

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/11122735/publications.pdf>

Version: 2024-02-01

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	The Sequence Alignment/Map format and SAMtools. <i>Bioinformatics</i> , 2009, 25, 2078-2079.	1.8	49,124
2	Fast and accurate short read alignment with Burrows-Wheeler transform. <i>Bioinformatics</i> , 2009, 25, 1754-1760.	1.8	43,062
3	Initial sequencing and analysis of the human genome. <i>Nature</i> , 2001, 409, 860-921.	13.7	21,074
4	Fast and accurate long-read alignment with Burrows-Wheeler transform. <i>Bioinformatics</i> , 2010, 26, 589-595.	1.8	10,002
5	Systematic functional analysis of the <i>Caenorhabditis elegans</i> genome using RNAi. <i>Nature</i> , 2003, 421, 231-237.	13.7	3,343
6	Accurate whole human genome sequencing using reversible terminator chemistry. <i>Nature</i> , 2008, 456, 53-59.	13.7	3,118
7	The Pfam protein families database. <i>Nucleic Acids Research</i> , 2004, 32, 138D-141.	6.5	3,084
8	A reference panel of 64,976 haplotypes for genotype imputation. <i>Nature Genetics</i> , 2016, 48, 1279-1283.	9.4	2,421
9	Mapping short DNA sequencing reads and calling variants using mapping quality scores. <i>Genome Research</i> , 2008, 18, 1851-1858.	2.4	2,275
10	GeneWise and Genomewise. <i>Genome Research</i> , 2004, 14, 988-995.	2.4	2,128
11	The Pfam Protein Families Database. <i>Nucleic Acids Research</i> , 2002, 30, 276-280.	6.5	2,067
12	Inference of human population history from individual whole-genome sequences. <i>Nature</i> , 2011, 475, 493-496.	13.7	2,053
13	Genome sequence of the Brown Norway rat yields insights into mammalian evolution. <i>Nature</i> , 2004, 428, 493-521.	13.7	1,943
14	Mouse genomic variation and its effect on phenotypes and gene regulation. <i>Nature</i> , 2011, 477, 289-294.	13.7	1,461
15	Reference-based phasing using the Haplotype Reference Consortium panel. <i>Nature Genetics</i> , 2016, 48, 1443-1448.	9.4	1,357
16	EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. <i>Genome Research</i> , 2009, 19, 327-335.	2.4	1,058
17	Pfam: A comprehensive database of protein domain families based on seed alignments. , 1997, 28, 405-420.		1,036
18	The DNA sequence of the human X chromosome. <i>Nature</i> , 2005, 434, 325-337.	13.7	985

#	ARTICLE	IF	CITATIONS
19	Ensembl 2012. <i>Nucleic Acids Research</i> , 2012, 40, D84-D90.	6.5	840
20	The Genome Sequence of <i>Caenorhabditis briggsae</i> : A Platform for Comparative Genomics. <i>PLoS Biology</i> , 2003, 1, e45.	2.6	812
21	Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. <i>Nature Genetics</i> , 2008, 40, 722-729.	9.4	736
22	Mapping cis- and trans-regulatory effects across multiple tissues in twins. <i>Nature Genetics</i> , 2012, 44, 1084-1089.	9.4	701
23	Efficient de novo assembly of large genomes using compressed data structures. <i>Genome Research</i> , 2012, 22, 549-556.	2.4	649
24	The InterPro Database, 2003 brings increased coverage and new features. <i>Nucleic Acids Research</i> , 2003, 31, 315-318.	6.5	640
25	The Sequence Ontology: a tool for the unification of genome annotations. <i>Genome Biology</i> , 2005, 6, R44.	13.9	638
26	Ensembl 2011. <i>Nucleic Acids Research</i> , 2011, 39, D800-D806.	6.5	630
27	A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. <i>Nature Biotechnology</i> , 2008, 26, 779-785.	9.4	619
28	WormBase: a multi-species resource for nematode biology and genomics. <i>Nucleic Acids Research</i> , 2004, 32, 411D-417.	6.5	610
29	Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. <i>GigaScience</i> , 2013, 2, 10.	3.3	582
30	The consensus coding sequence (CCDS) project: Identifying a common protein-coding gene set for the human and mouse genomes. <i>Genome Research</i> , 2009, 19, 1316-1323.	2.4	476
31	Systematic Analysis of Human Protein Complexes Identifies Chromosome Segregation Proteins. <i>Science</i> , 2010, 328, 593-599.	6.0	465
32	An Overview of Ensembl. <i>Genome Research</i> , 2004, 14, 925-928.	2.4	391
33	Dindel: Accurate indel calls from short-read data. <i>Genome Research</i> , 2011, 21, 961-973.	2.4	383
34	Using GeneWise in the <i>Drosophila</i> Annotation Experiment. <i>Genome Research</i> , 2000, 10, 547-548.	2.4	338
35	Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. <i>Nature Communications</i> , 2015, 6, 8111.	5.8	300
36	Health and population effects of rare gene knockouts in adult humans with related parents. <i>Science</i> , 2016, 352, 474-477.	6.0	272

#	ARTICLE	IF	CITATIONS
37	Ensembl's 10th year. <i>Nucleic Acids Research</i> , 2010, 38, D557-D562.	6.5	251
38	Prepublication data sharing. <i>Nature</i> , 2009, 461, 168-170.	13.7	243
39	WormBase: a comprehensive data resource for <i>Caenorhabditis</i> biology and genomics. <i>Nucleic Acids Research</i> , 2004, 33, D383-D389.	6.5	155
40	SNP detection and genotyping from low-coverage sequencing data on multiple diploid samples. <i>Genome Research</i> , 2011, 21, 952-960.	2.4	142
41	Comparative ab initio prediction of gene structures using pair HMMs. <i>Bioinformatics</i> , 2002, 18, 1309-1318.	1.8	114
42	Dynamic Programming Alignment Accuracy. <i>Journal of Computational Biology</i> , 1998, 5, 493-504.	0.8	97
43	GAZE: A Generic Framework for the Integration of Gene-Prediction Data by Dynamic Programming. <i>Genome Research</i> , 2002, 12, 1418-1427.	2.4	82
44	Whole-genome view of the consequences of a population bottleneck using 2926 genome sequences from Finland and United Kingdom. <i>European Journal of Human Genetics</i> , 2017, 25, 477-484.	1.4	60
45	Enhanced protein domain discovery by using language modeling techniques from speech recognition. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2003, 100, 4516-4520.	3.3	47
46	Copy number variant detection in inbred strains from short read sequence data. <i>Bioinformatics</i> , 2010, 26, 565-567.	1.8	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. <i>Genome Biology</i> , 2009, 10, R112.	13.9	36
48	Vertebrate gene finding from multiple-species alignments using a two-level strategy. <i>Genome Biology</i> , 2006, 7, S6.	13.9	14
49	Base Qualities Help Sequencing Software. <i>Genome Research</i> , 1998, 8, 161-162.	2.4	8
50	Pathway-Based Factor Analysis of Gene Expression Data Produces Highly Heritable Phenotypes That Associate with Age. <i>G3: Genes, Genomes, Genetics</i> , 2015, 5, 839-847.	0.8	7