

Nick Goldman

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/1670334/publications.pdf>

Version: 2024-02-01

109
papers

30,690
citations

34016

52
h-index

28224

105
g-index

127
all docs

127
docs citations

127
times ranked

38661
citing authors

#	ARTICLE	IF	CITATIONS
1	Initial sequencing and comparative analysis of the mouse genome. <i>Nature</i> , 2002, 420, 520-562.	13.7	6,319
2	Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. <i>Nature</i> , 2007, 447, 799-816.	13.7	4,709
3	A General Empirical Model of Protein Evolution Derived from Multiple Protein Families Using a Maximum-Likelihood Approach. <i>Molecular Biology and Evolution</i> , 2001, 18, 691-699.	3.5	2,599
4	Codon-Substitution Models for Heterogeneous Selection Pressure at Amino Acid Sites. <i>Genetics</i> , 2000, 155, 431-449.	1.2	2,064
5	Likelihood-Based Tests of Topologies in Phylogenetics. <i>Systematic Biology</i> , 2000, 49, 652-670.	2.7	1,032
6	A high-resolution map of human evolutionary constraint using 29 mammals. <i>Nature</i> , 2011, 478, 476-482.	13.7	1,016
7	From The Cover: An algorithm for progressive multiple alignment of sequences with insertions. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2005, 102, 10557-10562.	3.3	871
8	Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. <i>Nature</i> , 2013, 494, 77-80.	13.7	787
9	Statistical tests of models of DNA substitution. <i>Journal of Molecular Evolution</i> , 1993, 36, 182-198.	0.8	746
10	Phylogeny-Aware Gap Placement Prevents Errors in Sequence Alignment and Evolutionary Analysis. <i>Science</i> , 2008, 320, 1632-1635.	6.0	737
11	Insights into hominid evolution from the gorilla genome sequence. <i>Nature</i> , 2012, 483, 169-175.	13.7	663
12	Accuracy and Power of Statistical Methods for Detecting Adaptive Evolution in Protein Coding Sequences and for Identifying Positively Selected Sites. <i>Genetics</i> , 2004, 168, 1041-1051.	1.2	543
13	Systematic evaluation of spliced alignment programs for RNA-seq data. <i>Nature Methods</i> , 2013, 10, 1185-1191.	9.0	467
14	webPRANK: a phylogeny-aware multiple sequence aligner with interactive alignment browser. <i>BMC Bioinformatics</i> , 2010, 11, 579.	1.2	418
15	Molecular phylogenetics: state-of-the-art methods for looking into the past. <i>Trends in Genetics</i> , 2001, 17, 262-272.	2.9	376
16	Conservation and divergence in Toll-like receptor 4-regulated gene expression in primary human versus mouse macrophages. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2012, 109, E944-53.	3.3	332
17	Genome sequencing of normal cells reveals developmental lineages and mutational processes. <i>Nature</i> , 2014, 513, 422-425.	13.7	315
18	Covariation in Frequencies of Substitution, Deletion, Transposition, and Recombination During Eutherian Evolution. <i>Genome Research</i> , 2003, 13, 13-26.	2.4	263

#	ARTICLE	IF	CITATIONS
19	Transcriptional diversity during lineage commitment of human blood progenitors. <i>Science</i> , 2014, 345, 1251033.	6.0	253
20	Models of Molecular Evolution and Phylogeny: Table 1.. <i>Genome Research</i> , 1998, 8, 1233-1244.	2.4	252
21	Coevolving protein residues: maximum likelihood identification and relationship to structure 1 Edited by G. Von Heijne. <i>Journal of Molecular Biology</i> , 1999, 287, 187-198.	2.0	242
22	Maximum Likelihood Trees from DNA Sequences: A Peculiar Statistical Estimation Problem. <i>Systematic Biology</i> , 1995, 44, 384-399.	2.7	239
23	Assessing the Impact of Secondary Structure and Solvent Accessibility on Protein Evolution. <i>Genetics</i> , 1998, 149, 445-458.	1.2	228
24	Detecting Amino Acid Sites Under Positive Selection and Purifying Selection. <i>Genetics</i> , 2005, 169, 1753-1762.	1.2	221
25	Current Methods for Automated Filtering of Multiple Sequence Alignments Frequently Worsen Single-Gene Phylogenetic Inference. <i>Systematic Biology</i> , 2015, 64, 778-791.	2.7	200
26	Maximum Likelihood Inference of Phylogenetic Trees, with Special Reference to a Poisson Process Model of DNA Substitution and to Parsimony Analyses. <i>Systematic Zoology</i> , 1990, 39, 345.	1.6	191
27	RNAcode: Robust discrimination of coding and noncoding regions in comparative sequence data. <i>Rna</i> , 2011, 17, 578-594.	1.6	188
28	Analyses of deep mammalian sequence alignments and constraint predictions for 1% of the human genome. <i>Genome Research</i> , 2007, 17, 760-774.	2.4	184
29	Protein Evolution with Dependence Among Codons Due to Tertiary Structure. <i>Molecular Biology and Evolution</i> , 2003, 20, 1692-1704.	3.5	180
30	An Empirical Codon Model for Protein Sequence Evolution. <i>Molecular Biology and Evolution</i> , 2007, 24, 1464-1479.	3.5	166
31	Different Versions of the Dayhoff Rate Matrix. <i>Molecular Biology and Evolution</i> , 2005, 22, 193-199.	3.5	165
32	Genomic DNA k-mer spectra: models and modalities. <i>Genome Biology</i> , 2009, 10, R108.	13.9	165
33	The Effects of Alignment Error and Alignment Filtering on the Sitewise Detection of Positive Selection. <i>Molecular Biology and Evolution</i> , 2012, 29, 1125-1139.	3.5	164
34	Statistical Tests of Gamma-Distributed Rate Heterogeneity in Models of Sequence Evolution in Phylogenetics. <i>Molecular Biology and Evolution</i> , 2000, 17, 975-978.	3.5	128
35	Accurate extension of multiple sequence alignments using a phylogeny-aware graph algorithm. <i>Bioinformatics</i> , 2012, 28, 1684-1691.	1.8	126
36	Phylogenetic Quantification of Intra-tumour Heterogeneity. <i>PLoS Computational Biology</i> , 2014, 10, e1003535.	1.5	126

#	ARTICLE	IF	CITATIONS
37	Simple diagnostic statistical tests of models for DNA substitution. <i>Journal of Molecular Evolution</i> , 1993, 37, 650-61.	0.8	118
38	Phylogenetic information and experimental design in molecular systematics. <i>Proceedings of the Royal Society B: Biological Sciences</i> , 1998, 265, 1779-1786.	1.2	116
39	Using Evolutionary Trees in Protein Secondary Structure Prediction and Other Comparative Sequence Analyses. <i>Journal of Molecular Biology</i> , 1996, 263, 196-208.	2.0	113
40	SMIM1 underlies the Vel blood group and influences red blood cell traits. <i>Nature Genetics</i> , 2013, 45, 542-545.	9.4	96
41	Want to track pandemic variants faster? Fix the bioinformatics bottleneck. <i>Nature</i> , 2021, 591, 30-33.	13.7	92
42	Stability of SARS-CoV-2 phylogenies. <i>PLoS Genetics</i> , 2020, 16, e1009175.	1.5	92
43	Nucleotide, dinucleotide and trinucleotide frequencies explain patterns observed in chaos game representations of DNA sequences. <i>Nucleic Acids Research</i> , 1993, 21, 2487-2491.	6.5	87
44	Variation in Evolutionary Processes at Different Codon Positions. <i>Molecular Biology and Evolution</i> , 2006, 24, 513-521.	3.5	83
45	Mutation Rates and Selection on Synonymous Mutations in SARS-CoV-2. <i>Genome Biology and Evolution</i> , 2021, 13, .	1.1	83
46	Estimating the Frequency of Events That Cause Multiple-Nucleotide Changes. <i>Genetics</i> , 2004, 167, 2027-2043.	1.2	81
47	Genomic reconstruction of the SARS-CoV-2 epidemic in England. <i>Nature</i> , 2021, 600, 506-511.	13.7	80
48	Evolutionary footprints of nucleosome positions in yeast. <i>Trends in Genetics</i> , 2008, 24, 583-587.	2.9	70
49	Alignment Modulates Ancestral Sequence Reconstruction Accuracy. <i>Molecular Biology and Evolution</i> , 2018, 35, 1783-1797.	3.5	70
50	A Daily-Updated Database and Tools for Comprehensive SARS-CoV-2 Mutation-Annotated Trees. <i>Molecular Biology and Evolution</i> , 2021, 38, 5819-5824.	3.5	69
51	METHODS FOR DISCRETE CODING OF MORPHOLOGICAL CHARACTERS FOR NUMERICAL ANALYSIS. <i>Cladistics</i> , 1988, 4, 59-71.	1.5	67
52	Experimental Design Criteria in Phylogenetics: Where to Add Taxa. <i>Systematic Biology</i> , 2007, 56, 609-622.	2.7	65
53	Variance to Mean Ratio, $R(t)$, for Poisson Processes on Phylogenetic Trees. <i>Molecular Phylogenetics and Evolution</i> , 1994, 3, 230-239.	1.2	61
54	Species Choice for Comparative Genomics: Being Greedy Works. <i>PLoS Genetics</i> , 2005, 1, e71.	1.5	60

#	ARTICLE	IF	CITATIONS
55	A new criterion and method for amino acid classification. <i>Journal of Theoretical Biology</i> , 2004, 228, 97-106.	0.8	59
56	PANDIT: an evolution-centric database of protein and associated nucleotide domains with inferred trees. <i>Nucleic Acids Research</i> , 2006, 34, D327-D331.	6.5	59
57	A Penalized-Likelihood Method to Estimate the Distribution of Selection Coefficients from Phylogenetic Data. <i>Genetics</i> , 2014, 197, 257-271.	1.2	55
58	PhyloSim - Monte Carlo simulation of sequence evolution in the R statistical computing environment. <i>BMC Bioinformatics</i> , 2011, 12, 104.	1.2	54
59	Amino Acid Changes in Disease-Associated Variants Differ Radically from Variants Observed in the 1000 Genomes Project Dataset. <i>PLoS Computational Biology</i> , 2013, 9, e1003382.	1.5	54
60	Further results on error minimization in the genetic code. <i>Journal of Molecular Evolution</i> , 1993, 37, 662-4.	0.8	51
61	Clustering Genes of Common Evolutionary History. <i>Molecular Biology and Evolution</i> , 2016, 33, 1590-1605.	3.5	51
62	XRate: a fast prototyping, training and annotation tool for phylo-grammars. <i>BMC Bioinformatics</i> , 2006, 7, 428.	1.2	49
63	More on the Best Evolutionary Rate for Phylogenetic Analysis. <i>Systematic Biology</i> , 2017, 66, 769-785.	2.7	48
64	Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk. <i>PLoS Computational Biology</i> , 2021, 17, e1008561.	1.5	48
65	A Novel Use of Equilibrium Frequencies in Models of Sequence Evolution. <i>Molecular Biology and Evolution</i> , 2002, 19, 1821-1831.	3.5	46
66	The human blood DNA methylome displays a highly distinctive profile compared with other somatic tissues. <i>Epigenetics</i> , 2015, 10, 274-281.	1.3	46
67	Who Watches the Watchmen? An Appraisal of Benchmarks for Multiple Sequence Alignment. <i>Methods in Molecular Biology</i> , 2014, 1079, 59-73.	0.4	45
68	Pandit: a database of protein and associated nucleotide domains with inferred trees. <i>Bioinformatics</i> , 2003, 19, 1556-1563.	1.8	44
69	A model of evolution and structure for multiple sequence alignment. <i>Philosophical Transactions of the Royal Society B: Biological Sciences</i> , 2008, 363, 3913-3919.	1.8	44
70	Phylogenetic analysis of the rpoB gene from the plastid-like DNA of <i>Plasmodium falciparum</i> . <i>Molecular and Biochemical Parasitology</i> , 1994, 66, 221-231.	0.5	41
71	Maximum Likelihood Trees from DNA Sequences: A Peculiar Statistical Estimation Problem. <i>Systematic Biology</i> , 1995, 44, 384.	2.7	41
72	All your base: a fast and accurate probabilistic approach to base calling. <i>Genome Biology</i> , 2012, 13, R13.	13.9	39

#	ARTICLE	IF	CITATIONS
73	Phylogenomics and bioinformatics of SARS-CoV. <i>Trends in Microbiology</i> , 2004, 12, 106-111.	3.5	36
74	In defense of statistical methods for detecting positive selection. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2009, 106, E95; author reply E96.	3.3	36
75	Resource-Aware Taxon Selection for Maximizing Phylogenetic Diversity. <i>Systematic Biology</i> , 2007, 56, 431-444.	2.7	35
76	Short template switch events explain mutation clusters in the human genome. <i>Genome Research</i> , 2017, 27, 1039-1049.	2.4	31
77	Integrated structural and evolutionary analysis reveals common mechanisms underlying adaptive evolution in mammals. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2020, 117, 5977-5986.	3.3	31
78	Looking for Darwin in Genomic Sequences—Validity and Success of Statistical Methods. <i>Molecular Biology and Evolution</i> , 2012, 29, 2889-2893.	3.5	28
79	Maximum Likelihood Phylogenetic Inference is Consistent on Multiple Sequence Alignments, with or without Gaps. <i>Systematic Biology</i> , 2016, 65, 328-333.	2.7	28
80	Uniting Alignments and Trees. <i>Science</i> , 2009, 324, 1528-1529.	6.0	24
81	Addressing Inter-Gene Heterogeneity in Maximum Likelihood Phylogenomic Analysis: Yeasts Revisited. <i>PLoS ONE</i> , 2011, 6, e22783.	1.1	24
82	Determination and validation of principal gene products. <i>Bioinformatics</i> , 2008, 24, 11-17.	1.8	23
83	Markovian and Non-Markovian Protein Sequence Evolution: Aggregated Markov Process Models. <i>Journal of Molecular Biology</i> , 2011, 411, 910-923.	2.0	23
84	Genetic Variability of the SARS-CoV-2 Pocketome. <i>Journal of Proteome Research</i> , 2021, 20, 4212-4215.	1.8	23
85	Maximum Likelihood Inference of Small Trees in the Presence of Long Branches. <i>Systematic Biology</i> , 2014, 63, 798-811.	2.7	21
86	Are big trees indeed easy?. <i>Trends in Ecology and Evolution</i> , 1997, 12, 357.	4.2	19
87	Rapidly evolving human promoter regions. <i>Nature Genetics</i> , 2008, 40, 1262-1263.	9.4	18
88	Simple chained guide trees give poorer multiple sequence alignments than inferred trees in simulation and phylogenetic benchmarks. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2015, 112, E99-100.	3.3	18
89	Effects of sequence alignment procedures on estimates of phylogeny. <i>BioEssays</i> , 1998, 20, 287-290.	1.2	17
90	Genetics and geography. <i>Nature</i> , 1992, 357, 440-441.	13.7	15

#	ARTICLE	IF	CITATIONS
91	Sequence Bundles: a novel method for visualising, discovering and exploring sequence motifs. BMC Proceedings, 2014, 8, S8.	1.8	13
92	phastSim: Efficient simulation of sequence evolution for pandemic-scale datasets. PLoS Computational Biology, 2022, 18, e1010056.	1.5	13
93	Modeling Mitochondrial Protein Evolution Using Structural Information. Journal of Molecular Evolution, 2002, 54, 519-529.	0.8	12
94	Introduction. Statistical and computational challenges in molecular phylogenetics and evolution. Philosophical Transactions of the Royal Society B: Biological Sciences, 2008, 363, 3889-3892.	1.8	12
95	What's in a Likelihood? Simple Models of Protein Evolution and the Contribution of Structurally Viable Reconstructions to the Likelihood. Systematic Biology, 2011, 60, 161-174.	2.7	12
96	Short-range template switching in great ape genomes explored using pair hidden Markov models. PLoS Genetics, 2021, 17, e1009221.	1.5	11
97	Error-correcting properties of the SOLiD Exact Call Chemistry. BMC Bioinformatics, 2012, 13, 145.	1.2	10
98	An Improved Protocol for Sequencing of Repetitive Genomic Regions and Structural Variations Using Mutagenesis and Next Generation Sequencing. PLoS ONE, 2012, 7, e43359.	1.1	10
99	Modeling Structural Constraints on Protein Evolution via Side-Chain Conformational States. Molecular Biology and Evolution, 2019, 36, 2086-2103.	3.5	10
100	ALVIS: interactive non-aggregative visualization and explorative analysis of multiple sequence alignments. Nucleic Acids Research, 2016, 44, e77-e77.	6.5	9
101	Modeling Mitochondrial Protein Evolution Using Structural Information. Journal of Molecular Evolution, 2002, 54, 519-529.	0.8	7
102	Fewest Variables Coding Method for Multistate Characters. Systematic Zoology, 1989, 38, 79.	1.6	5
103	Improving communication for interdisciplinary teams working on storage of digital information in DNA. F1000Research, 2018, 7, 39.	0.8	4
104	A phylogenetic approach for weighting genetic sequences. BMC Bioinformatics, 2021, 22, 285.	1.2	3
105	Ambiguity Coding Allows Accurate Inference of Evolutionary Parameters from Alignments in an Aggregated State-Space. Systematic Biology, 2021, 70, 21-32.	2.7	1
106	Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk. , 2021, 17, e1008561.		0
107	Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk. , 2021, 17, e1008561.		0
108	Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk. , 2021, 17, e1008561.		0

#	ARTICLE	IF	CITATIONS
109	Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk. , 2021, 17, e1008561.		0