

List of Publications by Citations

Source: <https://exaly.com/author-pdf/247226/kevin-y-yip-publications-by-citations.pdf>
Version: 2024-04-10

This document has been generated based on the publications and citations recorded by exaly.com. For the latest version of this publication list, visit the link given above.
The third column is the impact factor (IF) of the journal, and the fourth column is the number of citations of the article.

107 papers	18,339 citations	37 h-index	135 g-index
146 ext. papers	22,536 ext. citations	12 avg, IF	8.16 L-index

#	Paper	IF	Citations
107	An integrated encyclopedia of DNA elements in the human genome. <i>Nature</i> , 2012 , 489, 57-74	50.4	11449
106	Architecture of the human regulatory network derived from ENCODE data. <i>Nature</i> , 2012 , 489, 91-100	50.4	1104
105	A user's guide to the encyclopedia of DNA elements (ENCODE). <i>PLoS Biology</i> , 2011 , 9, e1001046	9.7	1060
104	Integrative analysis of the <i>Caenorhabditis elegans</i> genome by the modENCODE project. <i>Science</i> , 2010 , 330, 1775-87	33.3	744
103	The International Human Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery. <i>Cell</i> , 2016 , 167, 1145-1149	56.2	232
102	FunSeq2: a framework for prioritizing noncoding regulatory variants in cancer. <i>Genome Biology</i> , 2014 , 15, 480	18.3	209
101	Comparative analysis of the transcriptome across distant species. <i>Nature</i> , 2014 , 512, 445-8	50.4	207
100	Classification of human genomic regions based on experimentally determined binding sites of more than 100 transcription-related factors. <i>Genome Biology</i> , 2012 , 13, R48	18.3	194
99	Exome and genome sequencing of nasopharynx cancer identifies NF- κ B pathway activating mutations. <i>Nature Communications</i> , 2017 , 8, 14121	17.4	163
98	Extensive in vivo metabolite-protein interactions revealed by large-scale systematic analyses. <i>Cell</i> , 2010 , 143, 639-50	56.2	162
97	Understanding transcriptional regulation by integrative analysis of transcription factor binding data. <i>Genome Research</i> , 2012 , 22, 1658-67	9.7	133
96	Whole-genome bisulfite sequencing of multiple individuals reveals complementary roles of promoter and gene body methylation in transcriptional regulation. <i>Genome Biology</i> , 2014 , 15, 408	18.3	131
95	Reconstruction of enhancer-target networks in 935 samples of human primary cells, tissues and cell lines. <i>Nature Genetics</i> , 2017 , 49, 1428-1436	36.3	124
94	Decision Trees for Uncertain Data. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2011 , 23, 64-78	4.2	116
93	Genome-Wide Structural Variation Detection by Genome Mapping on Nanochannel Arrays. <i>Genetics</i> , 2016 , 202, 351-62	4	103
92	Efficient Clustering of Uncertain Data. <i>IEEE International Conference on Data Mining</i> , 2006 ,		102
91	A statistical framework for modeling gene expression using chromatin features and application to modENCODE datasets. <i>Genome Biology</i> , 2011 , 12, R15	18.3	96

90	HARP: a practical projected clustering algorithm. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2004 , 16, 1387-1397	4.2	96
89	Sustained antidiabetic effects of a berberine-containing Chinese herbal medicine through regulation of hepatic gene expression. <i>Diabetes</i> , 2012 , 61, 933-43	0.9	88
88	YeastHub: a semantic web use case for integrating data in the life sciences domain. <i>Bioinformatics</i> , 2005 , 21 Suppl 1, i85-96	7.2	74
87	Improved reconstruction of in silico gene regulatory networks by integrating knockout and perturbation data. <i>PLoS ONE</i> , 2010 , 5, e8121	3.7	73
86	Genome maps across 26 human populations reveal population-specific patterns of structural variation. <i>Nature Communications</i> , 2019 , 10, 1025	17.4	70
85	HCLS 2.0/3.0: health care and life sciences data mashup using Web 2.0/3.0. <i>Journal of Biomedical Informatics</i> , 2008 , 41, 694-705	10.2	67
84	Identification of a major determinant for serine-threonine kinase phosphoacceptor specificity. <i>Molecular Cell</i> , 2014 , 53, 140-7	17.6	64
83	Establishment and characterization of new tumor xenografts and cancer cell lines from EBV-positive nasopharyngeal carcinoma. <i>Nature Communications</i> , 2018 , 9, 4663	17.4	63
82	Mining periodic patterns with gap requirement from sequences. <i>ACM Transactions on Knowledge Discovery From Data</i> , 2007 , 1, 7	4	62
81	The tYNA platform for comparative interactomics: a web tool for managing, comparing and mining multiple networks. <i>Bioinformatics</i> , 2006 , 22, 2968-70	7.2	61
80	An integrated system for studying residue coevolution in proteins. <i>Bioinformatics</i> , 2008 , 24, 290-2	7.2	60
79	CLUSTAG: hierarchical clustering and graph methods for selecting tag SNPs. <i>Bioinformatics</i> , 2005 , 21, 1735-6	7.2	59
78	Machine learning and genome annotation: a match meant to be?. <i>Genome Biology</i> , 2013 , 14, 205	18.3	56
77	Improved prediction of RNA secondary structure by integrating the free energy model with restraints derived from experimental probing data. <i>Nucleic Acids Research</i> , 2015 , 43, 7247-59	20.1	55
76	Prediction and characterization of noncoding RNAs in C. elegans by integrating conservation, secondary structure, and high-throughput sequencing and array data. <i>Genome Research</i> , 2011 , 21, 276-85	9.7	55
75	Comparing classical pathways and modern networks: towards the development of an edge ontology. <i>Trends in Biochemical Sciences</i> , 2007 , 32, 320-31	10.3	49
74	Aberrant enhancer hypomethylation contributes to hepatic carcinogenesis through global transcriptional reprogramming. <i>Nature Communications</i> , 2019 , 10, 335	17.4	49
73	Decision Trees for Uncertain Data. <i>Proceedings - International Conference on Data Engineering</i> , 2009 ,	2	45

72	Progressive skylining over Web-accessible databases. <i>Data and Knowledge Engineering</i> , 2006 , 57, 122-147.	17.5	40
71	Identification of Genes Associated With Hirschsprung Disease, Based on Whole-Genome Sequence Analysis, and Potential Effects on Enteric Nervous System Development. <i>Gastroenterology</i> , 2018 , 155, 1908-1922.e5	13.3	38
70	Identification of a recurrent transforming UBR5-ZNF423 fusion gene in EBV-associated nasopharyngeal carcinoma. <i>Journal of Pathology</i> , 2013 , 231, 158-67	9.4	35
69	Yin Yang 1-mediated epigenetic silencing of tumour-suppressive microRNAs activates nuclear factor- κ B in hepatocellular carcinoma. <i>Journal of Pathology</i> , 2016 , 238, 651-64	9.4	35
68	Analysis of membrane proteins in metagenomics: networks of correlated environmental features and protein families. <i>Genome Research</i> , 2010 , 20, 960-71	9.7	33
67	Complete genomic sequence of Epstein-Barr virus in nasopharyngeal carcinoma cell line C666-1. <i>Infectious Agents and Cancer</i> , 2013 , 8, 29	3.5	32
66	ACT: aggregation and correlation toolbox for analyses of genome tracks. <i>Bioinformatics</i> , 2011 , 27, 1152-4.	4.2	31
65	EBV-encoded miRNAs target ATM-mediated response in nasopharyngeal carcinoma. <i>Journal of Pathology</i> , 2018 , 244, 394-407	9.4	30
64	An integrative ENCODE resource for cancer genomics. <i>Nature Communications</i> , 2020 , 11, 3696	17.4	28
63	A common set of distinct features that characterize noncoding RNAs across multiple species. <i>Nucleic Acids Research</i> , 2015 , 43, 104-14	20.1	27
62	On discovery of extremely low-dimensional clusters using semi-supervised projected clustering		26
61	Mining periodic patterns with gap requirement from sequences		26
60	Supervised enhancer prediction with epigenetic pattern recognition and targeted validation. <i>Nature Methods</i> , 2020 , 17, 807-814	21.6	24
59	Are special read alignment strategies necessary and cost-effective when handling sequencing reads from patient-derived tumor xenografts?. <i>BMC Genomics</i> , 2014 , 15, 1172	4.5	22
58	LinkHub: a Semantic Web system that facilitates cross-database queries and information retrieval in proteomics. <i>BMC Bioinformatics</i> , 2007 , 8 Suppl 3, S5	3.6	22
57	OMBlast: alignment tool for optical mapping using a seed-and-extend approach. <i>Bioinformatics</i> , 2017 , 33, 311-319	7.2	22
56	A comprehensive web tool for toehold switch design. <i>Bioinformatics</i> , 2018 , 34, 2862-2864	7.2	19
55	Training set expansion: an approach to improving the reconstruction of biological networks from limited and uneven reliable interactions. <i>Bioinformatics</i> , 2009 , 25, 243-50	7.2	18

54	Development of grid-like applications for public health using Web 2.0 mashup techniques. <i>Journal of the American Medical Informatics Association: JAMIA</i> , 2008 , 15, 783-6	8.6	18
53	Jellyfish genomes reveal distinct homeobox gene clusters and conservation of small RNA processing. <i>Nature Communications</i> , 2020 , 11, 3051	17.4	17
52	On mining micro-array data by Order-Preserving Submatrix. <i>International Journal of Bioinformatics Research and Applications</i> , 2007 , 3, 42-64	0.9	17
51	Loss of tumor suppressor IGFBP4 drives epigenetic reprogramming in hepatic carcinogenesis. <i>Nucleic Acids Research</i> , 2018 , 46, 8832-8847	20.1	16
50	Genome-wide analysis of chromatin features identifies histone modification sensitive and insensitive yeast transcription factors. <i>Genome Biology</i> , 2011 , 12, R111	18.3	16
49	Interactome-transcriptome analysis discovers signatures complementary to GWAS Loci of Type 2 Diabetes. <i>Scientific Reports</i> , 2016 , 6, 35228	4.9	16
48	OMSV enables accurate and comprehensive identification of large structural variations from nanochannel-based single-molecule optical maps. <i>Genome Biology</i> , 2017 , 18, 230	18.3	15
47	OMTools: a software package for visualizing and processing optical mapping data. <i>Bioinformatics</i> , 2017 , 33, 2933-2935	7.2	14
46	Identification of specificity determining residues in peptide recognition domains using an information theoretic approach applied to large-scale binding maps. <i>BMC Biology</i> , 2011 , 9, 53	7.3	14
45	G9a Plays Distinct Roles in Maintaining DNA Methylation, Retrotransposon Silencing, and Chromatin Looping. <i>Cell Reports</i> , 2020 , 33, 108315	10.6	14
44	A web services choreography scenario for interoperating bioinformatics applications. <i>BMC Bioinformatics</i> , 2004 , 5, 25	3.6	11
43	A selective HDAC8 inhibitor potentiates antitumor immunity and efficacy of immune checkpoint blockade in hepatocellular carcinoma. <i>Science Translational Medicine</i> , 2021 , 13,	17.5	11
42	A network approach to exploring the functional basis of gene-gene epistatic interactions in disease susceptibility. <i>Bioinformatics</i> , 2018 , 34, 1741-1749	7.2	10
41	The Essential Component in DNA-Based Information Storage System: Robust Error-Tolerating Module. <i>Frontiers in Bioengineering and Biotechnology</i> , 2014 , 2, 49	5.8	10
40	Whole-genome profiling of nasopharyngeal carcinoma reveals viral-host co-operation in inflammatory NF- κ B activation and immune escape. <i>Nature Communications</i> , 2021 , 12, 4193	17.4	10
39	Systematic exploration of autonomous modules in noisy microRNA-target networks for testing the generality of the ceRNA hypothesis. <i>BMC Genomics</i> , 2014 , 15, 1178	4.5	9
38	Identifying projected clusters from gene expression profiles. <i>Journal of Biomedical Informatics</i> , 2004 , 37, 345-57	10.2	8
37	New guidelines for DNA methylome studies regarding 5-hydroxymethylcytosine for understanding transcriptional regulation. <i>Genome Research</i> , 2019 , 29, 543-553	9.7	8

36	Horseshoe crab genomes reveal the evolution of genes and microRNAs after three rounds of whole genome duplication. <i>Communications Biology</i> , 2021 , 4, 83	6.7	8
35	A novel method for discovering local spatial clusters of genomic regions with functional relationships from DNA contact maps. <i>Bioinformatics</i> , 2016 , 32, i111-i120	7.2	7
34	Towards a More Accurate Error Model for BioNano Optical Maps. <i>Lecture Notes in Computer Science</i> , 2016 , 67-79	0.9	7
33	Multi-level learning: improving the prediction of protein, domain and residue interactions by allowing information flow between levels. <i>BMC Bioinformatics</i> , 2009 , 10, 241	3.6	7
32	Mining Order-Preserving Submatrices from Data with Repeated Measurements 2008 ,		7
31	Biosphere: the interoperation of web services in microarray cluster analysis. <i>Applied Bioinformatics</i> , 2004 , 3, 253-6		7
30	Analysis of sequencing data for probing RNA secondary structures and protein-RNA binding in studying posttranscriptional regulations. <i>Briefings in Bioinformatics</i> , 2016 , 17, 1032-1043	13.4	6
29	Metric and trigonometric pruning for clustering of uncertain data in 2D geometric space. <i>Information Systems</i> , 2011 , 36, 476-497	2.7	6
28	A semi-supervised approach to projected clustering with applications to microarray data. <i>International Journal of Data Mining and Bioinformatics</i> , 2009 , 3, 229-59	0.5	6
27	Comparative analysis of single-cell parallel sequencing approaches in oocyte application. <i>International Journal of Biochemistry and Cell Biology</i> , 2019 , 107, 1-5	5.6	6
26	Mining Order-Preserving Submatrices from Data with Repeated Measurements. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2013 , 25, 1587-1600	4.2	5
25	Computational identification of protein binding sites on RNAs using high-throughput RNA structure-probing data. <i>Bioinformatics</i> , 2014 , 30, 1049-1055	7.2	5
24	A cross-organism framework for supervised enhancer prediction with epigenetic pattern recognition and targeted validation		5
23	Case Report: Exome sequencing reveals recurrent mutations and a loss-of-function mutation in a rare undifferentiated tongue sarcoma. <i>F1000Research</i> , 2018 , 7, 499	3.6	4
22	Whole-genome analysis of noncoding genetic variations identifies multiscale regulatory element perturbations associated with Hirschsprung disease. <i>Genome Research</i> , 2020 , 30, 1618-1632	9.7	4
21	Integrating Information in Biological Ontologies and Molecular Networks to Infer Novel Terms. <i>Scientific Reports</i> , 2016 , 6, 39237	4.9	3
20	OMMA enables population-scale analysis of complex genomic features and phylogenomic relationships from nanochannel-based optical maps. <i>GigaScience</i> , 2019 , 8,	7.6	3
19	Shaping the nebulous enhancer in the era of high-throughput assays and genome editing. <i>Briefings in Bioinformatics</i> , 2020 , 21, 836-850	13.4	3

18	Identification and characterization of a novel Epstein-Barr Virus-encoded circular RNA from LMP-2 Gene. <i>Scientific Reports</i> , 2021 , 11, 14392	4.9	3
17	START: a system for flexible analysis of hundreds of genomic signal tracks in few lines of SQL-like queries. <i>BMC Genomics</i> , 2017 , 18, 749	4.5	2
16	Input Validation for Semi-supervised Clustering 2006 ,		2
15	Sirtuin 7 super-enhancer drives epigenomic reprogramming in hepatocarcinogenesis. <i>Cancer Letters</i> , 2022 , 525, 115-130	9.9	2
14	An integrative ENCODE resource for cancer genomics		2
13	VAS: a convenient web portal for efficient integration of genomic features with millions of genetic variants. <i>BMC Genomics</i> , 2014 , 15, 886	4.5	1
12	On Mining Micro-array data by Order-Preserving Submatrix 2005 ,		1
11	Identifying projected clusters from gene expression profiles		1
10	A unified framework for integrative study of heterogeneous gene regulatory mechanisms. <i>Nature Machine Intelligence</i> , 2020 , 2, 447-456	22.5	1
9	A proof-of-concept study for the pathogenetic role of enhancer hypomethylation of MYBPHL in multiple myeloma. <i>Scientific Reports</i> , 2021 , 11, 7009	4.9	1
8	Deep learning identifies and quantifies recombination hotspot determinants		1
7	A Survey of the Computational Methods for Enhancers and Enhancer-target Predictions 2016 , 3-27		1
6	Quantifying full-length circular RNAs in cancer		1
5	Reusability report: Capturing properties of biological objects and their relationships using graph neural networks. <i>Nature Machine Intelligence</i> , 2022 , 4, 222-226	22.5	1
4	A general near-exact k-mer counting method with low memory consumption enables de novo assembly of 106 human sequence data in 2.7 hours. <i>Bioinformatics</i> , 2020 , 36, i625-i633	7.2	0
3	Flexible k-mers with variable-length indels for identifying binding sequences of protein dimers. <i>Briefings in Bioinformatics</i> , 2020 , 21, 1787-1797	13.4	0
2	Reply to Tnflated performance measures in enhancer-promoter interaction-prediction methodsT <i>Nature Genetics</i> , 2019 , 51, 1201-1202	36.3	
1	ECplot: an online tool for making standardized plots from large datasets for bioinformatics publications. <i>Bioinformatics</i> , 2014 , 30, 1467-8	7.2	

