Mark E Diekhans

List of Publications by Year
in descending order

Source: https://exaly.com/author-pdf/2600411/publications.pdf

Version: 2024-02-01

64 papers	29,083 citations	41344 49 h-index	114465 63 g-index
74 all docs	74 docs citations	74 times ranked	46157 citing authors

#	Article	IF	Citations
1	Initial sequencing and comparative analysis of the mouse genome. Nature, 2002, 420, 520-562.	27.8	6,319
2	GENCODE: The reference human genome annotation for The ENCODE Project. Genome Research, 2012, 22, 1760-1774.	5.5	4,217
3	GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Research, 2019, 47, D766-D773.	14.5	2,350
4	Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics, 2008, 24, 637-644.	4.1	1,618
5	The complete sequence of a human genome. Science, 2022, 376, 44-53.	12.6	1,222
6	Towards complete and error-free genome assemblies of all vertebrate species. Nature, 2021, 592, 737-746.	27.8	1,139
7	The UCSC Genome Browser database: 2015 update. Nucleic Acids Research, 2015, 43, D670-D681.	14.5	891
8	The UCSC Genome Browser database: 2019 update. Nucleic Acids Research, 2019, 47, D853-D858.	14.5	699
9	GENCODE 2021. Nucleic Acids Research, 2021, 49, D916-D923.	14.5	633
10	The UCSC Genome Browser database: 2014 update. Nucleic Acids Research, 2014, 42, D764-D770.	14.5	619
11	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
12	The UCSC Genome Browser database: 2018 update. Nucleic Acids Research, 2018, 46, D762-D769.	14.5	476
13	Multiple evidence strands suggest that there may be as few as 19 000 human protein-coding genes. Human Molecular Genetics, 2014, 23, 5866-5878.	2.9	463
14	Assemblathon 1: A competitive assessment of de novo short read assembly methods. Genome Research, 2011, 21, 2224-2241.	5.5	443
15	The UCSC Genome Browser database: 2016 update. Nucleic Acids Research, 2016, 44, D717-D725.	14.5	376
16	A Discriminative Framework for Detecting Remote Protein Homologies. Journal of Computational Biology, 2000, 7, 95-114.	1.6	370
17	Long-read sequence assembly of the gorilla genome. Science, 2016, 352, aae0344.	12.6	368
18	The UCSC Genome Browser database: 2021 update. Nucleic Acids Research, 2021, 49, D1046-D1057.	14.5	354

#	Article	IF	Citations
19	OUP accepted manuscript. Nucleic Acids Research, 2017, 45, D626-D634.	14.5	308
20	High-resolution comparative analysis of great ape genomes. Science, 2018, 360, .	12.6	304
21	The UCSC Cancer Genomics Browser: update 2015. Nucleic Acids Research, 2015, 43, D812-D817.	14.5	300
22	The UCSC Genome Browser database: extensions and updates 2011. Nucleic Acids Research, 2012, 40, D918-D923.	14.5	294
23	The GENCODE pseudogene resource. Genome Biology, 2012, 13, R51.	9.6	273
24	Covariation in Frequencies of Substitution, Deletion, Transposition, and Recombination During Eutherian Evolution. Genome Research, 2003, 13, 13-26.	5.5	263
25	Progressive Cactus is a multiple-genome aligner for the thousand-genome era. Nature, 2020, 587, 246-251.	27.8	256
26	Dense sampling of bird diversity increases power of comparative genomics. Nature, 2020, 587, 252-257.	27.8	251
27	Cactus: Algorithms for genome multiple sequence alignment. Genome Research, 2011, 21, 1512-1528.	5.5	245
28	28-Way vertebrate alignment and conservation track in the UCSC Genome Browser. Genome Research, 2007, 17, 1797-1808.	5.5	237
29	ENCODE whole-genome data in the UCSC Genome Browser: update 2012. Nucleic Acids Research, 2012, 40, D912-D917.	14.5	220
30	Complete genomic and epigenetic maps of human centromeres. Science, 2022, 376, eabl4178.	12.6	204
31	The UCSC Genome Browser database: 2022 update. Nucleic Acids Research, 2022, 50, D1115-D1122.	14.5	175
32	The UCSC Cancer Genomics Browser: update 2013. Nucleic Acids Research, 2013, 41, D949-D954.	14.5	172
33	Sixteen diverse laboratory mouse reference genomes define strain-specific haplotypes and novel functional loci. Nature Genetics, 2018, 50, 1574-1583.	21.4	169
34	Genomic legacy of the African cheetah, Acinonyx jubatus. Genome Biology, 2015, 16, 277.	8.8	167
35	Current status and new features of the Consensus Coding Sequence database. Nucleic Acids Research, 2014, 42, D865-D872.	14.5	140
36	UCSC Genome Browser enters 20th year. Nucleic Acids Research, 2020, 48, D756-D761.	14.5	138

#	Article	IF	Citations
37	Segmental duplications and their variation in a complete human genome. Science, 2022, 376, eabj6965.	12.6	130
38	The completion of the Mammalian Gene Collection (MGC). Genome Research, 2009, 19, 2324-2333.	5.5	125
39	Perspectives on ENCODE. Nature, 2020, 583, 693-698.	27.8	123
40	Sequence diversity analyses of an improved rhesus macaque genome enhance its biomedical utility. Science, 2020, 370, .	12.6	105
41	Comparative Genomics Search for Losses of Long-Established Genes on the Human Lineage. PLoS Computational Biology, 2007, 3, e247.	3.2	103
42	Consensus coding sequence (CCDS) database: a standardized set of human and mouse protein-coding regions supported by expert curation. Nucleic Acids Research, 2018, 46, D221-D228.	14.5	97
43	Retrocopy contributions to the evolution of the human genome. BMC Genomics, 2008, 9, 466.	2.8	93
44	Cactus Graphs for Genome Comparisons. Journal of Computational Biology, 2011, 18, 469-481.	1.6	93
45	Comparative Annotation Toolkit (CAT)â€”simultaneous clade and personal genome annotation. Genome Research, 2018, 28, 1029-1038.	5.5	86
46	Predicting protein structure using only sequence information. Proteins: Structure, Function and Bioinformatics, 1999, 37, 121-125.	2.6	76
47	Comparative analysis of pseudogenes across three phyla. Proceedings of the National Academy of Sciences of the United States of America, 2014, 111, 13361-13366.	7.1	72
48	MuPIT interactive: webserver for mapping variant positions to annotated, interactive 3D structures. Human Genetics, 2013, 132, 1235-1243.	3.8	68
49	Whole-Genome Alignment and Comparative Annotation. Annual Review of Animal Biosciences, 2019, 7, 41-64.	7.4	62
50	AMELIE speeds Mendelian diagnosis by matching patient phenotype and genotype to primary literature. Science Translational Medicine, 2020, 12, .	12.4	60
51	Predicting protein structure using only sequence information. Proteins: Structure, Function and Bioinformatics, 1999, 37, 121-125.	2.6	52
52	Regions of extreme synonymous codon selection in mammalian genes. Nucleic Acids Research, 2006, 34, 1700-1710.	14.5	47
53	Targeted discovery of novel human exons by comparative genomics. Genome Research, 2007, 17, 1763-1773.	5.5	42
54	Evaluating recovery potential of the northern white rhinoceros from cryopreserved somatic cells. Genome Research, 2018, 28, 780-788.	5.5	39

#	Article	IF	Citations
55	A high-quality bonobo genome refines the analysis of hominid evolution. Nature, 2021, 594, 77-81.	27.8	39
56	The NIH BD2K center for big data in translational genomics. Journal of the American Medical Informatics Association: JAMIA, 2015, 22, 1143-1147.	4.4	30
57	Re-annotation of 191 developmental and epileptic encephalopathy-associated genes unmasks de novo variants in SCN1A. Npj Genomic Medicine, 2019, 4, 31.	3.8	27
58	Analysis of Human mRNAs With the Reference Genome Sequence Reveals Potential Errors, Polymorphisms, and RNA Editing. Genome Research, 2004, 14, 2034-2040.	5.5	26
59	Transcriptional activity and strain-specific history of mouse pseudogenes. Nature Communications, 2020, 11, 3695.	12.8	17
60	Matching phenotypes to whole genomes: Lessons learned from four iterations of the personal genome project community challenges. Human Mutation, 2017, 38, 1266-1276.	2.5	14
61	Identification of high-confidence human poly(A) RNA isoform scaffolds using nanopore sequencing. Rna, 2022, 28, 162-176.	3.5	12
62	halSynteny: a fast, easy-to-use conserved synteny block construction method for multiple whole-genome alignments. GigaScience, 2020, 9, .	6.4	10
63	Score Functions for Determining Regional Conservation in Two-Species Local Alignments. Journal of Computational Biology, 2004, 11, 395-411.	1.6	2
64	Predicting protein structure using only sequence information. , 1999, 37, 121.		1