeal globins - a revised perspective. **2013**, 1834, 1789-800	82
507	Biology, Controls and Models of Tree Volatile Organic Compound Emissions. **2013**,	24
506	eIF5A dimerizes not only in vitro but also in vivo and its molecular envelope is similar to the EF-P monomer. **2013**, 44, 631-44	9
505	Sequence and structural investigation of a novel psychrophilic □-amylase from Glaciozyma antarctica PI12 for cold-adaptation analysis. **2013**, 19, 3369-83	26
504	ccm2-like is required for cardiovascular development as a novel component of the Heg-CCM pathway. **2013**, 376, 74-85	24
503	Characterization of a serine hydrolase targeted by acyl-protein thioesterase inhibitors in Toxoplasma gondii. **2013**, 288, 27002-27018	19

502	Exonic transcription factor binding directs codon choice and affects protein evolution. **2013**, 342, 1367-72	201
501	The Biochemistry and Molecular Biology of Volatile Messengers in Trees. **2013**, 47-93	23
500	Comparative transcriptome analysis of geographically distinct virulent and attenuated Babesia bovis strains reveals similar gene expression changes through attenuation. **2013**, 14, 763	39
499	Microbial eukaryote globins. **2013**, 63, 391-446	36
498	Sequence polymorphism in the Trypanosoma rangeli HSP70 coding genes allows typing of the parasite KP1(+) and KP1(-) groups. **2013**, 133, 447-53	4
497	Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. **2013**, 34, 57-65	723
496	Anti-viral immune responses in a primitive lung: characterization and expression analysis of interferon-inducible immunoproteasome subunits LMP2, LMP7 and MECL-1 in a sarcopterygian fish, the Nigerian spotted lungfish (Protopterus dolloi). **2013**, 41, 657-65	5
495	Rooted phylogeny of the three superkingdoms. **2013**, 95, 1593-604	35
494	Competition sensing: the social side of bacterial stress responses. **2013**, 11, 285-93	239
493	Evolutionary optimization of protein folding. **2013**, 9, e1002861	35
492	Protein complexes are under evolutionary selection to assemble via ordered pathways. **2013**, 153, 461-70	153
491	A daily-updated tree of (sequenced) life as a reference for genome research. **2013**, 3, 2015	39
490	Searching for likeness in a database of macromolecular complexes. **2013**, 53, 2634-47	1
489	Automated Sequence-Based Approaches for Identifying Domain Families. **2013**, 1-24	1
488	Structural Annotations of Genomes with Superfamily and G3D. **2013**, 69-97	
487	Predicting the functional consequences of cancer-associated amino acid substitutions. **2013**, 29, 1504-10	154
486	A network of HMG-box transcription factors regulates sexual cycle in the fungus Podospora anserina. **2013**, 9, e1003642	39
485	Enzyme reaction annotation using cloud techniques. **2013**, 2013, 140237	3

484	Exploring fold space preferences of new-born and ancient protein superfamilies. **2013**, 9, e1003325	24
483	Masculinization of gene expression is associated with exaggeration of male sexual dimorphism. **2013**, 9, e1003697	76
482	Trypanosome CNOT10 is essential for the integrity of the NOT deadenylase complex and for degradation of many mRNAs. **2013**, 41, 1211-22	23
481	Genomic and secretomic analyses reveal unique features of the lignocellulolytic enzyme system of Penicillium decumbens. **2013**, 8, e55185	119
480	The challenge of increasing Pfam coverage of the human proteome. **2013**, 2013, bat023	16
479	DcGO: database of domain-centric ontologies on functions, phenotypes, diseases and more. **2013**, 41, D536-44	72
478	Transcriptomic analysis of Ustilago maydis infecting Arabidopsis reveals important aspects of the fungus pathogenic mechanisms. **2013**, 8,	19
477	PlantTFcat: an online plant transcription factor and transcriptional regulator categorization and analysis tool. **2013**, 14, 321	87
476	mRNA-Seq and microarray development for the Grooved Carpet shell clam, Ruditapes decussatus: a functional approach to unravel host-parasite interaction. **2013**, 14, 741	31
475	Genomics-guided analysis of NAD recycling yields functional elucidation of COG1058 as a new family of pyrophosphatases. **2013**, 8, e65595	13
474	Comparative analysis of proteomes and functionomes provides insights into origins of cellular diversification. **2013**, 2013, 648746	17
473	Comparative analysis of barophily-related amino acid content in protein domains of Pyrococcus abyssi and Pyrococcus furiosus. **2013**, 2013, 680436	6
472	Comparative transcriptome analysis of four prymnesiophyte algae. **2014**, 9, e97801	25
471	Genome-wide identification and tissue-specific expression analysis of UDP-glycosyltransferases genes confirm their abundance in Cicer arietinum (Chickpea) genome. **2014**, 9, e109715	27
470	Structure Based Functional Annotation of Putative Conserved Proteins from Treponema pallidum: Search for a Potential Drug Target. **2014**, 12, 46-59	13
469	Duplication of a promiscuous transcription factor drives the emergence of a new regulatory network. **2014**, 5, 4868	44
468	Tracing the evolution of FERM domain of Kindlins. **2014**, 80, 193-204	11
467	The diversity of zinc-finger genes on human chromosome 19 provides an evolutionary mechanism for defense against inherited endogenous retroviruses. **2014**, 21, 381-7	43

466	Comparative Modeling of Drug Target Proteins?. **2014**,	2
465	De novo assembly of Aureococcus anophagefferens transcriptomes reveals diverse responses to the low nutrient and low light conditions present during blooms. **2014**, 5, 375	36
464	Cyclic di-GMP-dependent signaling pathways in the pathogenic Firmicute Listeria monocytogenes. **2014**, 10, e1004301	56
463	Global patterns of protein domain gain and loss in superkingdoms. **2014**, 10, e1003452	48
462	Protein flexibility facilitates quaternary structure assembly and evolution. **2014**, 12, e1001870	56
461	Pre-metazoan origins and evolution of the cadherin adhesome. **2014**, 3, 1183-95	34
460	Structure, evolution and virtual screening of NDM-1 strain from Kolkata. **2014**, 10, 235-63	2
459	Revisiting amino acid substitution matrices for identifying distantly related proteins. **2014**, 30, 317-25	31
458	From workstations to workbenches: Towards predicting physicochemically viable protein-protein interactions across a host and a pathogen. **2014**, 66, 759-74	9
457	Pairwise decomposition of an MMGBSA energy function for computational protein design. **2014**, 35, 1371-87	32
456	An allosteric switch for pro-HGF/Met signaling using zymogen activator peptides. **2014**, 10, 567-73	20
455	In silico characterization of an atypical MAPK phosphatase of Plasmodium falciparum as a suitable target for drug discovery. **2014**, 84, 158-68	5
454	The Structure-Function Linkage Database. **2014**, 42, D521-30	169
453	BuT2 is a member of the third major group of hAT transposons and is involved in horizontal transfer events in the genus Drosophila. **2014**, 6, 352-65	11
452	From sequence to enzyme mechanism using multi-label machine learning. **2014**, 15, 150	14
451	MMDB and VAST+: tracking structural similarities between macromolecular complexes. **2014**, 42, D297-303	194
450	SCOPe: Structural Classification of Proteins--extended, integrating SCOP and ASTRAL data and classification of new structures. **2014**, 42, D304-9	438
449	In Silico Prediction of T and B Cell Epitopes of Der f 25 in Dermatophagoides farinae. **2014**, 2014, 483905	19

448	Cloning, purification and preliminary crystallographic analysis of the Helicobacter pylori pseudaminic acid biosynthesis N-acetyltransferase PseH. **2014**, 70, 1276-9		7
447	Frequent gene fissions associated with human pathogenic bacteria. **2014**, 103, 65-75		3
446	Performance of protein disorder prediction programs on amino acid substitutions. **2014**, 35, 794-804		19
445	Tools for advanced and targeted genetic manipulation of the β-lactam antibiotic producer Acremonium chrysogenum. *Journal of Biotechnology*, **2014**, 169, 51-62	3.7	25
444	Heat Shock Proteins of Malaria. **2014**,		4
443	Functional divergence for every paralog. **2014**, 31, 984-92		38
442	Gene3D: Multi-domain annotations for protein sequence and comparative genome analysis. **2014**, 42, D240-5		42
441	A novel flavivirus in the soybean cyst nematode. **2014**, 95, 1272-1280		40
440	Evolution of the ribosome at atomic resolution. **2014**, 111, 10251-6		129
439	Comparative Protein Structure Modeling Using MODELLER. **2014**, 47, 5.6.1-32		754
438	General mechanism of two-state protein folding kinetics. **2014**, 136, 11420-7		44
437	An integrative computational model for large-scale identification of metalloproteins in microbial genomes: a focus on iron-sulfur cluster proteins. **2014**, 6, 1913-30		11
436	Fast pseudolikelihood maximization for direct-coupling analysis of protein structure from many homologous amino-acid sequences. **2014**, 276, 341-356		97
435	Ranking non-synonymous single nucleotide polymorphisms based on disease concepts. **2014**, 8, 11		105
434	Using hidden Markov models to predict DNA-binding proteins with sequence and structure information. **2014**, 18, 2365-2376		5
433	Think globally and solve locally: secondary memory-based network learning for automated multi-species function prediction. **2014**, 3, 5		11
432	Classification of intrinsically disordered regions and proteins. **2014**, 114, 6589-631		1141
431	Benchmarking of methods for genomic taxonomy. **2014**, 52, 1529-39		141

430	Hemoglobins in the genome of the cryptomonad Guillardia theta. **2014**, 9, 7		2
429	Dnmt1-independent CG methylation contributes to nucleosome positioning in diverse eukaryotes. **2014**, 156, 1286-1297		111
428	Emerging roles for protein S-palmitoylation in Toxoplasma biology. **2014**, 44, 121-31		20
427	Structural and functional analysis of a novel psychrophilic -mannanase from Glaciozyma antarctica PI12. **2014**, 28, 685-98		24
426	The use of soluble protein structures in modeling helical proteins in a layered membrane. **2014**, 32, 308-18		1
425	Draft Genome Sequence of the Fungus Trametes hirsuta 072. **2015**, 3,		10
424	The value of protein structure classification information-Surveying the scientific literature. **2015**, 83, 2025-38		14
423	Quantitative differential proteomics of yeast extracellular matrix: there is more to it than meets the eye. **2015**, 15, 271		11
422	Mean field theory for biology inspired duplication-divergence network model. **2015**, 25, 083106		5
421	Origin and implications of zero degeneracy in networks spectra. **2015**, 25, 043110		16
420	Structure prediction of Fe(II) 2-oxoglutarate dioxygenase from a psychrophilic yeast Glaciozyma antarctica PI12. **2015**,		1
419	Homology-Based Prediction of Potential Protein-Protein Interactions between Human Erythrocytes and Plasmodium falciparum. *Bioinformatics and Biology Insights*, **2015**, 9, 195-206	5.3	11
418	Why do Sequence Signatures Predict Enzyme Mechanism? Homology versus Chemistry. **2015**, 11, 267-74		3
417	The Haemonchus contortus kinome--a resource for fundamental molecular investigations and drug discovery. **2015**, 8, 623		12
416	Following the Footsteps of Chlamydial Gene Regulation. **2015**, 32, 3035-46		17
415	Evolution of the EKA family of powdery mildew avirulence-effector genes from the ORF 1 of a LINE retrotransposon. **2015**, 16, 917		21
414	Coiled-coil length: Size does matter. **2015**, 83, 2162-9		9
413	SInCRe-structural interactome computational resource for Mycobacterium tuberculosis. **2015**, 2015, bav060		5

412	Selection of recombinant anti-SH3 domain antibodies by high-throughput phage display. **2015**, 24, 1890-900	12
411	’M□age □trois’: a selfish genetic element uses a virus to propagate within Thermotogales. **2015**, 17, 3278-88	16
410	Interactive Naive Bayesian network: A new approach of constructing gene-gene interaction network for cancer classification. **2015**, 26 Suppl 1, S1929-36	2
409	Large-Scale Evolutionary Analysis of Genes and Supergene Clusters from Terpenoid Modular Pathways Provides Insights into Metabolic Diversification in Flowering Plants. **2015**, 10, e0128808	12
408	The Evolution of the Secreted Regulatory Protein Progranulin. **2015**, 10, e0133749	37
407	Activity Augmentation of Amphioxus Peptidoglycan Recognition Protein BbtPGRP3 via Fusion with a Chitin Binding Domain. **2015**, 10, e0140953	3
406	Function-selective domain architecture plasticity potentials in eukaryotic genome evolution. **2015**, 119, 269-77	3
405	Protein function from its emergence to diversity in contemporary proteins. **2015**, 12, 045002	41
404	Staphylococcus aureus Survives with a Minimal Peptidoglycan Synthesis Machine but Sacrifices Virulence and Antibiotic Resistance. **2015**, 11, e1004891	55
403	xiNET: cross-link network maps with residue resolution. **2015**, 14, 1137-47	167
402	Bioinformatics and Biomedical Engineering. **2015**,	3
401	Splice junctions are constrained by protein disorder. **2015**, 43, 4814-22	13
400	Structural biology and genome evolution: An introduction. **2015**, 119, 205-8	6
399	Enriching the annotation of Mycobacterium tuberculosis H37Rv proteome using remote homology detection approaches: insights into structure and function. **2015**, 95, 14-25	8
398	Online resources for genomic analysis using high-throughput sequencing. **2015**, 2015, 324-35	5
397	Structural prediction of a novel laminarinase from the psychrophilic Glaciozyma antarctica PI12 and its temperature adaptation analysis. **2015**, 21, 63	11
396	Untangling the origin of viruses and their impact on cellular evolution. **2015**, 1341, 61-74	25
395	The evaluation of tools used to predict the impact of missense variants is hindered by two types of circularity. **2015**, 36, 513-23	178

394	Rab5-family guanine nucleotide exchange factors bind retromer and promote its recruitment to endosomes. **2015**, 26, 1119-28	23
393	In-depth characterisation of the lamb meat proteome from longissimus lumborum. **2015**, 6, 28-41	14
392	Mitochondrial genomes are retained by selective constraints on protein targeting. **2015**, 112, 10154-61	52
391	Identification of functional candidates amongst hypothetical proteins of Mycobacterium leprae Br4923, a causative agent of leprosy. **2015**, 58, 25-42	13
390	Predictive and comparative analysis of Ebolavirus proteins. **2015**, 14, 2785-97	10
389	The human RBPome: from genes and proteins to human disease. **2015**, 127, 61-70	77
388	Solution structure of a soluble fragment derived from a membrane protein by shotgun proteolysis. **2015**, 28, 445-50	2
387	A novel transposable Mu-like prophage in Bacillus alcalophilus CGMCC 1.3604 (ATCC 27647). **2015**, 30, 63-5	1
386	Structure-based function analysis of putative conserved proteins with isomerase activity from Haemophilus influenzae. **2015**, 5, 741-763	11
385	Comparative transcriptome profiling approach to glean virulence and immunomodulation-related genes of Fasciola hepatica. **2015**, 16, 366	14
384	Transcriptome-wide analysis of SAMe superfamily to novelty phosphoethanolamine N-methyltransferase copy in Lonicera japonica. **2014**, 16, 521-34	7
383	Long non-coding RNA discovery across the genus anopheles reveals conserved secondary structures within and beyond the Gambiae complex. **2015**, 16, 337	56
382	MPIC: a mitochondrial protein import components database for plant and non-plant species. **2015**, 56, e10	18
381	Evolutionary patterns in coiled-coils. **2015**, 7, 545-56	15
380	The ribosome challenge to the RNA world. **2015**, 80, 143-61	57
379	Prediction of Human Gene - Phenotype Associations by Exploiting the Hierarchical Structure of the Human Phenotype Ontology. **2015**, 66-77	3
378	Molecular characterization and expression analysis of hsp60 gene homologue of sheep blowfly, Lucilia cuprina. **2015**, 52, 24-37	7
377	The SUPERFAMILY 1.75 database in 2014: a doubling of data. **2015**, 43, D227-33	62

376	In silico approaches for the identification of virulence candidates amongst hypothetical proteins of Mycoplasma pneumoniae 309. **2015**, 59 Pt A, 67-80	22
375	Defining the Domain Arrangement of the Mammalian Target of Rapamycin Complex Component Rictor Protein. **2015**, 22, 876-86	11
374	A phylogenomic data-driven exploration of viral origins and evolution. **2015**, 1, e1500527	110
373	The mouse wellhaarig (we) mutations result from defects in epidermal-type transglutaminase 3 (Tgm3). **2015**, 116, 187-91	7
372	A proteome quality index. **2015**, 17, 4-9	4
371	Draft Genome Sequence of Sporidiobolus salmonicolor CBS 6832, a Red-Pigmented Basidiomycetous Yeast. **2015**, 3,	6
370	A protein domain-based view of the virosphere-host relationship. **2015**, 119, 231-43	4
369	Expansion and Functional Divergence of Jumonji C-Containing Histone Demethylases: Significance of Duplications in Ancestral Angiosperms and Vertebrates. **2015**, 168, 1321-37	35
368	The phylogenomics of protein structures: The backstory. **2015**, 119, 284-302	11
367	Overview of computational vaccinology: vaccine development through information technology. **2015**, 56, 381-91	9
366	Nme family of proteins--clues from simple animals. **2015**, 388, 133-42	9
365	Structure-based functional annotation of putative conserved proteins having lyase activity from Haemophilus influenzae. **2015**, 5, 317-336	11
364	Arguments Reinforcing the Three-Domain View of Diversified Cellular Life. **2016**, 2016, 1851865	25
363	The Transcriptome of during Skin Model Infection. **2016**, 6, 136	18
362	Comparing three stochastic search algorithms for computational protein design: Monte Carlo, replica exchange Monte Carlo, and a multistart, steepest-descent heuristic. **2016**, 37, 1781-93	23
361	The crystal structure of maleylacetate reductase from Rhizobium sp. strain MTP-10005 provides insights into the reaction mechanism of enzymes in its original family. **2016**, 84, 1029-42	3
360	The Ensembl gene annotation system. **2016**, 2016,	537
359	An expressed, endogenous Nodavirus-like element captured by a retrotransposon in the genome of the plant parasitic nematode Bursaphelenchus xylophilus. **2016**, 6, 39749	9

358	Evolutionary Histories of Gene Families in Angiosperm Trees. **2016**, 121-137	
357	Two hAT transposon genes were transferred from Brassicaceae to broomrapes and are actively expressed in some recipients. **2016**, 6, 30192	10
356	Analyses of Compact Trichinella Kinomes Reveal a MOS-Like Protein Kinase with a Unique N-Terminal Domain. **2016**, 6, 2847-56	5
355	From Genomes to Phenotypes: Traitar, the Microbial Trait Analyzer. **2016**, 1,	62
354	Protein sequence-similarity search acceleration using a heuristic algorithm with a sensitive matrix. **2016**, 17, 147-154	1
353	MOCAT2: a metagenomic assembly, annotation and profiling framework. **2016**, 32, 2520-3	117
352	Expression profiling and characterization of cold, freezing-related genes from Brassica rapa cultivars. **2016**, 63, 105-115	2
351	DOGMA: domain-based transcriptome and proteome quality assessment. **2016**, 32, 2577-81	28
350	Topics in Grammatical Inference. **2016**,	9
349	Benchmarking the next generation of homology inference tools. **2016**, 32, 2636-41	10
348	Genomic analysis of Ugandan and Rwandan chicken ecotypes using a 600 k genotyping array. **2016**, 17, 407	24
347	Evolution of the Calcium-Based Intracellular Signaling System. **2016**, 8, 2118-32	26
346	Did Viruses Evolve As a Distinct Supergroup from Common Ancestors of Cells?. **2016**, 8, 2474-81	10
345	In vivo protein interaction network analysis reveals porin-localized antibiotic inactivation in Acinetobacter baumannii strain AB5075. **2016**, 7, 13414	57
344	Comparative Protein Structure Modeling Using MODELLER. **2016**, 86, 2.9.1-2.9.37	301
343	Evidence for the sexual origin of heterokaryosis in arbuscular mycorrhizal fungi. **2016**, 1, 16033	89
342	Rabifier2: an improved bioinformatic classifier of Rab GTPases. **2017**, 33, 568-570	3
341	Mapping the membrane proteome of anaerobic gut fungi identifies a wealth of carbohydrate binding proteins and transporters. **2016**, 15, 212	17

340	Distinguishing Hidden Markov Chains. **2016**,		2
339	Learning node labels with multi-category Hopfield networks. **2016**, 27, 1677-1692		5
338	Fra a 1.02 Is the Most Potent Isoform of the Bet´v´1-like Allergen in Strawberry Fruit. **2016**, 64, 3688-96		20
337	The effector candidate repertoire of the arbuscular mycorrhizal fungus Rhizophagus clarus. **2016**, 17, 101		52
336	Hologenome analysis of two marine sponges with different microbiomes. **2016**, 17, 158		40
335	Comparative Protein Structure Modeling Using MODELLER. **2016**, 54, 5.6.1-5.6.37		1267
334	The TIM Barrel Architecture Facilitated the Early Evolution of Protein-Mediated Metabolism. **2016**, 82, 17-26		46
333	Protein Structure Is Related to RNA Structural Reactivity In Vivo. *Journal of Molecular Biology*, **2016**, 428, 758-766	6.5	10
332	Hidden Markov models for gene sequence classification. **2016**, 19, 793-805		4
331	Rapid and enhanced remote homology detection by cascading hidden Markov model searches in sequence space. **2016**, 32, 338-44		3
330	De novo assembly of a tadpole shrimp (Triops newberryi) transcriptome and preliminary differential gene expression analysis. **2017**, 17, 161-171		7
329	Hidden Markov Models for Protein Domain Homology Identification and Analysis. **2017**, 1555, 47-58		3
328	Proteome wide identification of iron binding proteins of Xanthomonas translucens pv. undulosa: focus on secretory virulent proteins. **2017**, 30, 127-141		9
327	Evolutionary genomics of the cold-adapted diatom Fragilariopsis cylindrus. **2017**, 541, 536-540		226
326	Protein Structure Classification and Loop Modeling Using Multiple Ramachandran Distributions. **2017**, 15, 243-254		13
325	Comparison of Leptospira interrogans and Leptospira biflexa genomes: analysis of potential leptospiral-host interactions. **2017**, 13, 883-891		5
324	In silico characterization of hypothetical proteins obtained from Mycobacterium tuberculosis H37Rv. **2017**, 6, 1		3
323	Computational Design of the Tiam1 PDZ Domain and Its Ligand Binding. **2017**, 13, 2271-2289		10

322	Cilia- and Flagella-Associated Protein 69 Regulates Olfactory Transduction Kinetics in Mice. **2017**, 37, 5699-5710	24
321	A generalized mixture model applied to diabetes incidence data. **2017**, 59, 826-842	1
320	Empirical genome evolution models root the tree of life. **2017**, 138, 137-155	18
319	Akaryotes and Eukaryotes are independent descendants of a universal common ancestor. **2017**, 138, 168-183	17
318	Protein Function Prediction. **2017**,	9
317	Crosstalk of two-component signal transduction systems in regulating central carbohydrate and energy metabolism during autotrophic and photomixotrophic growth of Synechocystis sp. PCC 6803. **2017**, 9, 485-496	4
316	A Subset of Ubiquitin-Conjugating Enzymes Is Essential for Plant Immunity. **2017**, 173, 1371-1390	32
315	Floral nectar of the obligate outcrossing Canavalia gladiata (Jacq.) DC. (Fabaceae) contains only one predominant protein, a´class III acidic chitinase. **2017**, 19, 749-759	5
314	The ProFunc Function Prediction Server. **2017**, 1611, 75-95	15
313	Whipworm kinomes reflect a unique biology and adaptation to the host animal. **2017**, 47, 857-866	9
312	The H3 loop of antibodies shows unique structural characteristics. **2017**, 85, 1311-1318	52
311	Functional and biochemical characterization of a T cell-associated anti-apoptotic protein, GIMAP6. **2017**, 292, 9305-9319	7
310	Revised domain structure of ulvan lyase and characterization of the first ulvan binding domain. **2017**, 7, 44115	14
309	Structural and functional insight into pan-endopeptidase inhibition by �-macroglobulins. **2017**, 398, 975-994	13
308	Potential and pitfalls in the genetic diagnosis of kidney diseases. **2017**, 10, 581-585	6
307	Full Protein Sequence Redesign with an MMGBSA Energy Function. **2017**, 13, 4932-4943	8
306	Transcriptional response of the harmful raphidophyte Heterosigma akashiwo to nitrate and phosphate stress. **2017**, 68, 258-270	15
305	Compositional and expression analyses of the glideosome during the life cycle reveal an additional myosin light chain required for maximum motility. **2017**, 292, 17857-17875	26

304	A New Reference Genome Assembly for the Microcrustacean. **2017**, 7, 1405-1416		54
303	Mitochondria are not captive bacteria. **2017**, 434, 88-98		29
302	Comparing pairwise-additive and many-body generalized Born models for acid/base calculations and protein design. **2017**, 38, 2396-2410		16
301	Metabolic and evolutionary patterns in the extremely acidophilic archaeon Ferroplasma acidiphilum Y. **2017**, 7, 3682		11
300	Comparative and Evolutionary Genomics of Angiosperm Trees. **2017**,		5
299	The pomegranate (Punica granatum L.) genome and the genomics of punicalagin biosynthesis. *Plant Journal*, **2017**, 91, 1108-1128	6.9	66
298	AnkPlex: algorithmic structure for refinement of near-native ankyrin-protein docking. **2017**, 18, 220		
297	Simple adjustment of the sequence weight algorithm remarkably enhances PSI-BLAST performance. **2017**, 18, 288		10
296	Recognition of a structural domain (RWDBD) in Gcn1 proteins that interacts with the RWD domain containing proteins. **2017**, 12, 12		1
295	Forecasting residue-residue contact prediction accuracy. **2017**, 33, 3405-3414		3
294	Protein social behavior makes a stronger signal for partner identification than surface geometry. **2017**, 85, 137-154		9
293	Structural and functional annotation of human FAM26F: A multifaceted protein having a critical role in the immune system. **2017**, 597, 66-75		12
292	A whole genome analysis reveals the presence of a plant PR1 sequence in the potato pathogen Streptomyces scabies and other Streptomyces species. **2017**, 114, 346-352		6
291	Family-specific scaling laws in bacterial genomes. **2017**, 45, 7615-7622		3
290	Second-order autoregressive Hidden Markov Model. **2017**, 31,		0
289	Transcription Factor Repertoire of Necrotrophic Fungal Phytopathogen : Predominance of MYB Transcription Factors As Potential Regulators of Secretome. *Frontiers in Plant Science*, **2017**, 8, 1037	6.2	9
288	Autoantibody Repertoire in APECED Patients Targets Two Distinct Subgroups of Proteins. **2017**, 8, 976		27
287	The Enigmatic Origin of Papillomavirus Protein Domains. **2017**, 9,		5

286	Identification of Capsid/Coat Related Protein Folds and Their Utility for Virus Classification. **2017**, 8, 380	17
285	Conserved Transcriptional Responses to Nutrient Stress in Bloom-Forming Algae. **2017**, 8, 1279	20
284	Do Viruses Exchange Genes across Superkingdoms of Life?. **2017**, 8, 2110	13
283	Tum1 is involved in the metabolism of sterol esters in Saccharomyces cerevisiae. **2017**, 17, 181	2
282	Homology modeling and prediction of B-cell and T-cell epitopes of the house dust mite allergen Der´f´20. **2018**, 17, 1807-1812	2
281	Mutations in CFAP43 and CFAP44 cause male infertility and flagellum defects in Trypanosoma and human. **2018**, 9, 686	106
280	Severe Fertility Effects of Sperm Caused by Failure To Enter Female Sperm Storage Organs in. **2018** , 8, 149-160	2
279	Pan-Genome Storage and Analysis Techniques. **2018**, 1704, 29-53	14
278	Phase separation of a yeast prion protein promotes cellular fitness. **2018**, 359,	344
277	Comparative systems analysis of the secretome of the opportunistic pathogen Aspergillus fumigatus and other Aspergillus species. **2018**, 8, 6617	30
276	Organization of feed-forward loop motifs reveals architectural principles in natural and engineered networks. **2018**, 4, eaap9751	20
275	Structure Prediction of a Novel Exo-�-1,3-Glucanase: Insights into the Cold Adaptation of Psychrophilic Yeast Glaciozyma antarctica PI12. **2018**, 10, 157-168	10
274	Gene expression profiles of the Southern house mosquito Culex quinquefasciatus during exposure to permethrin. **2018**, 25, 439-453	12
273	Functional elucidation of hypothetical proteins for their indispensable roles toward drug designing targets from Helicobacter pylori strain HPAG1. **2018**, 36, 906-918	6
272	Semiparametric hidden Markov model with non-parametric regression. **2018**, 47, 5196-5204	2
271	The linear neighborhood propagation method for predicting long non-coding RNA�protein interactions. **2018**, 273, 526-534	104
270	The crystal structure and catalytic mechanism of hydroxynitrile lyase from passion fruit, Passiflora edulis. **2018**, 285, 313-324	10
269	New insights into the evolution and functional divergence of the SWEET family in Saccharum based on comparative genomics. **2018**, 18, 270	21

268	Ancestrality and Mosaicism of Giant Viruses Supporting the Definition of the Fourth TRUC of Microbes. **2018**, 9, 2668		30
267	SFPEL-LPI: Sequence-based feature projection ensemble learning for predicting LncRNA-protein interactions. **2018**, 14, e1006616		91
266	Genome-wide identification of AP2/ERF transcription factors in mungbean (Vigna radiata) and expression profiling of the VrDREB subfamily under drought stress. **2018**, 69, 1009		12
265	Integrating Multifaceted Information to Predict Mycobacterium tuberculosis-Human Protein-Protein Interactions. **2018**, 17, 3810-3823		5
264	Localization Under Topological Uncertainty for Lane Identification of Autonomous Vehicles. **2018**,		1
263	Mayr Versus Woese: Akaryotes and Eukaryotes. **2018**, 13-54		
262	A gene-rich fraction analysis of the Passiflora edulis genome reveals highly conserved microsyntenic regions with two related Malpighiales species. **2018**, 8, 13024		13
261	Genomic Approach to Asthma. **2018**,		1
260	Genome-wide survey of remote homologues for protein domain superfamilies of known structure reveals unequal distribution across structural classes. **2018**, 14, 266-280		0
259	Functional annotation of hypothetical proteins from the Exiguobacterium antarcticum strain B7 reveals proteins involved in adaptation to extreme environments, including high arsenic resistance. **2018**, 13, e0198965		22
258	Mining of potential drug targets through the identification of essential and analogous enzymes in the genomes of pathogens of Glycine max, Zea mays and Solanum lycopersicum. **2018**, 13, e0197511		1
257	Methods for Physical Characterization of Phase-Separated Bodies and Membrane-less Organelles. *Journal of Molecular Biology*, **2018**, 430, 4773-4805	6.5	74
256	Nitrosocaldus cavascurensis, an Ammonia Oxidizing, Extremely Thermophilic Archaeon with a Highly Mobile Genome. **2018**, 9, 28		51
255	Characterizing Phage Genomes for Therapeutic Applications. **2018**, 10,		70
254	Temperature adaptation analysis of a psychrophilic mannanase through structural, functional and molecular dynamics simulation. **2018**, 44, 1270-1277		1
253	De novo profile generation based on sequence context specificity with the long short-term memory network. **2018**, 19, 272		4
252	Homology modeling and epitope prediction of Der f 33. **2018**, 51, e6213		4
251	Improved strategy for the curation and classification of kinases, with broad applicability to other eukaryotic protein groups. **2018**, 8, 6808		8

250	Positive and Negative Regulation of Angiogenesis by Soluble Vascular Endothelial Growth Factor Receptor-1. **2018**, 19,	44
249	The mechanism by which a distinguishing arabinofuranosidase can cope with internal di-substitutions in arabinoxylans. **2018**, 11, 223	15
248	Homology Inference Based on a Reconciliation Approach for the Comparative Genomics of Protozoa. **2018**, 14, 1176934318785138	
247	Comprehensive analysis of Verticillium nonalfalfae in silico secretome uncovers putative effector proteins expressed during hop invasion. **2018**, 13, e0198971	12
246	HMMER Cut-off Threshold Tool (HMMERCTTER): Supervised classification of superfamily protein sequences with a reliable cut-off threshold. **2018**, 13, e0193757	11
245	Biological Databases. **2018**, 303-337	
244	GeneMANIA update 2018. **2018**, 46, W60-W64	309
243	Intelligent Agents: Multi-Agent Systems. **2019**, 315-320	3
242	IgE binding activities and in silico epitope prediction of Der f 32 in Dermatophagoides farinae. **2019** , 213, 46-54	4
241	Nascent Polypeptide Domain Topology and Elongation Rate Direct the Cotranslational Hierarchy of Hsp70 and TRiC/CCT. **2019**, 75, 1117-1130.e5	35
240	High potassium seawater inhibits ascidian sperm chemotaxis, but does not affect the male gamete chemotaxis of a brown alga. **2019**, 27, 225-231	1
239	Short- and Long-Term Effects of UVA on Arabidopsis Are Mediated by a Novel cGMP Phosphodiesterase. **2019**, 29, 2580-2585.e4	16
238	The Hsp70 Chaperone System Stabilizes a Thermo-sensitive Subproteome in E.´coli. **2019**, 28, 1335-1345.e6	20
237	The draft genome of reveals insights into toxin evolution. **2019**, 9, 11314-11328	14
236	The mixed liver and heart transcriptome dataset of the New Zealand brushtail possum. **2019**, 27, 104577	1
235	Structural insights into the molecular mechanisms of pectinolytic enzymes. **2019**, 10, 325-344	2
234	Comparative genomic analysis suggests that the sperm-specific sodium/proton exchanger and soluble adenylyl cyclase are key regulators of CatSper among the Metazoa. **2019**, 5, 25	10
233	Structural Basis of the Subcellular Topology Landscape of. **2019**, 10, 1670	8

232	Protein Abundance Biases the Amino Acid Composition of Disordered Regions to Minimize Non-functional Interactions. *Journal of Molecular Biology*, **2019**, 431, 4978-4992	6.5	11
231	Multivariate Information Fusion With Fast Kernel Learning to Kernel Ridge Regression in Predicting LncRNA-Protein Interactions. **2018**, 9, 716		15
230	Genome structure and evolution of Antirrhinum majus L. **2019**, 5, 174-183		39
229	FipoQ/FBXO33, a Cullin-1-based ubiquitin ligase complex component modulates ubiquitination and solubility of polyglutamine disease protein. **2019**, 149, 781-798		10
228	. **2019**, 7, 13486-13496		38
227	Genetic variant pathogenicity prediction trained using disease-specific clinical sequencing data sets. **2019**, 29, 1144-1151		7
226	Protein Structure-Guided Hidden Markov Models (HMMs) as A Powerful Method in the Detection of Ancestral Endogenous Viral Elements. **2019**, 11,		3
225	Testing Empirical Support for Evolutionary Models that Root the Tree of Life. **2019**, 87, 131-142		4
224	Differential proteostatic regulation of insoluble and abundant proteins. **2019**, 35, 4098-4107		5
223	Comparative Genomics Between and Applied to Identify Mechanisms Involved in Adaptation. **2019**, 10, 187		8
222	PDZ Domains Across the Microbial World: Molecular Link to the Proteases, Stress Response, and Protein Synthesis. **2019**, 11, 644-659		3
221	MultiDomainBenchmark: a multi-domain query and subject database suite. **2019**, 20, 77		1
220	Pathogenicity and Drug Resistance of Human Pathogens. **2019**,		1
219	Rab5 GTPases are required for optimal TORC2 function. **2019**, 218, 961-976		9
218	Repurposing Drugs Based on Evolutionary Relationships Between Targets of Approved Drugs and Proteins of Interest. **2019**, 1903, 45-59		4
217	Comprehensive Transcriptional Profiling of the Gastrointestinal Tract of Ruminants from Birth to Adulthood Reveals Strong Developmental Stage Specific Gene Expression. **2019**, 9, 359-373		6
216	Genome-wide analysis reveals the evolution and structural features of WRINKLED1 in plants. **2019**, 294, 329-341		9
215	A -Variant in a Litter of Shorthaired Standard Dachshunds with X-Linked Hypohidrotic Ectodermal Dysplasia. **2019**, 9, 95-104		5

214	The SUPERFAMILY 2.0 database: a significant proteome update and a new webserver. **2019**, 47, D490-D494	64
213	Genomic resources for the study of echinoderm development and evolution. **2019**, 151, 65-88	3
212	ACCBN: ant-Colony-clustering-based bipartite network method for predicting long non-coding RNA-protein interactions. **2019**, 20, 16	11
211	Towards region-specific propagation of protein functions. **2019**, 35, 1737-1744	3
210	Protein function prediction as a graph-transduction game. **2020**, 134, 96-105	5
209	The SCOP database in 2020: expanded classification of representative family and superfamily domains of known protein structures. **2020**, 48, D376-D382	87
208	Proteome-wide analysis of protein disorder in Triticum aestivum and Hordeum vulgare. **2020**, 84, 107138	0
207	Functional and Structural Analysis of Predicted Proteins Obtained from Minisatellite 33.15-Tagged Transcript pAKT-45 Variants. **2020**, 2020, 2562950	1
206	Integrated Co-functional Network Analysis on the Resistance and Virulence Features in. **2020**, 11, 598380	3
205	An Educational Bioinformatics Project to Improve Genome Annotation. **2020**, 11, 577497	2
204	Screening of Potential Bioremediation Enzymes from Hot Spring Isolates and Characterization of Amylase Protein Sequences of Bacteria and Fungi: An In Silico Study. **2020**, 24, 04020030	
203	In silico Functional Annotation and Characterization of Hypothetical Proteins from FGI94. **2020**, 47, 319-331	3
202	Functional annotation of an ecologically important protein from Chloroflexus aurantiacus involved in polyhydroxyalkanoates (PHA) biosynthetic pathway. **2020**, 2, 1	3
201	Identification of the Sesquiterpene Cyclase Involved in the Biosynthesis of (+)-4-Epi-eremophil-9-en-11-ol Derivatives Isolated from. **2020**, 15, 2775-2782	2
200	Genome-wide identification and expression analysis of SnRK2 gene family in mungbean (Vigna radiata) in response to drought stress. **2020**, 71, 469	6
199	The TERB1-TERB2-MAJIN complex of mouse meiotic telomeres dates back to the common ancestor of metazoans. **2020**, 20, 55	7
198	Effector Repertoire of : In Search of Possible Virulence Factors Responsible for Its Host Specificity. **2020**, 11, 579	3
197	Cell phenotypes as macrostates of the GRN dynamics. **2020**, 334, 213-224	2

196	PredictSuperFam-PSS-3D1D: A server for predicting superfamily for the annotation of twilight zone protein sequences. **2020**, 210, 107479	0
195	An inventory of interactors of the human HSP60/HSP10 chaperonin in the mitochondrial matrix space. **2020**, 25, 407-416	8
194	Bacterial Origin and Reductive Evolution of the CPR Group. **2020**, 12, 103-121	6
193	Antimicrobial Activity of Protein Fraction from Venom Against. **2020**, 25,	14
192	Chimeras of P4-ATPase and Guanylate Cyclase in Pathogenic Protists. **2020**, 36, 382-392	2
191	Methods and Tools for Long Non-Coding RNA Detection and their Application in Systems Medicine. **2021**, 40-51	1
190	Signatures of conserved and unique molecular features in Afrotheria. **2021**, 11, 1011	1
189	Computational structural genomics unravels common folds and predicted functions in the secretome of fungal phytopathogen Magnaporthe oryzae.	4
188	Prediction of neuropeptide precursors and differential expression of adipokinetic hormone/corazonin-related peptide, hugin and corazonin in the brain of malaria vector Nyssorhynchus albimanus during a Plasmodium berghei infection. **2021**, 1, 100014	
187	The N-terminal region of Jaw1 has a role to inhibit the formation of organized smooth endoplasmic reticulum as an intrinsically disordered region. **2021**, 11, 753	1
186	Hydrogen bonds meet self-attention: all you need for general-purpose protein structure embedding.	1
185	TMEM106B in humans and Vac7 and Tag1 in yeast are predicted to be lipid transfer proteins.	
184	Transcriptomic Responses of Four Pelagophytes to Nutrient (N, P) and Light Stress. **2021**, 8,	1
183	XlinkCyNET: A Cytoscape Application for Visualization of Protein Interaction Networks Based on Cross-Linking Mass Spectrometry Identifications. **2021**, 20, 1943-1950	1
182	The tree of life describes a tripartite cellular world. **2021**, 43, e2000343	6
181	SARS-CoV-2 Entry Protein TMPRSS2 and Its Homologue, TMPRSS4 Adopts Structural Fold Similar to Blood Coagulation and Complement Pathway Related Proteins.	2
180	Genomic insights into longan evolution from a chromosome-level genome assembly and population analysis of longan accessions.	0
179	Nucleoporin TPR Affects C2C12 Myogenic Differentiation via Regulation of Expression. **2021**, 10,	0

178	The Chaperonin GroESL Facilitates Caulobacter crescentus Cell Division by Supporting the Functions of the Z-Ring Regulators FtsA and FzlA. **2021**, 12,	1
177	Schistosomiasis Drug Discovery in the Era of Automation and Artificial Intelligence. **2021**, 12, 642383	2
176	DOE JGI Metagenome Workflow. **2021**, 6,	12
175	Searching sequence databases for functional homologs using profile HMMs: how to set bit score thresholds?.	0
174	Genome-wide association study of signature genetic alterations among pseudomonas aeruginosa cystic fibrosis isolates. **2021**, 17, e1009681	2
173	Unraveling the functions of uncharacterized transcription factors in Escherichia coli using ChIP-exo.	
172	Bioinformatic Analysis of the Type VI Secretion System and Effector Prediction. **2021**, 12, 694824	3
171	RNA polymerases in strict endosymbiont bacteria with extreme genome reduction show distinct erosions that might result in limited and differential promoter recognition. **2021**, 16, e0239350	0
170	Nonrefoldability is Pervasive Across the Proteome. **2021**, 143, 11435-11448	9
169	Unraveling the functions of uncharacterized transcription factors in Escherichia coli using ChIP-exo. **2021**, 49, 9696-9710	6
168	TMEM106B in humans and Vac7 and Tag1 in yeast are predicted to be lipid transfer proteins. **2021**,	1
167	Capturing the VirA/VirG TCS of Agrobacterium tumefaciens. **2008**, 631, 161-77	9
166	Predicting Gene Function using Predictive Clustering Trees. **2010**, 365-387	3
165	Protein structure prediction based on sequence similarity. **2009**, 569, 129-56	15
164	A Fuzzy Viterbi Algorithm for Improved Sequence Alignment and Searching of Proteins. **2005**, 11-21	2
163	Domains mediate protein-protein interactions and nucleate protein assemblies. **2008**, 383-405	6
162	Chaperones and Proteases of Plasmodium falciparum. **2014**, 161-187	4
161	Draft genomic and transcriptome resources for marine chelicerate Tachypleus tridentatus. **2019**, 6, 190029	10

160	DeepTFactor: A deep learning-based tool for the prediction of transcription factors. **2021**, 118,	17
159	Modeling methyl-sensitive transcription factor motifs with an expanded epigenetic alphabet.	9
158	Transforming the Language of Life: Transformer Neural Networks for Protein Prediction Tasks.	8
157	RefPlantNLR: a comprehensive collection of experimentally validated plant NLRs.	16
156	Non-Refoldability is Pervasive Across the E. coli Proteome.	2
155	The DOE JGI Metagenome Workflow.	1
154	Identification of potential inhibitory analogs of metastasis tumor antigens (MTAs) using bioactive compounds: revealing therapeutic option to prevent malignancy.	2
153	Genomic and morphologic characterization of a planktonic Thiovulum (Campylobacterota) dominating the surface waters of the sulfidic Movile Cave, Romania.	1
152	The chaperonin GroESL facilitates Caulobacter crescentus cell division by supporting the function of the actin homologue FtsA.	1
151	Structural and functional diversity of the microbial kinome. **2007**, 5, e17	205
150	Improvement in Protein Domain Identification Is Reached by Breaking Consensus, with the Agreement of Many Profiles and Domain Co-occurrence. **2016**, 12, e1005038	21
149	Diversity and distribution of nuclease bacteriocins in bacterial genomes revealed using Hidden Markov Models. **2017**, 13, e1005652	33
148	Analysis of the human kinome using methods including fold recognition reveals two novel kinases. **2008**, 3, e1597	9
147	Evolutionary descent of prion genes from the ZIP family of metal ion transporters. **2009**, 4, e7208	94
146	A statistical model of protein sequence similarity and function similarity reveals overly-specific function predictions. **2009**, 4, e7546	23
145	Computational protein design: validation and possible relevance as a tool for homology searching and fold recognition. **2010**, 5, e10410	16
144	The Vein Patterning 1 (VEP1) gene family laterally spread through an ecological network. **2011**, 6, e22279	15
143	Characterization of the small RNA transcriptome of the diatom, Thalassiosira pseudonana. **2011**, 6, e22870	34

142	Selection in coastal Synechococcus (cyanobacteria) populations evaluated from environmental metagenomes. **2011**, 6, e24249	15
141	Identification of Plasmodium vivax proteins with potential role in invasion using sequence redundancy reduction and profile hidden Markov models. **2011**, 6, e25189	10
140	Respiratory membrane endo-hydrogenase activity in the microaerophile Azorhizobium caulinodans is bidirectional. **2012**, 7, e36744	2
139	Hybrid sequencing approach applied to human fecal metagenomic clone libraries revealed clones with potential biotechnological applications. **2012**, 7, e47654	5
138	Structural phylogenomics reveals gradual evolutionary replacement of abiotic chemistries by protein enzymes in purine metabolism. **2013**, 8, e59300	19
137	Developmental gene discovery in a hemimetabolous insect: de novo assembly and annotation of a transcriptome for the cricket Gryllus bimaculatus. **2013**, 8, e61479	34
136	A method for WD40 repeat detection and secondary structure prediction. **2013**, 8, e65705	40
135	Link clustering reveals structural characteristics and biological contexts in signed molecular networks. **2013**, 8, e67089	7
134	Structural phylogenomics retrodicts the origin of the genetic code and uncovers the evolutionary impact of protein flexibility. **2013**, 8, e72225	47
133	Functional annotation of conserved hypothetical proteins from Haemophilus influenzae Rd KW20. **2013**, 8, e84263	59
132	De novo transcriptomes of a mixotrophic and a heterotrophic ciliate from marine plankton. **2014**, 9, e101418	20
131	Bioinformatic indications that COPI- and clathrin-based transport systems are not present in chloroplasts: an Arabidopsis model. **2014**, 9, e104423	6
130	Acanthamoeba castellanii STAT protein. **2014**, 9, e111345	4
129	Crystal structure of Helicobacter pylori pseudaminic acid biosynthesis N-acetyltransferase PseH: implications for substrate specificity and catalysis. **2015**, 10, e0115634	15
128	Biomimetics: From Bioinformatics to Rational Design of Dendrimers as Gene Carriers. **2015**, 10, e0138392	14
127	Cel5I, a SLH-Containing Glycoside Hydrolase: Characterization and Investigation on Its Role in Ruminiclostridium cellulolyticum. **2016**, 11, e0160812	5
126	Genome-wide identification and characterization of cacao WRKY transcription factors and analysis of their expression in response to witches' broom disease. **2017**, 12, e0187346	13
125	Identification of the S-transferase like superfamily bacillithiol transferases encoded by Bacillus subtilis. **2018**, 13, e0192977	5

124	Isolation and characterization of the full-length cDNA encoding a member of a novel cytochrome p450 family (CYP320A1) from the tropical freshwater snail, Biomphalaria glabrata, intermediate host for Schistosoma mansoni. **2005**, 100, 259-62	8
123	S18 family of mitochondrial ribosomal proteins: evolutionary history and Gly132 polymorphism in colon carcinoma. **2016**, 7, 55649-55662	6
122	Cancer-predicting transcriptomic and epigenetic signatures revealed for ulcerative colitis in patient-derived epithelial organoids. **2018**, 9, 28717-28730	14
121	Prediction of Long Non-Coding RNAs Based on RNA-Seq*. **2013**, 39, 1156-1166	3
120	Application of Fireworks Algorithm in Bioinformatics. **2020**, 233-262	1
119	PASS2. **2011**, 2, 53-66	3
118	Functional Prediction of Hypothetical Proteins from Shigella flexneri and Validation of the Predicted Models by Using ROC Curve Analysis. **2018**, 16, e26	5
117	Analysis of osmotin, a PR protein as metabolic modulator in plants. **2011**, 5, 336-40	35
116	Homology modeling and assigned functional annotation of an uncharacterized antitoxin protein from Streptomyces xinghaiensis. **2015**, 11, 493-500	2
115	Fold combinations in multi-domain proteins. **2019**, 15, 342-350	3
114	Ankyrin domains across the Tree of Life. **2014**, 2, e264	55
113	What is an archaeon and are the Archaea really unique?. **2018**, 6, e5770	6
112	Tandem-repeat protein domains across the tree of life. **2015**, 3, e732	37
111	In silico functional prediction of hypothetical proteins from the core genome of biovar. **2020**, 8, e9643	5
110	RefPlantNLR is a comprehensive collection of experimentally validated plant disease resistance proteins from the NLR family. **2021**, 19, e3001124	7
109	Proteins from the core genome of Corynebacterium ulcerans respond for pathogenicity and reveal promising vaccine targets for diphtheria. **2021**, 161, 105263	1
108	Protein Sequence Database Methods. **2004**, 13-17	
107	Development of Integrated System for Motif and Domain Search. **2004**, 14, 991-996	2

106	Gene and Protein Sequence Databases. **2007**, 349-372	
105	Relational Sequence Learning. **2008**, 28-55	6
104	The Bioverse API and web application. **2009**, 541, 511-34	2
103	A bioinformatics insight to rhizobial globins: gene identification and mapping, polypeptide sequence and phenetic analysis, and protein modeling. **2015**, 4, 117	1
102	Learning the Language of Biological Sequences. **2016**, 215-247	3
101	From genomes to phenotypes: Traitar, the microbial trait analyzer.	1
100	Did viruses evolve as a distinct supergroup from common ancestors of cells?.	0
99	Simple adjustment of the sequence weight algorithm remarkably enhances PSI-BLAST performance.	
98	HMMER Cut-off Threshold Tool (HMMERCTTER): Supervised Classification of Superfamily Protein Sequences with a reliable Cut-off Threshold.	0
97	PDZ Domains in Microorganisms: Link Between Stress Response and Protein Synthesis.	0
96	Comparative systems analysis of the secretome of the opportunistic pathogen Aspergillus fumigatus and other Aspergillus species.	
95	CandidatusNitrosocaldus cavascurensis, an ammonia oxidizing, extremely thermophilic archaeon with a highly mobile genome.	
94	De novo profile generation based on sequence context specificity with the long short-term memory network.	
93	Comprehensive analysis of Verticillium nonalfalfae in silico secretome uncovers putative effector proteins expressed during hop invasion.	
92	What is an archaeon and are the Archaea really unique?.	
91	Towards region-specific propagation of protein functions.	
90	Genetic variant pathogenicity prediction trained using disease-specific clinical sequencing datasets.	
89	Comprehensive transcriptional profiling of the gastrointestinal tract of ruminants from birth to adulthood reveals strong developmental stage specific gene expression.	1

88	A de novo EDA-variant in a litter of shorthaired standard Dachshunds with X-linked hypohidrotic ectodermal dysplasia.	
87	Discovery of Novel Drug Targets in Microbial Pathogens Among Hypothetical Proteins: Methods and Significance. **2019**, 377-391	
86	Genomic insights into the Archaea inhabiting an Australian radioactive legacy site.	
85	The TERB1-TERB2-MAJIN complex of mouse meiotic telomeres dates back to the common ancestor of metazoans.	
84	A physics-based energy function allows the computational redesign of a PDZ domain.	1
83	Effect of Bud3 phosphorylation site mutations on bud site selection and cell cycle progression in Saccharomyces cerevisiae. **2019**, 90, 59	
82	The Genome3D Consortium for Structural Annotations of Selected Model Organisms. **2020**, 2165, 27-67	1
81	XlinkCyNET: a Cytoscape application for visualization of protein interaction networks based on cross-linking mass-spectrometry identifications.	
80	Homology modeling and functional characterization of multidrug effluxor Mta protein from Bacillus Atrophaeus: An explanatory insilico approach.	
79	Genomic Insights Into the Inhabiting an Australian Radioactive Legacy Site. **2021**, 12, 732575	2
78	Identification of potential inhibitory analogs of metastasis tumor antigens (MTAs) using bioactive compounds: revealing therapeutic option to prevent malignancy. **2021**, 1	
77	Protein Structure Prediction. **2009**, 63-78	
76	In Silico Analysis of Glutaminase from Different Species of Escherichia and Bacillus. **2016**, 41, 406-14	8
75	Systematic Interrogation of Protein Refolding Under Cellular-Like Conditions.	2
74	Identification and functional annotation of hypothetical proteins of uropathogenic strain CFT073 towards designing antimicrobial drug targets. **2021**, 1-12	2
73	Reannotation of Mycoplasma hyopneumoniae hypothetical proteins revealed novel potential virulence factors. **2021**, 162, 105344	0
72	Sequence Divergence and Functional Specializations of the Ancient Spliceosomal SF3b: Implications in Flexibility and Adaptations of the Multi-Protein Complex.. **2021**, 12, 747344	0
71	Land-use type temporarily affects active pond community structure but not gene expression patterns.. **2022**,	1

70	ScrepYard: an online resource for disulfide-stabilised tandem repeat peptides.	0
69	Systematic analysis of HD-ZIP transcription factors in sesame genome and gene expression profiling of SiHD-ZIP class I entailing drought stress responses at early seedling stage.. **2022**, 49, 2059	0
68	Resources for computational prediction of intrinsic disorder in proteins.. **2022**,	1
67	Investigating the Functional Role of Hypothetical Proteins From an Antarctic Bacterium sp. Lz4W: Emphasis on Identifying Proteins Involved in Cold Adaptation.. **2022**, 13, 825269	
66	Quantifying and Cataloguing Unknown Sequences within Human Microbiomes.. **2022**, e0146821	0
65	Hypothetical protein predicted to be tumor suppressor: a protein functional analysis.. **2022**, 20, e6	0
64	Seqrutinator: Non-Functional Homologue Sequence Scrutiny for the Generation of large Datatsets for Protein Superfamily Analysis.	0
63	Phylogenetic study to analyse the evolutionary relationship of taxonomically diverse □-amylases. 1	
62	Hydrogen bonds meet self-attention: all you need for protein structure embedding. **2021**,	0
61	Regulation of KHNYN antiviral activity by the extended di-KH domain and nucleo-cytoplasmic trafficking.	
60	Characterization of a Secretory YML079-like Cupin Protein That Contributes to Pathogenicity.. **2021** , 9,	0
59	Image_1.TIF. **2019**,	
58	Image_2.TIF. **2019**,	
57	Table_1.XLSX. **2019**,	
56	Image_1.tif. **2020**,	
55	Table_1.xlsx. **2020**,	
54	Table_2.xlsx. **2020**,	
53	Table_3.xlsx. **2020**,	

52	Table_4.xlsx. **2020**,		
51	Table_5.xlsx. **2020**,		
50	Table_6.xlsx. **2020**,		
49	Data_Sheet_1.docx. **2018**,		
48	Table_1.xlsx. **2018**,		
47	Table_2.xlsx. **2018**,		
46	Table_3.xlsx. **2018**,		
45	Data_Sheet_1.docx. **2020**,		
44	Image_1.TIFF. **2018**,		
43	Image_2.PDF. **2018**,		
42	Table_1.XLSX. **2018**,		
41	Data_Sheet_1.docx. **2019**,		
40	Table_2.XLSX. **2019**,		
39	Data_Sheet_1.pdf. **2020**,		
38	The Plant Invertase/Pectin Methylesterase Inhibitor Superfamily.. *Frontiers in Plant Science*, **2022**, 13, 863892	6.2	0
37	In Silico Identification and Characterization of a Hypothetical Protein From Revealing S-Adenosylmethionine-Dependent Methyltransferase Activity.. *Bioinformatics and Biology Insights*, **2022**, 16, 11779322221094236	5.3	
36	The limitations of the current protein classification tools in identifying lipolytic features in putative bacterial lipase sequences.. *Journal of Biotechnology*, **2022**, 351, 30-37	3.7	0
35	Fungal Effector Proteins: Molecular Mediators of Fungal Symbionts of Plants. *Rhizosphere Biology*, **2022**, 297-321	0.8	

34	Chromosome-level genome and the identification of sex chromosomes in Uloborus diversus.		1
33	Genome-wide subcellular protein localisation in the flagellate parasite Trypanosoma brucei.		0
32	A fully phased interspecific grapevine rootstock genome sequence representing V. riparia and V. cinerea and allele-aware annotation of the phylloxera resistance locus Rdv1.		0
31	Spruce giga-genomes: structurally similar yet distinctive with differentially expanding gene families and rapidly evolving genes. *Plant Journal*,	6.9	0
30	Establishing the Taxa, Functional profile, and in-silico Ayurvedic Remedy of Microbiota implicated in West Nile Fever.		
29	Transformer Neural Networks for Protein Family and Interaction Prediction Tasks.		
28	Climate-Endangered Arctic Epishelf Lake Harbors Viral Assemblages with Distinct Genetic Repertoires.		1
27	Von Hippel□Lindau (VHL) disease and VHL-associated tumors in Indian subjects: VHL gene testing in a resource constraint setting. **2022**, 23,		
26	EventPointer 3.0: flexible and accurate splicing analysis that includes studying the differential usage of protein-domains. **2022**, 4,		0
25	In silico prediction of the animal susceptibility and virtual screening of natural compounds against SARS-CoV-2: Molecular dynamics simulation based analysis. 13,		0
24	Organizing the bacterial annotation space with amino acid sequence embeddings. **2022**, 23,		0
23	Maternal body condition and season influence RNA deposition in the oocytes of alfalfa leafcutting bees (Megachile rotundata).		0
22	In silico functional annotation of hypothetical proteins from the Bacillus paralicheniformis strain Bac84 reveals proteins with biotechnological potentials and adaptational functions to extreme environments. **2022**, 17, e0276085		0
21	SulfAtlas, the sulfatase database: state of the art and new developments.		1
20	Recombinant expression and characterisation of a lipase from the Antarctic zooplankton Salpa thompsoni.		0
19	Genomic basis of the giga-chromosomes and giga-genome of tree peony Paeonia ostii. **2022**, 13,		0
18	An improved reference of the grapevine genome supports reasserting the origin of the PN40024 highly-homozygous genotype.		1
17	Maternal body condition and season influence RNA deposition in the oocytes of alfalfa leafcutting bees (Megachile rotundata). 13,		0

16	Tracing the birth of structural domains from loops during protein evolution.	0
15	ScrepYard : an online resource for disulfide-stabilised tandem repeat peptides.	0
14	Chromosome-level genome and the identification of sex chromosomes in Uloborus diversus. **2022**, 12,	0
13	Addressing the pervasive scarcity of structural annotation in eukaryotic algae. **2023**, 13,	1
12	A novel chemogenomic discovery platform identifies bioactive hits with rapid bactericidal activity against Mycobacteroides Abscessus. **2023**, 139, 102317	0
11	The genome of Lignosus tigris: Uncovering its hidden nutraceutical potential. **2023**, 154, 108-119	1
10	Genomic Survey of Flavin Monooxygenases in Wild and Cultivated Rice Provides Insight into Evolution and Functional Diversities. **2023**, 24, 4190	0
9	Genome-wide subcellular protein map for the flagellate parasite Trypanosoma brucei. **2023**, 8, 533-547	0
8	Evolution is not Uniform Along Coding Sequences. **2023**, 40,	0
7	In silico Analysis of Diverse Endo-□-1,4-glucanases Reveals Their Molecular Evolution. **2023**, 59, 94-106	0
6	Homology Modeling in the Twilight Zone: Improved Accuracy by Sequence Space Analysis. **2023**, 1-23	0
5	An improved reference of the grapevine genome reasserts the origin of the PN40024 highly homozygous genotype.	0
4	Single-cell long-read mRNA isoform regulation is pervasive across mammalian brain regions, cell types, and development.	0
3	Targeting Essential Hypothetical Proteins of Pseudomonas aeruginosa PAO1 for Mining of Novel Therapeutics: An In Silico Approach. **2023**, 2023, 1-28	0
2	The dcGO domain-centric ontology database in 2023: new website and extended annotations for protein structural domains. **2023**, 168093	0
1	Annelid adult cell type diversity and their pluripotent cellular origins.	0