

# Nick Goldman

## List of Publications by Citations

**Source:** <https://exaly.com/author-pdf/1670334/nick-goldman-publications-by-citations.pdf>

**Version:** 2024-04-26

This document has been generated based on the publications and citations recorded by exaly.com. For the latest version of this publication list, visit the link given above.

The third column is the impact factor (IF) of the journal, and the fourth column is the number of citations of the article.

114  
papers

24,697  
citations

48  
h-index

127  
g-index

127  
ext. papers

28,431  
ext. citations

12.3  
avg, IF

6.72  
L-index

#	Paper	IF	Citations
114	Initial sequencing and comparative analysis of the mouse genome. <i>Nature</i> , <b>2002</b> , 420, 520-62	50.4	5376
113	Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. <i>Nature</i> , <b>2007</b> , 447, 799-816	50.4	4121
112	A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. <i>Molecular Biology and Evolution</i> , <b>2001</b> , 18, 691-9	8.3	2055
111	Codon-substitution models for heterogeneous selection pressure at amino acid sites. <i>Genetics</i> , <b>2000</b> , 155, 431-49	4	1740
110	Likelihood-based tests of topologies in phylogenetics. <i>Systematic Biology</i> , <b>2000</b> , 49, 652-70	8.4	930
109	A high-resolution map of human evolutionary constraint using 29 mammals. <i>Nature</i> , <b>2011</b> , 478, 476-82	50.4	802
108	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> , <b>2005</b> , 102, 10557-62	11.5	699
107	Statistical tests of models of DNA substitution. <i>Journal of Molecular Evolution</i> , <b>1993</b> , 36, 182-98	3.1	624
106	Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. <i>Science</i> , <b>2008</b> , 320, 1632-5	33.3	599
105	Insights into hominid evolution from the gorilla genome sequence. <i>Nature</i> , <b>2012</b> , 483, 169-75	50.4	517
104	Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. <i>Nature</i> , <b>2013</b> , 494, 77-80	50.4	501
103	Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. <i>Genetics</i> , <b>2004</b> , 168, 1041-51	4	443
102	Systematic evaluation of spliced alignment programs for RNA-seq data. <i>Nature Methods</i> , <b>2013</b> , 10, 1185-91.6	21.6	371
101	Molecular phylogenetics: state-of-the-art methods for looking into the past. <i>Trends in Genetics</i> , <b>2001</b> , 17, 262-72	8.5	311
100	webPRANK: a phylogeny-aware multiple sequence aligner with interactive alignment browser. <i>BMC Bioinformatics</i> , <b>2010</b> , 11, 579	3.6	267
99	Genome sequencing of normal cells reveals developmental lineages and mutational processes. <i>Nature</i> , <b>2014</b> , 513, 422-425	50.4	249
98	Covariation in frequencies of substitution, deletion, transposition, and recombination during eutherian evolution. <i>Genome Research</i> , <b>2003</b> , 13, 13-26	9.7	234

97	Models of molecular evolution and phylogeny. <i>Genome Research</i> , <b>1998</b> , 8, 1233-44	9.7	215
96	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> , <b>2012</b> , 109, E944-53	11.5	212
95	Coevolving protein residues: maximum likelihood identification and relationship to structure. <i>Journal of Molecular Biology</i> , <b>1999</b> , 287, 187-98	6.5	204
94	Maximum Likelihood Trees from DNA Sequences: A Peculiar Statistical Estimation Problem. <i>Systematic Biology</i> , <b>1995</b> , 44, 384-399	8.4	198
93	Transcriptional diversity during lineage commitment of human blood progenitors. <i>Science</i> , <b>2014</b> , 345, 1251033	33.3	187
92	Assessing the impact of secondary structure and solvent accessibility on protein evolution. <i>Genetics</i> , <b>1998</b> , 149, 445-58	4	172
91	Detecting amino acid sites under positive selection and purifying selection. <i>Genetics</i> , <b>2005</b> , 169, 1753-624		165
90	Analyses of deep mammalian sequence alignments and constraint predictions for 1% of the human genome. <i>Genome Research</i> , <b>2007</b> , 17, 760-74	9.7	163
89	Protein evolution with dependence among codons due to tertiary structure. <i>Molecular Biology and Evolution</i> , <b>2003</b> , 20, 1692-704	8.3	160
88	An empirical codon model for protein sequence evolution. <i>Molecular Biology and Evolution</i> , <b>2007</b> , 24, 1464-79	8.3	139
87	RNAcode: robust discrimination of coding and noncoding regions in comparative sequence data. <i>Rna</i> , <b>2011</b> , 17, 578-94	5.8	131
86	Current Methods for Automated Filtering of Multiple Sequence Alignments Frequently Worsen Single-Gene Phylogenetic Inference. <i>Systematic Biology</i> , <b>2015</b> , 64, 778-91	8.4	129
85	The effects of alignment error and alignment filtering on the sitewise detection of positive selection. <i>Molecular Biology and Evolution</i> , <b>2012</b> , 29, 1125-39	8.3	125
84	Genomic DNA k-mer spectra: models and modalities. <i>Genome Biology</i> , <b>2009</b> , 10, R108	18.3	118
83	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> , <b>1990</b> , 39, 345		113
82	Statistical tests of gamma-distributed rate heterogeneity in models of sequence evolution in phylogenetics. <i>Molecular Biology and Evolution</i> , <b>2000</b> , 17, 975-8	8.3	103
81	Different versions of the Dayhoff rate matrix. <i>Molecular Biology and Evolution</i> , <b>2005</b> , 22, 193-9	8.3	102
80	Using evolutionary trees in protein secondary structure prediction and other comparative sequence analyses. <i>Journal of Molecular Biology</i> , <b>1996</b> , 263, 196-208	6.5	96

79	Phylogenetic quantification of intra-tumour heterogeneity. <i>PLoS Computational Biology</i> , <b>2014</b> , 10, e1003535	5.5	91
78	Accurate extension of multiple sequence alignments using a phylogeny-aware graph algorithm. <i>Bioinformatics</i> , <b>2012</b> , 28, 1684-91	7.2	88
77	SMIM1 underlies the Vel blood group and influences red blood cell traits. <i>Nature Genetics</i> , <b>2013</b> , 45, 542-545	5.9	77
76	Phylogenetic information and experimental design in molecular systematics. <i>Proceedings of the Royal Society B: Biological Sciences</i> , <b>1998</b> , 265, 1779-86	4.4	76
75	Estimating the frequency of events that cause multiple-nucleotide changes. <i>Genetics</i> , <b>2004</b> , 167, 2027-43	4	69
74	Nucleotide, dinucleotide and trinucleotide frequencies explain patterns observed in chaos game representations of DNA sequences. <i>Nucleic Acids Research</i> , <b>1993</b> , 21, 2487-91	20.1	68
73	Simple diagnostic statistical tests of models for DNA substitution. <i>Journal of Molecular Evolution</i> , <b>1993</b> , 37, 650-61	3.1	68
72	Evolutionary footprints of nucleosome positions in yeast. <i>Trends in Genetics</i> , <b>2008</b> , 24, 583-7	8.5	64
71	Variation in evolutionary processes at different codon positions. <i>Molecular Biology and Evolution</i> , <b>2007</b> , 24, 513-21	8.3	61
70	METHODS FOR DISCRETE CODING OF MORPHOLOGICAL CHARACTERS FOR NUMERICAL ANALYSIS.. <i>Cladistics</i> , <b>1988</b> , 4, 59-71	3.5	56
69	PANDIT: an evolution-centric database of protein and associated nucleotide domains with inferred trees. <i>Nucleic Acids Research</i> , <b>2006</b> , 34, D327-31	20.1	50
68	Experimental design criteria in phylogenetics: where to add taxa. <i>Systematic Biology</i> , <b>2007</b> , 56, 609-22	8.4	49
67	Species choice for comparative genomics: being greedy works. <i>PLoS Genetics</i> , <b>2005</b> , 1, e71	6	48
66	A novel use of equilibrium frequencies in models of sequence evolution. <i>Molecular Biology and Evolution</i> , <b>2002</b> , 19, 1821-31	8.3	45
65	Amino acid changes in disease-associated variants differ radically from variants observed in the 1000 genomes project dataset. <i>PLoS Computational Biology</i> , <b>2013</b> , 9, e1003382	5	44
64	A new criterion and method for amino acid classification. <i>Journal of Theoretical Biology</i> , <b>2004</b> , 228, 97-106	6.3	43
63	Further results on error minimization in the genetic code. <i>Journal of Molecular Evolution</i> , <b>1993</b> , 37, 662-43	3.1	43
62	Stability of SARS-CoV-2 phylogenies. <i>PLoS Genetics</i> , <b>2020</b> , 16, e1009175	6	43

61	Alignment Modulates Ancestral Sequence Reconstruction Accuracy. <i>Molecular Biology and Evolution</i> , <b>2018</b> , 35, 1783-1797	8.3	40
60	XRate: a fast prototyping, training and annotation tool for phylo-grammars. <i>BMC Bioinformatics</i> , <b>2006</b> , 7, 428	3.6	40
59	Pandit: a database of protein and associated nucleotide domains with inferred trees. <i>Bioinformatics</i> , <b>2003</b> , 19, 1556-63	7.2	40
58	A penalized-likelihood method to estimate the distribution of selection coefficients from phylogenetic data. <i>Genetics</i> , <b>2014</b> , 197, 257-71	4	39
57	Phylogenetic analysis of the rpoB gene from the plastid-like DNA of Plasmodium falciparum. <i>Molecular and Biochemical Parasitology</i> , <b>1994</b> , 66, 221-31	1.9	38
56	Maximum Likelihood Trees from DNA Sequences: A Peculiar Statistical Estimation Problem. <i>Systematic Biology</i> , <b>1995</b> , 44, 384	8.4	37
55	More on the Best Evolutionary Rate for Phylogenetic Analysis. <i>Systematic Biology</i> , <b>2017</b> , 66, 769-785	8.4	36
54	The human blood DNA methylome displays a highly distinctive profile compared with other somatic tissues. <i>Epigenetics</i> , <b>2015</b> , 10, 274-81	5.7	36
53	PhyloSim - Monte Carlo simulation of sequence evolution in the R statistical computing environment. <i>BMC Bioinformatics</i> , <b>2011</b> , 12, 104	3.6	36
52	A model of evolution and structure for multiple sequence alignment. <i>Philosophical Transactions of the Royal Society B: Biological Sciences</i> , <b>2008</b> , 363, 3913-9	5.8	36
51	Clustering Genes of Common Evolutionary History. <i>Molecular Biology and Evolution</i> , <b>2016</b> , 33, 1590-605	8.3	35
50	In defense of statistical methods for detecting positive selection. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , <b>2009</b> , 106, E95; author reply E96	11.5	34
49	Phylogenomics and bioinformatics of SARS-CoV. <i>Trends in Microbiology</i> , <b>2004</b> , 12, 106-11	12.4	33
48	Who watches the watchmen? An appraisal of benchmarks for multiple sequence alignment. <i>Methods in Molecular Biology</i> , <b>2014</b> , 1079, 59-73	1.4	33
47	Resource-aware taxon selection for maximizing phylogenetic diversity. <i>Systematic Biology</i> , <b>2007</b> , 56, 431-44	8.4	32
46	Variance to mean ratio, R(t), for poisson processes on phylogenetic trees. <i>Molecular Phylogenetics and Evolution</i> , <b>1994</b> , 3, 230-9	4.1	28
45	Want to track pandemic variants faster? Fix the bioinformatics bottleneck. <i>Nature</i> , <b>2021</b> , 591, 30-33	50.4	27
44	Looking for Darwin in genomic sequences--validity and success of statistical methods. <i>Molecular Biology and Evolution</i> , <b>2012</b> , 29, 2889-93	8.3	26

43	Mutation Rates and Selection on Synonymous Mutations in SARS-CoV-2. <i>Genome Biology and Evolution</i> , <b>2021</b> , 13,	3.9	25
42	All Your Base: a fast and accurate probabilistic approach to base calling. <i>Genome Biology</i> , <b>2012</b> , 13, R13	18.3	24
41	Evolution. Uniting alignments and trees. <i>Science</i> , <b>2009</b> , 324, 1528-9	33.3	22
40	Addressing inter-gene heterogeneity in maximum likelihood phylogenomic analysis: yeasts revisited. <i>PLoS ONE</i> , <b>2011</b> , 6, e22783	3.7	20
39	Determination and validation of principal gene products. <i>Bioinformatics</i> , <b>2008</b> , 24, 11-7	7.2	20
38	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> , <b>2020</b> , 117, 5977-5986	11.5	18
37	Maximum Likelihood Phylogenetic Inference is Consistent on Multiple Sequence Alignments, with or without Gaps. <i>Systematic Biology</i> , <b>2016</b> , 65, 328-33	8.4	16
36	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> , <b>2015</b> , 112, E99-100	11.5	16
35	Maximum likelihood inference of small trees in the presence of long branches. <i>Systematic Biology</i> , <b>2014</b> , 63, 798-811	8.4	16
34	Markovian and non-Markovian protein sequence evolution: aggregated Markov process models. <i>Journal of Molecular Biology</i> , <b>2011</b> , 411, 910-23	6.5	16
33	Genomic reconstruction of the SARS-CoV-2 epidemic in England. <i>Nature</i> , <b>2021</b> ,	50.4	16
32	Rapidly evolving human promoter regions. <i>Nature Genetics</i> , <b>2008</b> , 40, 1262-3; author reply 1263-4	36.3	15
31	Effects of sequence alignment procedures on estimates of phylogeny. <i>BioEssays</i> , <b>1998</b> , 20, 287-290	4.1	14
30	Short template switch events explain mutation clusters in the human genome. <i>Genome Research</i> , <b>2017</b> , 27, 1039-1049	9.7	13
29	Sequence Bundles: a novel method for visualising, discovering and exploring sequence motifs. <i>BMC Proceedings</i> , <b>2014</b> , 8, S8	2.3	12
28	Are big trees indeed easy?. <i>Trends in Ecology and Evolution</i> , <b>1997</b> , 12, 357	10.9	12
27	Modeling Mitochondrial Protein Evolution Using Structural Information. <i>Journal of Molecular Evolution</i> , <b>2002</b> , 54, 519-529	3.1	12
26	Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk. <i>PLoS Computational Biology</i> , <b>2021</b> , 17, e1008561	5	12

25	A Daily-Updated Database and Tools for Comprehensive SARS-CoV-2 Mutation-Annotated Trees. <i>Molecular Biology and Evolution</i> , <b>2021</b> , 38, 5819-5824	8.3	12
24	What's in a likelihood? Simple models of protein evolution and the contribution of structurally viable reconstructions to the likelihood. <i>Systematic Biology</i> , <b>2011</b> , 60, 161-74	8.4	11
23	Introduction. Statistical and computational challenges in molecular phylogenetics and evolution. <i>Philosophical Transactions of the Royal Society B: Biological Sciences</i> , <b>2008</b> , 363, 3889-92	5.8	10
22	Error-correcting properties of the SOLiD Exact Call Chemistry. <i>BMC Bioinformatics</i> , <b>2012</b> , 13, 145	3.6	9
21	Genomic reconstruction of the SARS-CoV-2 epidemic in England		9
20	An improved protocol for sequencing of repetitive genomic regions and structural variations using mutagenesis and next generation sequencing. <i>PLoS ONE</i> , <b>2012</b> , 7, e43359	3.7	7
19	ALVIS: interactive non-aggregative visualization and explorative analysis of multiple sequence alignments. <i>Nucleic Acids Research</i> , <b>2016</b> , 44, e77	20.1	6
18	Mutation rates and selection on synonymous mutations in SARS-CoV-2 <b>2021</b> ,		6
17	Modeling Structural Constraints on Protein Evolution via Side-Chain Conformational States. <i>Molecular Biology and Evolution</i> , <b>2019</b> , 36, 2086-2103	8.3	5
16	Fewest Variables Coding Method for Multistate Characters. <i>Systematic Zoology</i> , <b>1989</b> , 38, 79		5
15	Modeling mitochondrial protein evolution using structural information. <i>Journal of Molecular Evolution</i> , <b>2002</b> , 54, 519-29	3.1	4
14	Improving communication for interdisciplinary teams working on storage of digital information in DNA. <i>F1000Research</i> , <b>2018</b> , 7, 39	3.6	4
13	A daily-updated database and tools for comprehensive SARS-CoV-2 mutation-annotated trees <b>2021</b> ,		4
12	Avoiding ascertainment bias in the maximum likelihood inference of phylogenies based on truncated data		3
11	Short-range template switching in great ape genomes explored using pair hidden Markov models. <i>PLoS Genetics</i> , <b>2021</b> , 17, e1009221	6	3
10	Genetic Variability of the SARS-CoV-2 Pocketome. <i>Journal of Proteome Research</i> , <b>2021</b> , 20, 4212-4215	5.6	2
9	phastSim: efficient simulation of sequence evolution for pandemic-scale datasets <b>2021</b> ,		1
8	Probabilistic Models for the Study of Protein Evolution <b>2019</b> , 347-30		1

- 7 A phylogenetic approach for weighting genetic sequences. *BMC Bioinformatics*, **2021**, 22, 285 3.6 ○
- 6 phastSim: Efficient simulation of sequence evolution for pandemic-scale datasets.. *PLoS Computational Biology*, **2022**, 18, e1010056 5 ○
- 5 Ambiguity Coding Allows Accurate Inference of Evolutionary Parameters from Alignments in an Aggregated State-Space. *Systematic Biology*, **2021**, 70, 21-32 8.4
- 4 Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk **2021**, 17, e1008561
- 3 Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk **2021**, 17, e1008561
- 2 Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk **2021**, 17, e1008561
- 1 Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk **2021**, 17, e1008561