Anders Krogh

List of Publications by Citations

140 papers	35,400 citations	61 h-index	150 g-index
150 ext. papers	41,413 ext. citations	11.7 avg, IF	6.92 L-index

#	Paper	IF	Citations
140	Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology*, **2001**, 305, 567-80	6.5	8822
139	Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature*, **1998**, 393, 537-44	50.4	6341
138	Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids **1998**,		2271
137	A combined transmembrane topology and signal peptide prediction method. *Journal of Molecular Biology*, **2004**, 338, 1027-36	6.5	1669
136	Hidden Markov models in computational biology. Applications to protein modeling. *Journal of Molecular Biology*, **1994**, 235, 1501-31	6.5	1331
135	Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. *Nucleic Acids Research*, **2007**, 35, W429-32	20.1	1033
134	Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18. *Nature*, **2001**, 413, 848-52	50.4	1010
133	Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Science*, **2003**, 12, 1652-62	6.3	880
132	Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells. *Journal of Biological Chemistry*, **2008**, 283, 1026-33	5.4	871
131	Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature Communications*, **2016**, 7, 11257	17.4	677
130	Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature*, **2010**, 463, 757-62	50.4	567
129	Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature*, **2013**, 499, 74-8	50.4	563
128	JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic Acids Research*, **2008**, 36, D102-6	20.1	541
127	An Aboriginal Australian genome reveals separate human dispersals into Asia. *Science*, **2011**, 334, 94-8	33.3	528
126	Whole-genome sequencing of cultivated and wild peppers provides insights into Capsicum domestication and specialization. *Proceedings of the National Academy of Sciences of the United States of America*, **2014**, 111, 5135-40	11.5	466
125	The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. *Nature Genetics*, **2009**, 41, 553-62	36.3	356
124	microRNA-101 is a potent inhibitor of autophagy. *EMBO Journal*, **2011**, 30, 4628-41	13	275

123	Prediction of human protein function from post-translational modifications and localization features. *Journal of Molecular Biology*, **2002**, 319, 1257-65	6.5	274
122	An HMM posterior decoder for sequence feature prediction that includes homology information. *Bioinformatics*, **2005**, 21 Suppl 1, i251-7	7.2	246
121	A hidden Markov model that finds genes in E. coli DNA. *Nucleic Acids Research*, **1994**, 22, 4768-78	20.1	219
120	What are artificial neural networks?. *Nature Biotechnology*, **2008**, 26, 195-7	44.5	207
119	The genome of the leaf-cutting ant Acromyrmex echinatior suggests key adaptations to advanced social life and fungus farming. *Genome Research*, **2011**, 21, 1339-48	9.7	183
118	miR-449 inhibits cell proliferation and is down-regulated in gastric cancer. *Molecular Cancer*, **2011**, 10, 29	42.1	181
117	Reliability measures for membrane protein topology prediction algorithms. *Journal of Molecular Biology*, **2003**, 327, 735-44	6.5	181
116	A code for transcription initiation in mammalian genomes. *Genome Research*, **2008**, 18, 1-12	9.7	179
115	Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics*, **2012**, 13, 178	4.5	178
114	On the total number of genes and their length distribution in complete microbial genomes. *Trends in Genetics*, **2001**, 17, 425-8	8.5	170
113	Genome analysis reveals insights into physiology and longevity of the Brandt's bat Myotis brandtii. *Nature Communications*, **2013**, 4, 2212	17.4	160
112	Molecular composition of IMP1 ribonucleoprotein granules. *Molecular and Cellular Proteomics*, **2007**, 6, 798-811	7.6	160
111	Hidden Markov models for sequence analysis: extension and analysis of the basic method. *Bioinformatics*, **1996**, 12, 95-107	7.2	160
110	A sequence-profile-based HMM for predicting and discriminating beta barrel membrane proteins. *Bioinformatics*, **2002**, 18 Suppl 1, S46-53	7.2	158
109	No evidence that mRNAs have lower folding free energies than random sequences with the same dinucleotide distribution. *Nucleic Acids Research*, **1999**, 27, 4816-22	20.1	149
108	Dirichlet mixtures: a method for improved detection of weak but significant protein sequence homology. *Bioinformatics*, **1996**, 12, 327-45	7.2	139
107	MicroRNA-145 targets YES and STAT1 in colon cancer cells. *PLoS ONE*, **2010**, 5, e8836	3.7	138
106	Adaptations to a subterranean environment and longevity revealed by the analysis of mole rat genomes. *Cell Reports*, **2014**, 8, 1354-64	10.6	124

105	Genomic Analyses from Non-invasive Prenatal Testing Reveal Genetic Associations, Patterns of Viral Infections, and Chinese Population History. *Cell*, **2018**, 175, 347-359.e14	56.2	123
104	EasyGene--a prokaryotic gene finder that ranks ORFs by statistical significance. *BMC Bioinformatics*, **2003**, 4, 21	3.6	120
103	Novel variation and de novo mutation rates in population-wide de novo assembled Danish trios. *Nature Communications*, **2015**, 6, 5969	17.4	119
102	Improving prediction of protein secondary structure using structured neural networks and multiple sequence alignments. *Journal of Computational Biology*, **1996**, 3, 163-83	1.7	117
101	Genome-wide detection and analysis of hippocampus core promoters using DeepCAGE. *Genome Research*, **2009**, 19, 255-65	9.7	116
100	Rapid detection of structural variation in a human genome using nanochannel-based genome mapping technology. *GigaScience*, **2014**, 3, 34	7.6	113
99	Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome. *Genome Research*, **2014**, 24, 454-66	9.7	113
98	Large-scale prokaryotic gene prediction and comparison to genome annotation. *Bioinformatics*, **2005**, 21, 4322-9	7.2	113
97	MicroRNA-143 down-regulates Hexokinase 2 in colon cancer cells. *BMC Cancer*, **2012**, 12, 232	4.8	109
96	Genome-wide adaptive complexes to underground stresses in blind mole rats Spalax. *Nature Communications*, **2014**, 5, 3966	17.4	101
95	Mammalian tissues defective in nonsense-mediated mRNA decay display highly aberrant splicing patterns. *Genome Biology*, **2012**, 13, R35	18.3	90
94	Sequencing and de novo assembly of 150 genomes from Denmark as a population reference. *Nature*, **2017**, 548, 87-91	50.4	87
93	Signatures of RNA binding proteins globally coupled to effective microRNA target sites. *Genome Research*, **2010**, 20, 1010-9	9.7	86
92	A generative, probabilistic model of local protein structure. *Proceedings of the National Academy of Sciences of the United States of America*, **2008**, 105, 8932-7	11.5	84
91	Comparing cancer vs normal gene expression profiles identifies new disease entities and common transcriptional programs in AML patients. *Blood*, **2014**, 123, 894-904	2.2	82
90	Statistical mechanics of ensemble learning. *Physical Review E*, **1997**, 55, 811-825	2.4	81
89	Comparative Metagenomics of Eight Geographically Remote Terrestrial Hot Springs. *Microbial Ecology*, **2015**, 70, 411-24	4.4	79
88	Reconstructing genome evolution in historic samples of the Irish potato famine pathogen. *Nature Communications*, **2013**, 4, 2172	17.4	79

87	microRNA-146a inhibits G protein-coupled receptor-mediated activation of NF-**B** by targeting CARD10 and COPS8 in gastric cancer. *Molecular Cancer*, **2012**, 11, 71	42.1	75
86	Genome organisation and chromatin structure in Escherichia coli. *Biochimie*, **2001**, 83, 201-12	4.6	71
85	Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature*, **1998**, 396, 190-190	50.4	70
84	RpoD promoters in Campylobacter jejuni exhibit a strong periodic signal instead of a -35 box. *Journal of Molecular Biology*, **2003**, 326, 1361-72	6.5	70
83	Introduction to the Theory of Neural Computation		68
82	Sampling realistic protein conformations using local structural bias. *PLoS Computational Biology*, **2006**, 2, e131	5	66
81	An introduction to hidden Markov models for biological sequences. *New Comprehensive Biochemistry*, **1998**, 32, 45-63		63
80	Naturally occurring nucleosome positioning signals in human exons and introns. *Journal of Molecular Biology*, **1996**, 263, 503-10	6.5	62
79	Generalization in a linear perceptron in the presence of noise. *Journal of Physics A*, **1992**, 25, 1135-1147		60
78	Computational evidence for hundreds of non-conserved plant microRNAs. *BMC Genomics*, **2005**, 6, 119	4.5	59
77	De novo assembly of a haplotype-resolved human genome. *Nature Biotechnology*, **2015**, 33, 617-22	44.5	57
76	MASTR: multiple alignment and structure prediction of non-coding RNAs using simulated annealing. *Bioinformatics*, **2007**, 23, 3304-11	7.2	57
75	MicroRNA transfection and AGO-bound CLIP-seq data sets reveal distinct determinants of miRNA action. *Rna*, **2011**, 17, 820-34	5.8	56
74	Sigma A recognition sites in the Bacillus subtilis genome. *Microbiology (United Kingdom)*, **2001**, 147, 2417-2424	2.9	56
73	Intragenomic matching reveals a huge potential for miRNA-mediated regulation in plants. *PLoS Computational Biology*, **2007**, 3, e238	5	52
72	Using database matches with for HMMGene for automated gene detection in Drosophila. *Genome Research*, **2000**, 10, 523-8	9.7	50
71	Bayesian transcriptome assembly. *Genome Biology*, **2014**, 15, 501	18.3	43
70	Phase transitions in simple learning. *Journal of Physics A*, **1989**, 22, 2133-2150		43

69	Novel viral genomes identified from six metagenomes reveal wide distribution of archaeal viruses and high viral diversity in terrestrial hot springs. *Environmental Microbiology*, **2016**, 18, 863-74	5.2	40
68	CCHMM_PROF: a HMM-based coiled-coil predictor with evolutionary information. *Bioinformatics*, **2009**, 25, 2757-63	7.2	38
67	Identification and analysis of miRNAs in human breast cancer and teratoma samples using deep sequencing. *BMC Medical Genomics*, **2009**, 2, 35	3.7	38
66	A short-read multiplex sequencing method for reliable, cost-effective and high-throughput genotyping in large-scale studies. *Human Mutation*, **2013**, 34, 1715-20	4.7	37
65	Measuring covariation in RNA alignments: physical realism improves information measures. *Bioinformatics*, **2006**, 22, 2988-95	7.2	37
64	Adaptable probabilistic mapping of short reads using position specific scoring matrices. *BMC Bioinformatics*, **2014**, 15, 100	3.6	36
63	PONGO: a web server for multiple predictions of all-alpha transmembrane proteins. *Nucleic Acids Research*, **2006**, 34, W169-72	20.1	36
62	Accurate genotyping across variant classes and lengths using variant graphs. *Nature Genetics*, **2018**, 50, 1054-1059	36.3	35
61	SHAPE Selection (SHAPES) enrich for RNA structure signal in SHAPE sequencing-based probing data. *Rna*, **2015**, 21, 1042-52	5.8	35
60	An evolutionary method for learning HMM structure: prediction of protein secondary structure. *BMC Bioinformatics*, **2007**, 8, 357	3.6	35
59	Asap: a framework for over-representation statistics for transcription factor binding sites. *PLoS ONE*, **2008**, 3, e1623	3.7	35
58	Cancers of unknown primary origin (CUP) are characterized by chromosomal instability (CIN) compared to metastasis of know origin. *BMC Cancer*, **2015**, 15, 151	4.8	34
57	Loss of miR-10a activates lpo and collaborates with activated Wnt signaling in inducing intestinal neoplasia in female mice. *PLoS Genetics*, **2013**, 9, e1003913	6	34
56	Highly accessible AU-rich regions in 3untranslated regions are hotspots for binding of regulatory factors. *PLoS Computational Biology*, **2017**, 13, e1005460	5	33
55	Training HMM structure with genetic algorithm for biological sequence analysis. *Bioinformatics*, **2004**, 20, 3613-9	7.2	33
54	Hidden neural networks. *Neural Computation*, **1999**, 11, 541-63	2.9	33
53	Bias of purine stretches in sequenced chromosomes. *Computers & Chemistry*, **2002**, 26, 531-41		31
52	Alterations in polyadenylation and its implications for endocrine disease. *Frontiers in Endocrinology*, **2013**, 4, 53	5.7	26

51	Drosophila Imp iCLIP identifies an RNA assemblage coordinating F-actin formation. *Genome Biology*, **2015**, 16, 123	18.3	23
50	Discovery of regulatory elements is improved by a discriminatory approach. *PLoS Computational Biology*, **2009**, 5, e1000562	5	23
49	Hierarchical associative networks. *Journal of Physics A*, **1987**, 20, 4449-4455		23
48	A hidden Markov model approach for determining expression from genomic tiling micro arrays. *BMC Bioinformatics*, **2006**, 7, 239	3.6	22
47	Learning with noise in a linear perceptron. *Journal of Physics A*, **1992**, 25, 1119-1133		21
46	Sensitive detection of circular DNAs at single-nucleotide resolution using guided realignment of partially aligned reads. *BMC Bioinformatics*, **2019**, 20, 663	3.6	21
45	Automatic generation of gene finders for eukaryotic species. *BMC Bioinformatics*, **2006**, 7, 263	3.6	20
44	Long- and short-term selective forces on malaria parasite genomes. *PLoS Genetics*, **2010**, 6, e1001099	6	20
43	cWords - systematic microRNA regulatory motif discovery from mRNA expression data. *Silence: A Journal of RNA Regulation*, **2013**, 4, 2		19
42	A quantitative study of pruning by optimal brain damage. *International Journal of Neural Systems*, **1993**, 4, 159-69	6.2	18
41	Mean-field analysis of hierarchical associative networks with Umagnetisation*U Journal of Physics A*, **1988**, 21, 2211-2224		17
40	Discovery, genotyping and characterization of structural variation and novel sequence at single nucleotide resolution from de novo genome assemblies on a population scale. *GigaScience*, **2015**, 4, 64	7.6	16
39	Novel overlapping coding sequences in Chlamydia trachomatis. *FEMS Microbiology Letters*, **2006**, 265, 106-17	2.9	15
38	Gene Finding: Putting the Parts Together **1998**, 261-274		14
37	Nonlinear backpropagation: doing backpropagation without derivatives of the activation function. *IEEE Transactions on Neural Networks*, **1997**, 8, 1321-7		13
36	Sugar Metabolism of the First Thermophilic Planctomycete : Comparative Genomic and Transcriptomic Approaches. *Frontiers in Microbiology*, **2017**, 8, 2140	5.7	12
35	miRMaid: a unified programming interface for microRNA data resources. *BMC Bioinformatics*, **2010**, 11, 29	3.6	12
34	Prediction of Structurally-Determined Coiled-Coil Domains with Hidden Markov Models **2007**, 292-302		12

33   Modeling promoter grammars with evolving hidden Markov models. *Bioinformatics*, **2008**, 24, 1669-75   7.2   11

32   Alternative polyadenylation of tumor suppressor genes in small intestinal neuroendocrine tumors. *Frontiers in Endocrinology*, **2014**, 5, 46   5.7   10

31   Hidden Markov Models for prediction of protein features. *Methods in Molecular Biology*, **2008**, 413, 173-98   1.4   10

30   Teaching computers to fold proteins. *Physical Review E*, **2004**, 70, 030903   2.4   10

29   Evolving the structure of hidden Markov models. *IEEE Transactions on Evolutionary Computation*, **2006**, 10, 39-49   15.6   9

28   Secreted breast tumor interstitial fluid microRNAs and their target genes are associated with triple-negative breast cancer, tumor grade, and immune infiltration. *Breast Cancer Research*, **2020**, 22, 73   8.3   8

27   RNA Sequencing of Trigeminal Ganglia in Rattus Norvegicus after Glyceryl Trinitrate Infusion with Relevance to Migraine. *PLoS ONE*, **2016**, 11, e0155039   3.7   7

26   SNPest: a probabilistic graphical model for estimating genotypes. *BMC Research Notes*, **2014**, 7, 698   2.3   6

25   BayesMD: flexible biological modeling for motif discovery. *Journal of Computational Biology*, **2008**, 15, 1347-63   1.7   6

24   Dynamics of Learning in Simple Perceptrons. *Physica Scripta*, **1989**, T25, 149-151   2.6   6

23   On the accuracy of short read mapping. *Methods in Molecular Biology*, **2013**, 1038, 39-59   1.4   6

22   Taxonomic classification method for metagenomics based on core protein families with Core-Kaiju. *Nucleic Acids Research*, **2020**, 48, e93   20.1   5

21   Early Pleistocene enamel proteome sequences from Dmanisi resolve Stephanorhinus phylogeny   5

20   High-throughput proteomics of breast cancer interstitial fluid: identification of tumor subtype-specific serologically relevant biomarkers. *Molecular Oncology*, **2021**, 15, 429-461   7.9   4

19   Transcriptome dynamics of the microRNA inhibition response. *Nucleic Acids Research*, **2015**, 43, 6207-21   20.1   3

18   . *IEEE Transactions on Information Theory*, **1994**, 40, 1215-1218   2.8   3

17   CAncer bioMarker Prediction Pipeline (CAMPP)-A standardized framework for the analysis of quantitative biological data. *PLoS Computational Biology*, **2020**, 16, e1007665   5   2

16   Sensitive detection of circular DNAs at single-nucleotide resolution using guided realignment of partially aligned reads   2

List of Publications

#	Publication		
15	Prediction of DNA from context using neural networks		2
14	A General Method for Combining Predictors Tested on Protein Secondary Structure Prediction. *Perspectives in Neural Computing*, **2000**, 259-264		2
13	Multiple alignment and structure prediction of non-coding RNA sequences. *BMC Bioinformatics*, **2007**, 8,	3.6	1
12	The Block Hidden Markov Model for Biological Sequence Analysis. *Lecture Notes in Computer Science*, **2004**, 64-70	0.9	1
11	Markov chains and hidden Markov models **1998**, 47-80		1
10	Context dependency of nucleotide probabilities and variants in human DNA.. *BMC Genomics*, **2022**, 23, 87	4.5	1
9	Low Pass Genomes of 141,431 Chinese Reveal Patterns of Viral Infection, Novel Phenotypic Associations, and the Genetic History of China. *SSRN Electronic Journal*,	1	1
8	Context dependency of nucleotide probabilities and variants in human DNA		1
7	Hidden Markov Models for Human Genes **1997**, 15-32		
6	Improving The Analysis Of Gene Expression Profiles By Comparing AML Blasts With Their Nearest Normal Counterparts. *Blood*, **2013**, 122, 2568-2568	2.2	
5	CAncer bioMarker Prediction Pipeline (CAMPP)A standardized framework for the analysis of quantitative biological data **2020**, 16, e1007665		
4	CAncer bioMarker Prediction Pipeline (CAMPP)A standardized framework for the analysis of quantitative biological data **2020**, 16, e1007665		
3	CAncer bioMarker Prediction Pipeline (CAMPP)A standardized framework for the analysis of quantitative biological data **2020**, 16, e1007665		
2	CAncer bioMarker Prediction Pipeline (CAMPP)A standardized framework for the analysis of quantitative biological data **2020**, 16, e1007665		
1	CAncer bioMarker Prediction Pipeline (CAMPP)A standardized framework for the analysis of quantitative biological data **2020**, 16, e1007665		