List of Publications by Year in descending order

50 papers	163,051 citations	61857 43 h-index	197535 49 g-index
53 all docs	53 docs citations	53 times ranked	170948 citing authors

#	Article	IF	Citations
1	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
2	Reference-based phasing using the Haplotype Reference Consortium panel. Nature Genetics, 2016, 48, 1443-1448.	9.4	1,357
3	A reference panel of 64,976 haplotypes for genotype imputation. Nature Genetics, 2016, 48, 1279-1283.	9.4	2,421
4	Health and population effects of rare gene knockouts in adult humans with related parents. Science, 2016, 352, 474-477.	6.0	272
5	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
6	Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. Nature Communications, 2015, 6, 8111.	5.8	300
7	Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. GigaScience, 2013, 2, 10.	3.3	582
8	Efficient de novo assembly of large genomes using compressed data structures. Genome Research, 2012, 22, 549-556.	2.4	649
9	Ensembl 2012. Nucleic Acids Research, 2012, 40, D84-D90.	6.5	840
10	Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nature Genetics, 2012, 44, 1084-1089.	9.4	701
11	Dindel: Accurate indel calls from short-read data. Genome Research, 2011, 21, 961-973.	2.4	383
12	Inference of human population history from individual whole-genome sequences. Nature, 2011, 475, 493-496.	13.7	2,053
13	Mouse genomic variation and its effect on phenotypes and gene regulation. Nature, 2011, 477, 289-294.	13.7	1,461
14	Ensembl 2011. Nucleic Acids Research, 2011, 39, D800-D806.	6.5	630
15	SNP detection and genotyping from low-coverage sequencing data on multiple diploid samples. Genome Research, 2011, 21, 952-960.	2.4	142
16	Ensembl's 10th year. Nucleic Acids Research, 2010, 38, D557-D562.	6.5	251
17	Copy number variant detection in inbred strains from short read sequence data. Bioinformatics, 2010, 26, 565-567.	1.8	47
18	Systematic Analysis of Human Protein Complexes Identifies Chromosome Segregation Proteins. Science, 2010, 328, 593-599.	6.0	465

#	Article	IF	Citations
19	Fast and accurate long-read alignment with Burrowsâ€“Wheeler transform. Bioinformatics, 2010, 26, 589-595.	1.8	10,002
20	EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. Genome Research, 2009, 19, 327-335.	2.4	1,058
21	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
22	Prepublication data sharing. Nature, 2009, 461, 168-170.	13.7	243
23	The Sequence Alignment/Map format and SAMtools. Bioinformatics, 2009, 25, 2078-2079.	1.8	49,124
24	Fast and accurate short read alignment with Burrowsâ€“Wheeler transform. Bioinformatics, 2009, 25, 1754-1760.	1.8	43,062
25	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
26	Accurate whole human genome sequencing using reversible terminator chemistry. Nature, 2008, 456, 53-59.	13.7	3,118
27	A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nature Biotechnology, 2008, 26, 779-785.	9.4	619
28	Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. Nature Genetics, 2008, 40, 722-729.	9.4	736
29	Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Research, 2008, 18, 1851-1858.	2.4	2,275
30	Vertebrate gene finding from multiple-species alignments using a two-level strategy. Genome Biology, 2006, 7, S6.	13.9	14
31	The DNA sequence of the human X chromosome. Nature, 2005, 434, 325-337.	13.7	985
32	The Sequence Ontology: a tool for the unification of genome annotations. Genome Biology, 2005, 6, R44.	13.9	638
33	WormBase: a comprehensive data resource for Caenorhabditis biology and genomics. Nucleic Acids Research, 2004, 33, D383-D389.	6.5	155
34	An Overview of Ensembl. Genome Research, 2004, 14, 925-928.	2.4	391
35	GeneWise and Genomewise. Genome Research, 2004, 14, 988-995.	2.4	2,128
36	WormBase: a multi-species resource for nematode biology and genomics. Nucleic Acids Research, 2004, 32, 411D-417.	6.5	610

#	Article	IF	Citations
37	Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature, 2004, 428, 493-521.	13.7	1,943
38	The Pfam protein families database. Nucleic Acids Research, 2004, 32, 138D-141.	6.5	3,084
39	Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature, 2003, 421, 231-237.	13.7	3,343
40	The InterPro Database, 2003 brings increased coverage and new features. Nucleic Acids Research, 2003, 31, 315-318.	6.5	640
41	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
42	The Genome Sequence of Caenorhabditis briggsae: A Platform for Comparative Genomics. PLoS Biology, 2003, 1, e45.	2.6	812
43	Comparative ab initio prediction of gene structures using pair HMMs. Bioinformatics, 2002, 18, 1309-1318.	1.8	114
44	GAZE: A Generic Framework for the Integration of Gene-Prediction Data by Dynamic Programming. Genome Research, 2002, 12, 1418-1427.	2.4	82
45	The Pfam Protein Families Database. Nucleic Acids Research, 2002, 30, 276-280.	6.5	2,067
46	Initial sequencing and analysis of the human genome. Nature, 2001, 409, 860-921.	13.7	21,074
47	Using GeneWise in the Drosophila Annotation Experiment. Genome Research, 2000, 10, 547-548.	2.4	338
48	Dynamic Programming Alignment Accuracy. Journal of Computational Biology, 1998, 5, 493-504.	0.8	97
49	Base Qualities Help Sequencingâ€‰Software. Genome Research, 1998, 8, 161-162.	2.4	8
50	Pfam: A comprehensive database of protein domain families based on seed alignments. , 1997, 28, 405-420.		1,036