Erik Ll L Sonnhammer

List of Publications by Citations

141 papers	**44,279** citations	**53** h-index	**159** g-index
159 ext. papers	**53,058** ext. citations	**9.3** avg, IF	**7.39** L-index

#	Paper	IF	Citations
141	Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology*, **2001**, 305, 567-80	6.5	8822
140	Pfam: the protein families database. *Nucleic Acids Research*, **2014**, 42, D222-30	20.1	3975
139	The Pfam protein families database. *Nucleic Acids Research*, **2012**, 40, D290-301	20.1	2844
138	The Pfam protein families database. *Nucleic Acids Research*, **2004**, 32, D138-41	20.1	2720
137	The Pfam protein families database. *Nucleic Acids Research*, **2010**, 38, D211-22	20.1	2467
136	The Pfam protein families database in 2019. *Nucleic Acids Research*, **2019**, 47, D427-D432	20.1	2298
135	The Pfam protein families database. *Nucleic Acids Research*, **2002**, 30, 276-80	20.1	1839
134	The Pfam protein families database. *Nucleic Acids Research*, **2008**, 36, D281-8	20.1	1806
133	Pfam: clans, web tools and services. *Nucleic Acids Research*, **2006**, 34, D247-51	20.1	1784
132	A combined transmembrane topology and signal peptide prediction method. *Journal of Molecular Biology*, **2004**, 338, 1027-36	6.5	1669
131	The Pfam protein families database. *Nucleic Acids Research*, **2000**, 28, 263-6	20.1	1074
130	Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. *Nucleic Acids Research*, **2007**, 35, W429-32	20.1	1033
129	Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *Journal of Molecular Biology*, **2001**, 314, 1041-52	6.5	959
128	Pfam: a comprehensive database of protein domain families based on seed alignments. *Proteins: Structure, Function and Bioinformatics*, **1997**, 28, 405-20	4.2	841
127	Pfam: The protein families database in 2021. *Nucleic Acids Research*, **2021**, 49, D412-D419	20.1	620
126	A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene*, **1995**, 167, GC1-10	3.8	563
125	Inparanoid: a comprehensive database of eukaryotic orthologs. *Nucleic Acids Research*, **2005**, 33, D476-80	20.1	541

124	Pfam: multiple sequence alignments and HMM-profiles of protein domains. *Nucleic Acids Research*, **1998**, 26, 320-2	20.1	527
123	InParanoid 7: new algorithms and tools for eukaryotic orthology analysis. *Nucleic Acids Research*, **2010**, 38, D196-203	20.1	500
122	Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins. *Nucleic Acids Research*, **1999**, 27, 260-2	20.1	488
121	Kalign--an accurate and fast multiple sequence alignment algorithm. *BMC Bioinformatics*, **2005**, 6, 298	3.6	476
120	InParanoid 8: orthology analysis between 273 proteomes, mostly eukaryotic. *Nucleic Acids Research*, **2015**, 43, D234-9	20.1	337
119	Orthology, paralogy and proposed classification for paralog subtypes. *Trends in Genetics*, **2002**, 18, 619-20	11.5	299
118	Sequence of the human immunoglobulin diversity (D) segment locus: a systematic analysis provides no evidence for the use of DIR segments, inverted D segments, "minor" D segments or D-D recombination. *Journal of Molecular Biology*, **1997**, 270, 587-97	6.5	265
117	An HMM posterior decoder for sequence feature prediction that includes homology information. *Bioinformatics*, **2005**, 21 Suppl 1, i251-7	7.2	246
116	Genomic gene clustering analysis of pathways in eukaryotes. *Genome Research*, **2003**, 13, 875-82	9.7	243
115	FAT: a novel domain in PIK-related kinases. *Trends in Biochemical Sciences*, **2000**, 25, 225-7	10.3	243
114	Kalign2: high-performance multiple alignment of protein and nucleotide sequences allowing external features. *Nucleic Acids Research*, **2009**, 37, 858-65	20.1	214
113	Automatic clustering of orthologs and inparalogs shared by multiple proteomes. *Bioinformatics*, **2006**, 22, e9-15	7.2	200
112	Spatial maps of prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity. *Nature Communications*, **2018**, 9, 2419	17.4	195
111	The imprint of somatic hypermutation on the repertoire of human germline V genes. *Journal of Molecular Biology*, **1996**, 256, 813-17	6.5	178
110	InParanoid 6: eukaryotic ortholog clusters with inparalogs. *Nucleic Acids Research*, **2008**, 36, D263-6	20.1	176
109	Modular arrangement of proteins as inferred from analysis of homology. *Protein Science*, **1994**, 3, 482-92	6.3	174
108	Membrane topology of the Drosophila OR83b odorant receptor. *FEBS Letters*, **2007**, 581, 5601-4	3.8	165
107	Automated ortholog inference from phylogenetic trees and calculation of orthology reliability. *Bioinformatics*, **2002**, 18, 92-9	7.2	148

106	A general model of G protein-coupled receptor sequences and its application to detect remote homologs. *Protein Science*, **2006**, 15, 509-21	6.3	139
105	Volume changes in protein evolution. *Journal of Molecular Biology*, **1994**, 236, 1067-78	6.5	139
104	Standardized benchmarking in the quest for orthologs. *Nature Methods*, **2016**, 13, 425-30	21.6	133
103	Global networks of functional coupling in eukaryotes from comprehensive data integration. *Genome Research*, **2009**, 19, 1107-16	9.7	123
102	Improved and automated prediction of effective siRNA. *Biochemical and Biophysical Research Communications*, **2004**, 319, 264-74	3.4	115
101	Analysis of protein domain families in Caenorhabditis elegans. *Genomics*, **1997**, 46, 200-16	4.3	106
100	Automatic assessment of alignment quality. *Nucleic Acids Research*, **2005**, 33, 7120-8	20.1	102
99	Comprehensive sequence analysis of the 182 predicted open reading frames of yeast chromosome III. *Protein Science*, **1992**, 1, 1677-90	6.3	98
98	Quality assessment of multiple alignment programs. *FEBS Letters*, **2002**, 529, 126-30	3.8	95
97	FunCoup 3.0: database of genome-wide functional coupling networks. *Nucleic Acids Research*, **2014**, 42, D380-8	20.1	89
96	Scoredist: a simple and robust protein sequence distance estimator. *BMC Bioinformatics*, **2005**, 6, 108	3.6	89
95	Kalign, Kalignvu and Mumsa: web servers for multiple sequence alignment. *Nucleic Acids Research*, **2006**, 34, W596-9	20.1	79
94	Domain tree-based analysis of protein architecture evolution. *Molecular Biology and Evolution*, **2008**, 25, 254-64	8.3	70
93	A comprehensive structural, biochemical and biological profiling of the human NUDIX hydrolase family. *Nature Communications*, **2017**, 8, 1541	17.4	62
92	Predicting protein function from domain content. *Bioinformatics*, **2008**, 24, 1681-7	7.2	60
91	Reliability of transmembrane predictions in whole-genome data. *FEBS Letters*, **2002**, 532, 415-8	3.8	58
90	siRNAdb: a database of siRNA sequences. *Nucleic Acids Research*, **2005**, 33, D131-4	20.1	51
89	OrthoDisease: a database of human disease orthologs. *Human Mutation*, **2004**, 24, 112-9	4.7	48

88   Network-based Identification of novel cancer genes. *Molecular and Cellular Proteomics*, **2010**, 9, 648-55   7.6   46

87   Letter to the editor: SeqXML and OrthoXML: standards for sequence and orthology information. *Briefings in Bioinformatics*, **2011**, 12, 485-8   13.4   44

86   A workbench for large-scale sequence homology analysis. *Bioinformatics*, **1994**, 10, 301-7   7.2   44

85   What@ in a genome?. *Nature*, **1992**, 358, 287   50.4   44

84   A novel transmembrane topology of presenilin based on reconciling experimental and computational evidence. *FEBS Journal*, **2005**, 272, 2727-33   5.7   43

83   Prognostic significance in breast cancer of a gene signature capturing stromal PDGF signaling. *American Journal of Pathology*, **2013**, 182, 2037-47   5.8   42

82   Comparative interactomics with Funcoup 2.0. *Nucleic Acids Research*, **2012**, 40, D821-8   20.1   41

81   Dynamic zebrafish interactome reveals transcriptional mechanisms of dioxin toxicity. *PLoS ONE*, **2010**, 5, e10465   3.7   41

80   Improved profile HMM performance by assessment of critical algorithmic features in SAM and HMMER. *BMC Bioinformatics*, **2005**, 6, 99   3.6   40

79   A Novel ACKR2-Dependent Role of Fibroblast-Derived CXCL14 in Epithelial-to-Mesenchymal Transition and Metastasis of Breast Cancer. *Clinical Cancer Research*, **2019**, 25, 3702-3717   12.9   37

78   FunShift: a database of function shift analysis on protein subfamilies. *Nucleic Acids Research*, **2005**, 33, D197-200   20.1   37

77   Computational antisense oligo prediction with a neural network model. *Bioinformatics*, **2002**, 18, 1567-75   5.2   37

76   Evolution of protein domain architectures. *Methods in Molecular Biology*, **2012**, 856, 187-216   1.4   35

75   Comprehensive analysis of orthologous protein domains using the HOPS database. *Genome Research*, **2003**, 13, 2353-62   9.7   34

74   Integrated graphical analysis of protein sequence features predicted from sequence composition. *Proteins: Structure, Function and Bioinformatics*, **2001**, 45, 262-73   4.2   34

73   Domain architecture conservation in orthologs. *BMC Bioinformatics*, **2011**, 12, 326   3.6   32

72   Hieranoid: hierarchical orthology inference. *Journal of Molecular Biology*, **2013**, 425, 2072-2081   6.5   31

71   A large-scale benchmark of gene prioritization methods. *Scientific Reports*, **2017**, 7, 46598   4.9   30

70	Functional association networks as priors for gene regulatory network inference. *Bioinformatics*, **2014**, 30, i130-8	7.2	30
69	FunCoup 4: new species, data, and visualization. *Nucleic Acids Research*, **2018**, 46, D601-D607	20.1	29
68	Statistical assessment of crosstalk enrichment between gene groups in biological networks. *PLoS ONE*, **2013**, 8, e54945	3.7	28
67	Large-scale prediction of function shift in protein families with a focus on enzymatic function. *Proteins: Structure, Function and Bioinformatics*, **2005**, 60, 758-68	4.2	26
66	Classification of transmembrane protein families in the Caenorhabditis elegans genome and identification of human orthologs. *Genome Research*, **2000**, 10, 1679-89	9.7	26
65	Orthology confers intron position conservation. *BMC Genomics*, **2010**, 11, 412	4.5	25
64	Gearing up to handle the mosaic nature of life in the quest for orthologs. *Bioinformatics*, **2018**, 34, 323-329	7.2	25
63	Improving profile HMM discrimination by adapting transition probabilities. *Journal of Molecular Biology*, **2004**, 338, 847-54	6.5	24
62	PathWAX: a web server for network crosstalk based pathway annotation. *Nucleic Acids Research*, **2016**, 44, W105-9	20.1	22
61	MaxLink: network-based prioritization of genes tightly linked to a disease seed set. *Bioinformatics*, **2014**, 30, 2689-90	7.2	22
60	Comparative analysis and unification of domain-domain interaction networks. *Bioinformatics*, **2009**, 25, 3020-5	7.2	22
59	MetaTM - a consensus method for transmembrane protein topology prediction. *BMC Bioinformatics* , **2009**, 10, 314	3.6	22
58	Profiled support vector machines for antisense oligonucleotide efficacy prediction. *BMC Bioinformatics*, **2004**, 5, 135	3.6	21
57	A novel method for crosstalk analysis of biological networks: improving accuracy of pathway annotation. *Nucleic Acids Research*, **2017**, 45, e8	20.1	20
56	Quality criteria for finding genes with high mRNA-protein expression correlation and coexpression correlation. *Gene*, **2012**, 497, 228-36	3.8	18
55	The Quest for Orthologs benchmark service and consensus calls in 2020. *Nucleic Acids Research*, **2020**, 48, W538-W545	20.1	17
54	A genome-wide IR-induced RAD51 foci RNAi screen identifies CDC73 involved in chromatin remodeling for DNA repair. *Cell Discovery*, **2015**, 1, 15034	22.3	17
53	Overview and comparison of ortholog databases. *Drug Discovery Today: Technologies*, **2006**, 3, 137-43	7.1	17

52	jSquid: a Java applet for graphical on-line network exploration. *Bioinformatics*, **2008**, 24, 1467-8	7.2	16
51	siRNA specificity searching incorporating mismatch tolerance data. *Bioinformatics*, **2008**, 24, 1316-7	7.2	15
50	PfamAlyzer: domain-centric homology search. *Bioinformatics*, **2007**, 23, 3382-3	7.2	14
49	Functional characterization in Caenorhabditis elegans of transmembrane worm-human orthologs. *BMC Genomics*, **2004**, 5, 85	4.5	14
48	Assessment of protein distance measures and tree-building methods for phylogenetic tree reconstruction. *Molecular Biology and Evolution*, **2005**, 22, 2257-64	8.3	13
47	The Chironomus tentans genome sequence and the organization of the Balbiani ring genes. *BMC Genomics*, **2014**, 15, 819	4.5	12
46	OrthoDisease: tracking disease gene orthologs across 100 species. *Briefings in Bioinformatics*, **2011**, 12, 463-73	13.4	12
45	Focusing on RISC assembly in mammalian cells. *Biochemical and Biophysical Research Communications*, **2008**, 368, 703-8	3.4	12
44	OrthoGUI: graphical presentation of Orthostrapper results. *Bioinformatics*, **2002**, 18, 1272-3	7.2	12
43	Discovering viral genomes in human metagenomic data by predicting unknown protein families. *Scientific Reports*, **2018**, 8, 28	4.9	11
42	Evolution of Protein Domain Architectures. *Methods in Molecular Biology*, **2019**, 1910, 469-504	1.4	11
41	Improved orthology inference with Hieranoid 2. *Bioinformatics*, **2017**, 33, 1154-1159	7.2	11
40	Optimal sparsity criteria for network inference. *Journal of Computational Biology*, **2013**, 20, 398-408	1.7	11
39	Benchmarking homology detection procedures with low complexity filters. *Bioinformatics*, **2009**, 25, 2500-5	7.2	11
38	Dynamic contact maps of protein structures. *Journal of Molecular Graphics and Modelling*, **1998**, 16, 1-5, 33	2.8	10
37	Prediction of function divergence in protein families using the substitution rate variation parameter alpha. *Molecular Biology and Evolution*, **2006**, 23, 1406-13	8.3	10
36	Conspicuous accumulation of transcription elongation repressor hrp130/CA150 on the intron-rich Balbiani ring 3 gene. *Chromosoma*, **2004**, 113, 244-57	2.8	10
35	Benchmarking the next generation of homology inference tools. *Bioinformatics*, **2016**, 32, 2636-41	7.2	10

34	GeneSPIDER - gene regulatory network inference benchmarking with controlled network and data properties. *Molecular BioSystems*, **2017**, 13, 1304-1312		9
33	HieranoiDB: a database of orthologs inferred by Hieranoid. *Nucleic Acids Research*, **2017**, 45, D687-D690	20.1	9
32	Avoiding pitfalls in L1-regularised inference of gene networks. *Molecular BioSystems*, **2015**, 11, 287-96		8
31	MGclus: network clustering employing shared neighbors. *Molecular BioSystems*, **2013**, 9, 1670-5		8
30	DASher: a stand-alone protein sequence client for DAS, the Distributed Annotation System. *Bioinformatics*, **2009**, 25, 1333-4	7.2	8
29	ChromoWheel: a new spin on eukaryotic chromosome visualization. *Bioinformatics*, **2004**, 20, 576-7	7.2	7
28	Fusion transcript detection using spatial transcriptomics. *BMC Medical Genomics*, **2020**, 13, 110	3.7	7
27	Avoiding pitfalls in gene (co)expression meta-analysis. *Genomics*, **2014**, 103, 21-30	4.3	6
26	Chromosomal clustering of nuclear genes encoding mitochondrial and chloroplast proteins in Arabidopsis. *Trends in Genetics*, **2006**, 22, 589-93	8.5	6
25	Widespread eukaryotic sequences, highly similar to bacterial DNA polymerase I, looking for functions. *Current Biology*, **1997**, 7, R463-5	6.3	5
24	Employing conservation of co-expression to improve functional inference. *BMC Systems Biology*, **2008**, 2, 81	3.5	5
23	Putative regulatory domains in the human and mouse CVADR genes. *Gene Function & Disease*, **2000**, 1, 82-86		5
22	A generalized framework for controlling FDR in gene regulatory network inference. *Bioinformatics*, **2019**, 35, 1026-1032	7.2	5
21	Genome-wide functional association networks: background, data & state-of-the-art resources. *Briefings in Bioinformatics*, **2020**, 21, 1224-1237	13.4	5
20	DNA Methylation Levels in Mononuclear Leukocytes from the Mother and Her Child Are Associated with IgE Sensitization to Allergens in Early Life. *International Journal of Molecular Sciences*, **2021**, 22,	6.3	5
19	Comparison of the human germline and rearranged VH repertoire reveals complementarity between germline variability and somatic mutation. *Annals of the New York Academy of Sciences*, **1995**, 764, 180-2	6.5	4
18	Automatic extraction of reliable regions from multiple sequence alignments. *BMC Bioinformatics*, **2007**, 8 Suppl 5, S9	3.6	4
17	Sfixem--graphical sequence feature display in Java. *Bioinformatics*, **2004**, 20, 2488-90	7.2	4

LIST OF PUBLICATIONS

#	Publication	IF	Cit
16	Pathway-specific model estimation for improved pathway annotation by network crosstalk. *Scientific Reports*, **2020**, 10, 13585	4.9	4
15	FunCoup 5: Functional Association Networks in All Domains of Life, Supporting Directed Links and Tissue-Specificity. *Journal of Molecular Biology*, **2021**, 433, 166835	6.5	4
14	Domainoid: domain-oriented orthology inference. *BMC Bioinformatics*, **2019**, 20, 523	3.6	3
13	transition priors for protein hidden Markov models: an empirical study towards maximum discrimination. *Journal of Computational Biology*, **2004**, 11, 181-93	1.7	3
12	The C. elegans expression pattern database: a beginning **1996**, 12, 370-370		3
11	TreeDom: a graphical web tool for analysing domain architecture evolution. *Bioinformatics*, **2016**, 32, 2384-5	7.2	2
10	Network analysis of functional genomics data: application to avian sex-biased gene expression. *Scientific World Journal, The*, **2012**, 2012, 130491	2.2	2
9	NovelFam3000--uncharacterized human protein domains conserved across model organisms. *BMC Genomics*, **2006**, 7, 48	4.5	2
8	Exploring the foundation of genomics: a northern blot reference set for the comparative analysis of transcript profiling technologies. *Comparative and Functional Genomics*, **2004**, 5, 584-95		2
7	Uncovering cancer gene regulation by accurate regulatory network inference from uninformative data. *Npj Systems Biology and Applications*, **2020**, 6, 37	5	1
6	Experimental validation of predicted cancer genes using FRET. *Methods and Applications in Fluorescence*, **2018**, 6, 035007	3.1	1
5	Drug repurposing improves disease targeting 11-fold and can be augmented by network module targeting, applied to COVID-19. *Scientific Reports*, **2021**, 11, 20687	4.9	1
4	MetaCNV - a consensus approach to infer accurate copy numbers from low coverage data. *BMC Medical Genomics*, **2020**, 13, 76	3.7	1
3	Perturbation-based gene regulatory network inference to unravel oncogenic mechanisms. *Scientific Reports*, **2020**, 10, 14149	4.9	0
2	Generation of Realistic Gene Regulatory Networks by Enriching for Feed-Forward Loops.. *Frontiers in Genetics*, **2022**, 13, 815692	4.5	
1	Network Crosstalk as a Basis for Drug Repurposing.. *Frontiers in Genetics*, **2022**, 13, 792090	4.5	