List of Publications by Year in descending order

50 papers	163,051 citations	57758 44 h-index	189892 50 g-index
53 all docs	53 docs citations	53 times ranked	170948 citing authors

#	Article	IF	Citations
1	The Sequence Alignment/Map format and SAMtools. Bioinformatics, 2009, 25, 2078-2079.	4.1	49,124
2	Fast and accurate short read alignment with Burrowsâ€“Wheeler transform. Bioinformatics, 2009, 25, 1754-1760.	4.1	43,062
3	Initial sequencing and analysis of the human genome. Nature, 2001, 409, 860-921.	27.8	21,074
4	Fast and accurate long-read alignment with Burrowsâ€“Wheeler transform. Bioinformatics, 2010, 26, 589-595.	4.1	10,002
5	Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature, 2003, 421, 231-237.	27.8	3,343
6	Accurate whole human genome sequencing using reversible terminator chemistry. Nature, 2008, 456, 53-59.	27.8	3,118
7	The Pfam protein families database. Nucleic Acids Research, 2004, 32, 138D-141.	14.5	3,084
8	A reference panel of 64,976 haplotypes for genotype imputation. Nature Genetics, 2016, 48, 1279-1283.	21.4	2,421
9	Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Research, 2008, 18, 1851-1858.	5.5	2,275
10	GeneWise and Genomewise. Genome Research, 2004, 14, 988-995.	5.5	2,128
11	The Pfam Protein Families Database. Nucleic Acids Research, 2002, 30, 276-280.	14.5	2,067
12	Inference of human population history from individual whole-genome sequences. Nature, 2011, 475, 493-496.	27.8	2,053
13	Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature, 2004, 428, 493-521.	27.8	1,943
14	Mouse genomic variation and its effect on phenotypes and gene regulation. Nature, 2011, 477, 289-294.	27.8	1,461
15	Reference-based phasing using the Haplotype Reference Consortium panel. Nature Genetics, 2016, 48, 1443-1448.	21.4	1,357
16	EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. Genome Research, 2009, 19, 327-335.	5.5	1,058
17	Pfam: A comprehensive database of protein domain families based on seed alignments. Proteins: Structure, Function and Bioinformatics, 1997, 28, 405-420.	2.6	1,036
18	The DNA sequence of the human X chromosome. Nature, 2005, 434, 325-337.	27.8	985

#	Article	IF	Citations
19	Ensembl 2012. Nucleic Acids Research, 2012, 40, D84-D90.	14.5	840
20	The Genome Sequence of Caenorhabditis briggsae: A Platform for Comparative Genomics. PLoS Biology, 2003, 1, e45.	5.6	812
21	Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. Nature Genetics, 2008, 40, 722-729.	21.4	736
22	Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nature Genetics, 2012, 44, 1084-1089.	21.4	701
23	Efficient de novo assembly of large genomes using compressed data structures. Genome Research, 2012, 22, 549-556.	5.5	649
24	The InterPro Database, 2003 brings increased coverage and new features. Nucleic Acids Research, 2003, 31, 315-318.	14.5	640
25	The Sequence Ontology: a tool for the unification of genome annotations. Genome Biology, 2005, 6, R44.	9.6	638
26	Ensembl 2011. Nucleic Acids Research, 2011, 39, D800-D806.	14.5	630
27	A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nature Biotechnology, 2008, 26, 779-785.	17.5	619
28	WormBase: a multi-species resource for nematode biology and genomics. Nucleic Acids Research, 2004, 32, 411D-417.	14.5	610
29	Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. GigaScience, 2013, 2, 10.	6.4	582
30	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.	5.5	476
31	Systematic Analysis of Human Protein Complexes Identifies Chromosome Segregation Proteins. Science, 2010, 328, 593-599.	12.6	465
32	An Overview of Ensembl. Genome Research, 2004, 14, 925-928.	5.5	391
33	Dindel: Accurate indel calls from short-read data. Genome Research, 2011, 21, 961-973.	5.5	383
34	Using GeneWise in the Drosophila Annotation Experiment. Genome Research, 2000, 10, 547-548.	5.5	338
35	Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. Nature Communications, 2015, 6, 8111.	12.8	300
36	Health and population effects of rare gene knockouts in adult humans with related parents. Science, 2016, 352, 474-477.	12.6	272

#	Article	IF	Citations
37	Ensembl's 10th year. Nucleic Acids Research, 2010, 38, D557-D562.	14.5	251
38	Prepublication data sharing. Nature, 2009, 461, 168-170.	27.8	243
39	WormBase: a comprehensive data resource for Caenorhabditis biology and genomics. Nucleic Acids Research, 2004, 33, D383-D389.	14.5	155
40	SNP detection and genotyping from low-coverage sequencing data on multiple diploid samples. Genome Research, 2011, 21, 952-960.	5.5	142
41	Comparative ab initio prediction of gene structures using pair HMMs. Bioinformatics, 2002, 18, 1309-1318.	4.1	114
42	Dynamic Programming Alignment Accuracy. Journal of Computational Biology, 1998, 5, 493-504.	1.6	97
43	GAZE: A Generic Framework for the Integration of Gene-Prediction Data by Dynamic Programming. Genome Research, 2002, 12, 1418-1427.	5.5	82
44	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.	2.8	60
45	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.	7.1	47
46	Copy number variant detection in inbred strains from short read sequence data. Bioinformatics, 2010, 26, 565-567.	4.1	47
47	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.	9.6	36
48	Vertebrate gene finding from multiple-species alignments using a two-level strategy. Genome Biology, 2006, 7, S6.	9.6	14
49	Base Qualities Help Sequencingâ€‰Software. Genome Research, 1998, 8, 161-162.	5.5	8
50	Pathway-Based Factor Analysis of Gene Expression Data Produces Highly Heritable Phenotypes That Associate with Age. G3: Genes, Genomes, Genetics, 2015, 5, 839-847.	1.8	7