Richard M Durbin

List of Publications by Year
in descending order

221 papers	244,327 citations	1094 112 h-index	1341 223 g-index
280 all docs	280 docs citations	280 times ranked	204813 citing authors

#	Article	IF	Citations
1	The Sequence Alignment/Map format and SAMtools. Bioinformatics, 2009, 25, 2078-2079.	1.8	49,124
2	Fast and accurate short read alignment with Burrowsâ€“Wheeler transform. Bioinformatics, 2009, 25, 1754-1760.	1.8	43,062
3	Initial sequencing and analysis of the human genome. Nature, 2001, 409, 860-921.	13.7	21,074
4	A global reference for human genetic variation. Nature, 2015, 526, 68-74.	13.7	13,998
5	The variant call format and VCFtools. Bioinformatics, 2011, 27, 2156-2158.	1.8	11,326
6	Fast and accurate long-read alignment with Burrowsâ€“Wheeler transform. Bioinformatics, 2010, 26, 589-595.	1.8	10,002
7	A map of human genome variation from population-scale sequencing. Nature, 2010, 467, 1061-1073.	13.7	7,209
8	An integrated map of genetic variation from 1,092 human genomes. Nature, 2012, 491, 56-65.	13.7	7,199
9	Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature, 2003, 421, 231-237.	13.7	3,343
10	Accurate whole human genome sequencing using reversible terminator chemistry. Nature, 2008, 456, 53-59.	13.7	3,118
11	The Pfam protein families database. Nucleic Acids Research, 2004, 32, 138D-141.	6.5	3,084
12	A reference panel of 64,976 haplotypes for genotype imputation. Nature Genetics, 2016, 48, 1279-1283.	9.4	2,421
13	Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Research, 2008, 18, 1851-1858.	2.4	2,275
14	GeneWise and Genomewise. Genome Research, 2004, 14, 988-995.	2.4	2,128
15	The Pfam Protein Families Database. Nucleic Acids Research, 2002, 30, 276-280.	6.5	2,067
16	Inference of human population history from individual whole-genome sequences. Nature, 2011, 475, 493-496.	13.7	2,053
17	Pfam: clans, web tools and services. Nucleic Acids Research, 2006, 34, D247-D251.	6.5	2,030
18	Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature, 2004, 428, 493-521.	13.7	1,943

#	Article	IF	Citations
19	Mouse genomic variation and its effect on phenotypes and gene regulation. Nature, 2011, 477, 289-294.	13.7	1,461
20	Population genomics of domestic and wild yeasts. Nature, 2009, 458, 337-341.	13.7	1,391
21	Reference-based phasing using the Haplotype Reference Consortium panel. Nature Genetics, 2016, 48, 1443-1448.	9.4	1,357
22	The complete sequence of a human genome. Science, 2022, 376, 44-53.	6.0	1,222
23	Identifying and removing haplotypic duplication in primary genome assemblies. Bioinformatics, 2020, 36, 2896-2898.	1.8	1,221
24	The Pfam Protein Families Database. Nucleic Acids Research, 2000, 28, 263-266.	6.5	1,173
25	Towards complete and error-free genome assemblies of all vertebrate species. Nature, 2021, 592, 737-746.	13.7	1,139
26	EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. Genome Research, 2009, 19, 327-335.	2.4	1,058
27	Pfam: A comprehensive database of protein domain families based on seed alignments. , 1997, 28, 405-420.		1,036
28	The UK10K project identifies rare variants in health and disease. Nature, 2015, 526, 82-90.	13.7	1,014
29	The DNA sequence of the human X chromosome. Nature, 2005, 434, 325-337.	13.7	985
30	Inferring human population size and separation history from multiple genome sequences. Nature Genetics, 2014, 46, 919-925.	9.4	870
31	Ensembl 2012. Nucleic Acids Research, 2012, 40, D84-D90.	6.5	840
32	The diploid genome sequence of an Asian individual. Nature, 2008, 456, 60-65.	13.7	834
33	The Genome Sequence of Caenorhabditis briggsae: A Platform for Comparative Genomics. PLoS Biology, 2003, 1, e45.	2.6	812
34	Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. Nature Protocols, 2012, 7, 500-507.	5.5	799
35	Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. Nature, 2010, 464, 721-727.	13.7	768
36	An analogue approach to the travelling salesman problem using an elastic net method. Nature, 1987, 326, 689-691.	13.7	740

#	Article	IF	Citations
37	Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. Nature Genetics, 2008, 40, 722-729.	9.4	736
38	Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. Genome Research, 2017, 27, 849-864.	2.4	728
39	RNA sequence analysis using covariance models. Nucleic Acids Research, 1994, 22, 2079-2088.	6.5	727
40	Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nature Genetics, 2012, 44, 1084-1089.	9.4	701
41	A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene, 1995, 167, GC1-GC10.	1.0	677
42	Insights into hominid evolution from the gorilla genome sequence. Nature, 2012, 483, 169-175.	13.7	663
43	A large genome center's improvements to the Illumina sequencing system. Nature Methods, 2008, 5, 1005-1010.	9.0	656
44	Earth BioGenome Project: Sequencing life for the future of life. Proceedings of the National Academy of Sciences of the United States of America, 2018, 115, 4325-4333.	3.3	652
45	Efficient de novo assembly of large genomes using compressed data structures. Genome Research, 2012, 22, 549-556.	2.4	649
46	The InterPro Database, 2003 brings increased coverage and new features. Nucleic Acids Research, 2003, 31, 315-318.	6.5	640
47	The Sequence Ontology: a tool for the unification of genome annotations. Genome Biology, 2005, 6, R44.	13.9	638
48	Ensembl 2011. Nucleic Acids Research, 2011, 39, D800-D806.	6.5	630
49	A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nature Biotechnology, 2008, 26, 779-785.	9.4	619
50	WormBase: a multi-species resource for nematode biology and genomics. Nucleic Acids Research, 2004, 32, 411D-417.	6.5	610
51	Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. GigaScience, 2013, 2, 10.	3.3	582
52	BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals. Nature Methods, 2008, 5, 409-415.	9.0	568
53	Insights into human genetic variation and population history from 929 diverse genomes. Science, 2020, 367, .	6.0	534
54	BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. Bioinformatics, 2016, 32, 1749-1751.	1.8	506

#	Article	IF	Citations
55	Common genetic variation drives molecular heterogeneity in human iPSCs. Nature, 2017, 546, 370-375.	13.7	491
56	Revising the human mutation rate: implications for understanding human evolution. Nature Reviews Genetics, 2012, 13, 745-753.	7.7	483
57	Wholeâ€genome sequencing identifies EN1 as a determinant of bone density and fracture. Nature, 2015, 526, 112-117.	13.7	483
58	InterPro, progress and status in 2005. Nucleic Acids Research, 2004, 33, D201-D205.	6.5	478
59	The consensus coding sequence (CCDS) project: Identifying a common protein-coding gene set for the human and mouse genomes. Genome Research, 2009, 19, 1316-1323.	2.4	476
60	TreeFam: a curated database of phylogenetic trees of animal gene families. Nucleic Acids Research, 2006, 34, D572-D580.	6.5	465
61	Systematic Analysis of Human Protein Complexes Identifies Chromosome Segregation Proteins. Science, 2010, 328, 593-599.	6.0	465
62	Genomic evidence for the Pleistocene and recent population history of Native Americans. Science, 2015, 349, aab3884.	6.0	449
63	Assemblathon 1: A competitive assessment of de novo short read assembly methods. Genome Research, 2011, 21, 2224-2241.	2.4	443
64	Patterns of Cis Regulatory Variation in Diverse Human Populations. PLoS Genetics, 2012, 8, e1002639.	1.5	439
65	A genomic history of Aboriginal Australia. Nature, 2016, 538, 207-214.	13.7	439
66	Variation graph toolkit improves read mapping by representing genetic variation in the reference. Nature Biotechnology, 2018, 36, 875-879.	9.4	435
67	A dimension reduction framework for understanding cortical maps. Nature, 1990, 343, 644-647.	13.7	422
68	A Bayesian Framework to Account for Complex Non-Genetic Factors in Gene Expression Levels Greatly Increases Power in eQTL Studies. PLoS Computational Biology, 2010, 6, e1000770.	1.5	408
69	The Architecture of Gene Regulatory Variation across Multiple Human Tissues: The MuTHER Study. PLoS Genetics, 2011, 7, e1002003.	1.5	392
70	An Overview of Ensembl. Genome Research, 2004, 14, 925-928.	2.4	391
71	Efficient haplotype matching and storage using the positional Burrowsâ€“Wheeler transform (PBWT). Bioinformatics, 2014, 30, 1266-1272.	1.8	387
72	Dindel: Accurate indel calls from short-read data. Genome Research, 2011, 21, 961-973.	2.4	383

#	Article	IF	Citations
73	Product Units: A Computationally Powerful and Biologically Plausible Extension to Backpropagation Networks. Neural Computation, 1989, 1, 133-142.	1.3	378
74	Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. Nature Ecology and Evolution, 2018, 2, 1940-1955.	3.4	358
75	Distribution and Medical Impact of Loss-of-Function Variants in the Finnish Founder Population. PLoS Genetics, 2014, 10, e1004494.	1.5	351
76	Contrasting evolutionary genome dynamics between domesticated and wild yeasts. Nature Genetics, 2017, 49, 913-924.	9.4	340
77	Using GeneWise in the Drosophila Annotation Experiment. Genome Research, 2000, 10, 547-548.	2.4	338
78	Global Analysis of DNA Methylation Variation in Adipose Tissue from Twins Reveals Links to Disease-Associated Variants in Distal Regulatory Elements. American Journal of Human Genetics, 2013, 93, 876-890.	2.6	330
79	Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. Science, 2015, 350, 1493-1498.	6.0	330
80	A High-Definition View of Functional Genetic Variation from Natural Yeast Genomes. Molecular Biology and Evolution, 2014, 31, 872-888.	3.5	328
81	WormBase: a comprehensive resource for nematode research. Nucleic Acids Research, 2010, 38, D463-D467.	6.5	325
82	Did Our Species Evolve in Subdivided Populations across Africa, and Why Does It Matter?. Trends in Ecology and Evolution, 2018, 33, 582-594.	4.2	315
83	Trait Variation in Yeast Is Defined by Population History. PLoS Genetics, 2011, 7, e1002111.	1.5	311
84	Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. Nature Communications, 2015, 6, 8111.	5.8	300
85	TreeFam: 2008 Update. Nucleic Acids Research, 2007, 36, D735-D740.	6.5	294
86	Health and population effects of rare gene knockouts in adult humans with related parents. Science, 2016, 352, 474-477.	6.0	272
87	Revealing the genetic structure of a trait by sequencing a population under selection. Genome Research, 2011, 21, 1131-1138.	2.4	263
88	Gene expression changes with age in skin, adipose tissue, blood and brain. Genome Biology, 2013, 14, R75.	13.9	263
89	The first horse herders and the impact of early Bronze Age steppe expansions into Asia. Science, 2018, 360, .	6.0	262
90	The population history of northeastern Siberia since the Pleistocene. Nature, 2019, 570, 182-188.	13.7	259

#	Article	IF	Citations
91	Ensembl's 10th year. Nucleic Acids Research, 2010, 38, D557-D562.	6.5	251
92	QuickTree: building huge Neighbour-Joining trees of protein sequences. Bioinformatics, 2002, 18, 1546-1547.	1.8	249
93	Structure and expression of the Huntington's disease gene: Evidence against simple inactivation due to an expanded CAG repeat. Somatic Cell and Molecular Genetics, 1994, 20, 27-38.	0.7	246
94	Prepublication data sharing. Nature, 2009, 461, 168-170.	13.7	243
95	Souporcell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. Nature Methods, 2020, 17, 615-620.	9.0	232
96	Maximum Discrimination Hidden Markov Models of Sequence Consensus. Journal of Computational Biology, 1995, 2, 9-23.	0.8	218
97	Gene-gene and gene-environment interactions detected by transcriptome sequence analysis in twins. Nature Genetics, 2015, 47, 88-91.	9.4	215
98	Efficient construction of an assembly string graph using the FM-index. Bioinformatics, 2010, 26, i367-i373.	1.8	198
99	Comparative Analysis of Noncoding Regions of 77 Orthologous Mouse and Human Gene Pairs. Genome Research, 1999, 9, 815-824.	2.4	180
100	Deficient methylation and formylation of mt-tRNAMet wobble cytosine in a patient carrying mutations in NSUN3. Nature Communications, 2016, 7, 12039.	5.8	178
101	WormBase 2012: more genomes, more data, new website. Nucleic Acids Research, 2012, 40, D735-D741.	6.5	175
102	The genomic and phenotypic diversity of Schizosaccharomyces pombe. Nature Genetics, 2015, 47, 235-241.	9.4	174
103	Sixteen diverse laboratory mouse reference genomes define strain-specific haplotypes and novel functional loci. Nature Genetics, 2018, 50, 1574-1583.	9.4	169
104	An Analysis of the Elastic Net Approach to the Traveling Salesman Problem. Neural Computation, 1989, 1, 348-358.	1.3	168
105	High-Resolution Mapping of Complex Traits with a Four-Parent Advanced Intercross Yeast Population. Genetics, 2013, 195, 1141-1155.	1.2	164
106	InterPro: An integrated documentation resource for protein families, domains and functional sites. Briefings in Bioinformatics, 2002, 3, 225-235.	3.2	155
107	WormBase: a comprehensive data resource for Caenorhabditis biology and genomics. Nucleic Acids Research, 2004, 33, D383-D389.	6.5	155
108	Tracing the Route of Modern Humans out of Africa by Using 225 Human Genome Sequences from Ethiopians and Egyptians. American Journal of Human Genetics, 2015, 96, 986-991.	2.6	152

#	Article	IF	Citations
109	Estimating telomere length from whole genome sequence data. Nucleic Acids Research, 2014, 42, e75-e75.	6.5	151
110	High levels of RNA-editing site conservation amongst 15 laboratory mouse strains. Genome Biology, 2012, 13, R26.	13.9	149
111	De novo yeast genome assemblies from MinION, PacBio and MiSeq platforms. Scientific Reports, 2017, 7, 3935.	1.6	146
112	Comparative sequence analysis of the human and pufferfish Huntington's disease genes. Nature Genetics, 1995, 10, 67-76.	9.4	144
113	Iron Age and Anglo-Saxon genomes from East England reveal British migration history. Nature Communications, 2016, 7, 10408.	5.8	144
114	SNP detection and genotyping from low-coverage sequencing data on multiple diploid samples. Genome Research, 2011, 21, 952-960.	2.4	142
115	Extending reference assembly models. Genome Biology, 2015, 16, 13.	3.8	139
116	Genetic interactions affecting human gene expression identified by variance association mapping. ELife, 2014, 3, e01381.	2.8	137
117	Whole-Genome Sequencing Coupled to Imputation Discovers Genetic Signals for Anthropometric Traits. American Journal of Human Genetics, 2017, 100, 865-884.	2.6	131
118	Analysis of Protein Domain Families inCaenorhabditis elegans. Genomics, 1997, 46, 200-216.	1.3	129
119	A computational scan for U12-dependent introns in the human genome sequence. Nucleic Acids Research, 2001, 29, 4006-4013.	6.5	127
120	The Earth BioGenome Project 2020: Starting the clock. Proceedings of the National Academy of Sciences of the United States of America, 2022, 119, .	3.3	124
121	A High-Quality De novo Genome Assembly from a Single Mosquito Using PacBio Sequencing. Genes, 2019, 10, 62.	1.0	121
122	Immunofluorescence Analysis and Diagnosis of Primary Ciliary Dyskinesia with Radial Spoke Defects. American Journal of Respiratory Cell and Molecular Biology, 2015, 53, 563-573.	1.4	120
123	Sequence locally, think globally: The Darwin Tree of Life Project. Proceedings of the National Academy of Sciences of the United States of America, 2022, 119, .	3.3	120
124	Comparative ab initio prediction of gene structures using pair HMMs. Bioinformatics, 2002, 18, 1309-1318.	1.8	114
125	WormBase: a cross-species database for comparative genomics. Nucleic Acids Research, 2003, 31, 133-137.	6.5	107
126	DNAH11 Localization in the Proximal Region of Respiratory Cilia Defines Distinct Outer Dynein Arm Complexes. American Journal of Respiratory Cell and Molecular Biology, 2016, 55, 213-224.	1.4	107

#	Article	IF	Citations
127	Mapping Trait Loci by Use of Inferred Ancestral Recombination Graphs. American Journal of Human Genetics, 2006, 79, 910-922.	2.6	105
128	Ancestral Hybridization Facilitated Species Diversification in the Lake Malawi Cichlid Fish Adaptive Radiation. Molecular Biology and Evolution, 2020, 37, 1100-1113.	3.5	98
129	Dynamic Programming Alignment Accuracy. Journal of Computational Biology, 1998, 5, 493-504.	0.8	97
130	WormBase 2007. Nucleic Acids Research, 2007, 36, D612-D617.	6.5	95
131	Optimal Numberings of an $N imes N$ Array. SIAM Journal on Algebraic and Discrete Methods, 1986, 7, 571-582.	0.8	93
132	Deficiency of <scp>ECHS</scp>1 causes mitochondrial encephalopathy with cardiac involvement. Annals of Clinical and Translational Neurology, 2015, 2, 492-509.	1.7	90
133	No evidence for maintenance of a sympatricHeliconiusspecies barrier by chromosomal inversions. Evolution Letters, 2017, 1, 138-154.	1.6	90
134	Efficiently Inferring the Demographic History of Many Populations With Allele Count Data. Journal of the American Statistical Association, 2020, 115, 1472-1487.	1.8	90
135	TTC25 Deficiency Results in Defects of the Outer Dynein Arm Docking Machinery and Primary Ciliary Dyskinesia with Left-Right Body Asymmetry Randomization. American Journal of Human Genetics, 2016, 99, 460-469.	2.6	88
136	Software for genome mapping by fingerprinting techniques. Bioinformatics, 1988, 4, 125-132.	1.8	87
137	GAZE: A Generic Framework for the Integration of Gene-Prediction Data by Dynamic Programming. Genome Research, 2002, 12, 1418-1427.	2.4	82
138	Estimating the human mutation rate from autozygous segments reveals population differences in human mutational processes. Nature Communications, 2017, 8, 303.	5.8	81
139	Late Quaternary dynamics of Arctic biota from ancient environmental genomics. Nature, 2021, 600, 86-92.	13.7	81
140	WormBase: new content and better access. Nucleic Acids Research, 2007, 35, D506-D510.	6.5	80
141	Gene structure conservation aids similarity based gene prediction. Nucleic Acids Research, 2004, 32, 776-783.	6.5	79
142	Detecting archaic introgression using an unadmixed outgroup. PLoS Genetics, 2018, 14, e1007641.	1.5	78
143	Bi-allelic Truncating Mutations in TANGO2 Cause Infancy-Onset Recurrent Metabolic Crises with Encephalocardiomyopathy. American Journal of Human Genetics, 2016, 98, 358-362.	2.6	77
144	Joint Genetic Analysis of Gene Expression Data with Inferred Cellular Phenotypes. PLoS Genetics, 2011, 7, e1001276.	1.5	76

#	Article	IF	Citations
145	Whole-genome sequence-based analysis of thyroid function. Nature Communications, 2015, 6, 5681.	5.8	75
146	WormBase: better software, richer content. Nucleic Acids Research, 2006, 34, D475-D478.	6.5	74
147	Clustering of phosphorylation site recognition motifs can be exploited to predict the targets of cyclin-dependent kinase. Genome Biology, 2007, 8, R23.	13.9	74
148	Human genomic regions with exceptionally high levels of population differentiation identified from 911 whole-genome sequences. Genome Biology, 2014, 15, R88.	13.9	72
149	A Workbench for large-scale sequence homology analysis. Bioinformatics, 1994, 10, 301-307.	1.8	69
150	Complete vertebrate mitogenomes reveal widespread repeats and gene duplications. Genome Biology, 2021, 22, 120.	3.8	69
151	Enrichment of low-frequency functional variants revealed by whole-genome sequencing of multiple isolated European populations. Nature Communications, 2017, 8, 15927.	5.8	64
152	Image analysis of restriction enzyme fingerprint autoradiograms. Bioinformatics, 1989, 5, 101-106.	1.8	62
153	Regulatory evolution in proteins by turnover and lineage-specific changes of cyclin-dependent kinase consensus sites. Proceedings of the National Academy of Sciences of the United States of America, 2007, 104, 17713-17718.	3.3	62
154	A systematic comparative and structural analysis of protein phosphorylation sites based on the mtcPTM database. Genome Biology, 2007, 8, R90.	13.9	62
155	A rare variant in APOC3 is associated with plasma triglyceride and VLDL levels in Europeans. Nature Communications, 2014, 5, 4871.	5.8	62
156	Whole-genome view of the consequences of a population bottleneck using 2926 genome sequences from Finland and United Kingdom. European Journal of Human Genetics, 2017, 25, 477-484.	1.4	60
157	A graph-based approach to diploid genome assembly. Bioinformatics, 2018, 34, i105-i114.	1.8	59
158	Haplotype-aware graph indexes. Bioinformatics, 2020, 36, 400-407.	1.8	59
159	Quantitative Genetics of CTCF Binding Reveal Local Sequence Effects and Different Modes of X-Chromosome Association. PLoS Genetics, 2014, 10, e1004798.	1.5	55
160	Why sequence all eukaryotes?. Proceedings of the National Academy of Sciences of the United States of America, 2022, 119, .	3.3	51
161	A probabilistic model of 3' end formation in Caenorhabditis elegans. Nucleic Acids Research, 2004, 32, 3392-3399.	6.5	50
162	Extent, Causes, and Consequences of Small RNA Expression Variation in Human Adipose Tissue. PLoS Genetics, 2012, 8, e1002704.	1.5	48

#	Article	IF	Citations
163	Enhanced protein domain discovery by using language modeling techniques from speech recognition. Proceedings of the National Academy of Sciences of the United States of America, 2003, 100, 4516-4520.	3.3	47
164	Copy number variant detection in inbred strains from short read sequence data. Bioinformatics, 2010, 26, 565-567.	1.8	47
165	Removing reference bias and improving indel calling in ancient DNA data analysis by mapping to a sequence variation graph. Genome Biology, 2020, 21, 250.	3.8	44
166	Environmental genomics of Late Pleistocene black bears and giant short-faced bears. Current Biology, 2021, 31, 2728-2736.e8.	1.8	42
167	A high-content platform to characterise human induced pluripotent stem cell lines. Methods, 2016, 96, 85-96.	1.9	41
168	Population-scale proteome variation in human induced pluripotent stem cells. ELife, 2020, 9, .	2.8	40
169	Managing clinically significant findings in research: the UK10K example. European Journal of Human Genetics, 2014, 22, 1100-1104.	1.4	38
170	A high-quality, chromosome-level genome assembly of the Black Soldier Fly (Hermetia illucens) Tj ETQq0 0 0 rgBT /Overlock 10 Tf	0.8	37
171	Deep short-read sequencing of chromosome 17 from the mouse strains A/J and CAST/Ei identifies significant germline variation and candidate genes that regulate liver triglyceride levels. Genome Biology, 2009, 10, R112.	13.9	36
172	Identifying Extrinsic versus Intrinsic Drivers of Variation in Cell Behavior in Human iPSC Lines from Healthy Donors. Cell Reports, 2019, 26, 2078-2087.e3.	2.9	36
173	Alfresco---A Workbench for Comparative Genomic Sequence Analysis. Genome Research, 2000, 10, 1148-1157.	2.4	35
174	Standards recommendations for the Earth BioGenome Project. Proceedings of the National Academy of Sciences of the United States of America, 2022, 119, .	3.3	33
175	The ACEDB Genome Database. , 1994, , 45-55.		32
176	Identityâ€byâ€descentâ€based phasing and imputation in founder populations using graphical models. Genetic Epidemiology, 2011, 35, 853-860.	0.6	31
177	Estimation of Epistatic Variance Components and Heritability in Founder Populations and Crosses. Genetics, 2014, 198, 1405-1416.	1.2	27
178	Homozygous loss-of-function variants in European cosmopolitan and isolate populations. Human Molecular Genetics, 2015, 24, 5464-5474.	1.4	27
179	Whole-exome sequencing of 228 patients with sporadic Parkinsonâ€™s disease. Scientific Reports, 2017, 7, 41188.	1.6	27
180	A Method for Checking Genomic Integrity in Cultured Cell Lines from SNP Genotyping Data. PLoS ONE, 2016, 11, e0155014.	1.1	26

#	Article	IF	Citations
181	The anatomy of successful computational biology software. Nature Biotechnology, 2013, 31, 894-897.	9.4	25
182	A Genome-Wide Survey of Genetic Variation in Gorillas Using Reduced Representation Sequencing. PLoS ONE, 2013, 8, e65066.	1.1	23
183	htsget: a protocol for securely streaming genomic data. Bioinformatics, 2019, 35, 119-121.	1.8	23
184	Placing Ancient DNA Sequences into Reference Phylogenies. Molecular Biology and Evolution, 2022, 39, .	3.5	23
185	A haplotype-awarede novoassembly of related individuals using pedigree sequence graph. Bioinformatics, 2020, 36, 2385-2392.	1.8	22
186	The C. elegans expression pattern database: a beginning. Trends in Genetics, 1996, 12, 370-371.	2.9	21
187	Crumble: reference free lossy compression of sequence quality values. Bioinformatics, 2019, 35, 337-339.	1.8	21
188	A haplotype-resolved,de novogenome assembly for the wood tiger moth (Arctia) Tj ETQq0 0 0 rgBT /Overlock 10 Tf 50 462 Td	3.3	20
189	Genomic consequences of domestication of the Siamese fighting fish. Science Advances, 2022, 8, eabm4950.	4.7	20
190	Using reference-free compressed data structures to analyze sequencing reads from thousands of human genomes. Genome Research, 2017, 27, 300-309.	2.4	19
191	Gene expression and development databases forC. elegans. Seminars in Cell and Developmental Biology, 1997, 8, 459-467.	2.3	18
192	Enhanced protein domain discovery using taxonomy. BMC Bioinformatics, 2004, 5, 56.	1.2	18
193	Mapping epigenetic divergence in the massive radiation of Lake Malawi cichlid fishes. Nature Communications, 2021, 12, 5870.	5.8	17
194	[19] Oscillation method with large unit cells. Methods in Enzymology, 1985, 114, 211-237.	0.4	16
195	Sequence Assembly withâ€‰CAFTOOLS. Genome Research, 1998, 8, 260-267.	2.4	16
196	Accounting for Non-genetic Factors Improves the Power of eQTL Studies. Lecture Notes in Computer Science, 2008, , 411-422.	1.0	16
197	The genome sequence of the brown trout, Salmo trutta Linnaeus 1758. Wellcome Open Research, 2021, 6, 108.	0.9	15
198	Vertebrate gene finding from multiple-species alignments using a two-level strategy. Genome Biology, 2006, 7, S6.	13.9	14

#	Article	IF	Citations
199	WormBase. Worm, 2012, 1, 15-21.	1.0	14
200	Method for Calculation of Probability of Matching a Bounded Regular Expression in a Random Data String. Journal of Computational Biology, 1995, 2, 25-31.	0.8	13
201	Genomix: a method for combining gene-finders' predictions, which uses evolutionary conservation of sequence and intronâ€“exon structure. Bioinformatics, 2007, 23, 1468-1475.	1.8	12
202	Efficient iterative Hi-C scaffolder based on N-best neighbors. BMC Bioinformatics, 2021, 22, 569.	1.2	12
203	Whole-exome sequencing in an isolated population from the Dalmatian island of Vis. European Journal of Human Genetics, 2016, 24, 1479-1487.	1.4	11
204	A Table-Driven, Full-Sensitivity Similarity Search Algorithm. Journal of Computational Biology, 2003, 10, 103-117.	0.8	9
205	Inferring Selection on Amino Acid Preference in Protein Domains. Molecular Biology and Evolution, 2008, 26, 527-536.	3.5	9
206	Base Qualities Help Sequencingâ€‰Software. Genome Research, 1998, 8, 161-162.	2.4	8
207	Ethical, legal, and social issues in the Earth BioGenome Project. Proceedings of the National Academy of Sciences of the United States of America, 2022, 119, e2115859119.	3.3	8
208	A conserved sequence motif in 3' untranslated regions of ribosomal protein mRNAs in nematodes. Rna, 2006, 12, 1786-1789.	1.6	7
209	Pathway-Based Factor Analysis of Gene Expression Data Produces Highly Heritable Phenotypes That Associate with Age. G3: Genes, Genomes, Genetics, 2015, 5, 839-847.	0.8	7
210	The genome sequence of the Eurasian river otter, Lutra lutra Linnaeus 1758. Wellcome Open Research, 2020, 5, 33.	0.9	6
211	The genome sequence of the ringlet, Aphantopus hyperantus Linnaeus 1758. Wellcome Open Research, 0, 6, 165.	0.9	4
212	The C. elegans expression pattern database: a beginning. , 1996, 12, 370-370.		4
213	The genome sequence of the eastern grey squirrel, Sciurus carolinensis Gmelin, 1788. Wellcome Open Research, 2020, 5, 27.	0.9	4
214	GFAKluge: A C++ library and command line utilities for the Graphical Fragment Assembly formats. Journal of Open Source Software, 2019, 4, 1083.	2.0	4
215	Learning from your neighbour. Nature, 1992, 355, 112-113.	13.7	3
216	Viral coinfection analysis using a MinHash toolkit. BMC Bioinformatics, 2019, 20, 389.	1.2	3

#	Article	IF	Citations
217	The genome sequence of the European golden eagle, Aquila chrysaetos chrysaetos Linnaeus 1758. Wellcome Open Research, 2021, 6, 112.	0.9	3
218	The genome sequence of the Eurasian red squirrel, Sciurus vulgaris Linnaeus 1758. Wellcome Open Research, 2020, 5, 18.	0.9	3
219	[X]uniqMAP: unique gene sequence regions in the human and mouse genomes. BMC Genomics, 2006, 7, 249.	1.2	1
220	Abstract 3273: rkmh: A MinHash toolbox for analyzing HPV coinfections. , 2018, , .		0
221	A Theoretical Analysis of Taxonomic Binning Accuracy. Molecular Ecology Resources, 2022, , .	2.2	0