

Pengyi Yang

List of Publications by Year in Descending Order

Source: <https://exaly.com/author-pdf/7706880/pengyi-yang-publications-by-year.pdf>

Version: 2024-04-25

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.

87
papers

2,399
citations

23
h-index

48
g-index

114
ext. papers

3,361
ext. citations

9.1
avg. IF

5.01
L-index

#	Paper	IF	Citations
87	Benchmarking clustering algorithms on estimating the number of cell types from single-cell RNA-sequencing data.. <i>Genome Biology</i> , 2022 , 23, 49	18.3	3
86	Functional analysis of the stable phosphoproteome reveals cancer vulnerabilities.. <i>Bioinformatics</i> , 2022 ,	7.2	1
85	scREMOTE: Using multimodal single cell data to predict regulatory gene relationships and to build a computational cell reprogramming model.. <i>NAR Genomics and Bioinformatics</i> , 2022 , 4, lqac023	3.7	0
84	Mouse organogenesis atlas at single-cell resolution.. <i>Cell</i> , 2022 , 185, 1625-1627	56.2	
83	Computational systems approach towards phosphoproteomics and their downstream regulation.. <i>Proteomics</i> , 2022 , e2200068	4.8	1
82	A benchmark study of simulation methods for single-cell RNA sequencing data. <i>Nature Communications</i> , 2021 , 12, 6911	17.4	3
81	Feature selection revisited in the single-cell era. <i>Genome Biology</i> , 2021 , 22, 321	18.3	4
80	Decoding the function of bivalent chromatin in development and cancer. <i>Genome Research</i> , 2021 ,	9.7	3
79	Global phosphoproteomics reveals DYRK1A regulates CDK1 activity in glioblastoma cells. <i>Cell Death Discovery</i> , 2021 , 7, 81	6.9	8
78	Defining cell identity beyond the premise of differential gene expression. <i>Cell Regeneration</i> , 2021 , 10, 20	2.5	0
77	Protocol for the processing and downstream analysis of phosphoproteomic data with PhosR. <i>STAR Protocols</i> , 2021 , 2, 100585	1.4	1
76	Akt phosphorylates insulin receptor substrate to limit PI3K-mediated PIP3 synthesis. <i>ELife</i> , 2021 , 10,	8.9	2
75	Enhanced structure and function of human pluripotent stem cell-derived beta-cells cultured on extracellular matrix. <i>Stem Cells Translational Medicine</i> , 2021 , 10, 492-505	6.9	4
74	PhosR enables processing and functional analysis of phosphoproteomic data. <i>Cell Reports</i> , 2021 , 34, 108776	10.6	7
73	Clonal evolution in liver cancer at single-cell and single-variant resolution. <i>Journal of Hematology and Oncology</i> , 2021 , 14, 22	22.4	8
72	Signaling Heterogeneity is Defined by Pathway Architecture and Intercellular Variability in Protein Expression. <i>IScience</i> , 2021 , 24, 102118	6.1	8
71	A hierarchical approach to removal of unwanted variation for large-scale metabolomics data. <i>Nature Communications</i> , 2021 , 12, 4992	17.4	2

70	Uncovering cell identity through differential stability with Cepo. <i>Nature Computational Science</i> , 2021 , 1, 784-790		0
69	CiteFuse enables multi-modal analysis of CITE-seq data. <i>Bioinformatics</i> , 2020 , 36, 4137-4143	7.2	19
68	scClassify: sample size estimation and multiscale classification of cells using single and multiple reference. <i>Molecular Systems Biology</i> , 2020 , 16, e9389	12.2	22
67	Transcriptional network dynamics during the progression of pluripotency revealed by integrative statistical learning. <i>Nucleic Acids Research</i> , 2020 , 48, 1828-1842	20.1	6
66	Temporal ordering of omics and multiomic events inferred from time-series data. <i>Npj Systems Biology and Applications</i> , 2020 , 6, 22	5	5
65	Ensemble deep learning in bioinformatics. <i>Nature Machine Intelligence</i> , 2020 , 2, 500-508	22.5	32
64	Data Independent Acquisition Proteomic Analysis Can Discriminate between Actinic Keratosis, Bowen's Disease, and Cutaneous Squamous Cell Carcinoma. <i>Journal of Investigative Dermatology</i> , 2020 , 140, 212-222.e11	4.3	13
63	Integrated Glycoproteomics Identifies a Role of N-Glycosylation and Galectin-1 on Myogenesis and Muscle Development. <i>Molecular and Cellular Proteomics</i> , 2020 , 20, 100030	7.6	13
62	Evaluating stably expressed genes in single cells. <i>GigaScience</i> , 2019 , 8,	7.6	23
61	Multi-omic Profiling Reveals Dynamics of the Phased Progression of Pluripotency. <i>Cell Systems</i> , 2019 , 8, 427-445.e10	10.6	50
60	QCMAP: An Interactive Web-Tool for Performance Diagnosis and Prediction of LC-MS Systems. <i>Proteomics</i> , 2019 , 19, e1900068	4.8	3
59	scMerge leverages factor analysis, stable expression, and pseudoreplication to merge multiple single-cell RNA-seq datasets. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2019 , 116, 9775-9784	11.5	67
58	MiR-93-5p is a novel predictor of coronary in-stent restenosis. <i>Heart Asia</i> , 2019 , 11, e011134	1.9	12
57	Impact of similarity metrics on single-cell RNA-seq data clustering. <i>Briefings in Bioinformatics</i> , 2019 , 20, 2316-2326	13.4	47
56	An Uncertainty Visual Analytics Framework for fMRI Functional Connectivity. <i>Neuroinformatics</i> , 2019 , 17, 211-223	3.2	2
55	AdaSampling for Positive-Unlabeled and Label Noise Learning With Bioinformatics Applications. <i>IEEE Transactions on Cybernetics</i> , 2019 , 49, 1932-1943	10.2	20
54	NF-Y controls fidelity of transcription initiation at gene promoters through maintenance of the nucleosome-depleted region. <i>Nature Communications</i> , 2019 , 10, 3072	17.4	25
53	An integrative systems genetic analysis of mammalian lipid metabolism. <i>Nature</i> , 2019 , 567, 187-193	50.4	63

52	Global redox proteome and phosphoproteome analysis reveals redox switch in Akt. <i>Nature Communications</i> , 2019 , 10, 5486	17.4	36
51	Autoencoder-based cluster ensembles for single-cell RNA-seq data analysis. <i>BMC Bioinformatics</i> , 2019 , 20, 660	3.6	18
50	scDC: single cell differential composition analysis. <i>BMC Bioinformatics</i> , 2019 , 20, 721	3.6	12
49	scReClassify: post hoc cell type classification of single-cell rNA-seq data. <i>BMC Genomics</i> , 2019 , 20, 913	4.5	9
48	Mitochondrial CoQ deficiency is a common driver of mitochondrial oxidants and insulin resistance. <i>ELife</i> , 2018 , 7,	8.9	61
47	Intragenic Enhancers Attenuate Host Gene Expression. <i>Molecular Cell</i> , 2017 , 68, 104-117.e6	17.6	54
46	Integrative analysis identifies co-dependent gene expression regulation of BRG1 and CHD7 at distal regulatory sites in embryonic stem cells. <i>Bioinformatics</i> , 2017 , 33, 1916-1920	7.2	9
45	An improved Akt reporter reveals intra- and inter-cellular heterogeneity and oscillations in signal transduction. <i>Journal of Cell Science</i> , 2017 , 130, 2757-2766	5.3	12
44	Positive unlabeled learning via wrapper-based adaptive sampling 2017 ,		10
43	Positive-unlabeled ensemble learning for kinase substrate prediction from dynamic phosphoproteomics data. <i>Bioinformatics</i> , 2016 , 32, 252-9	7.2	17
42	CNOT3-Dependent mRNA Deadenylation Safeguards the Pluripotent State. <i>Stem Cell Reports</i> , 2016 , 7, 897-910	8	18
41	Unraveling Kinase Activation Dynamics Using Kinase-Substrate Relationships from Temporal Large-Scale Phosphoproteomics Studies. <i>PLoS ONE</i> , 2016 , 11, e0157763	3.7	9
40	KinasePA: Phosphoproteomics data annotation using hypothesis driven kinase perturbation analysis. <i>Proteomics</i> , 2016 , 16, 1868-71	4.8	14
39	NoisyGOA: Noisy GO annotations prediction using taxonomic and semantic similarity. <i>Computational Biology and Chemistry</i> , 2016 , 65, 203-211	3.6	4
38	mTORC1 Is a Major Regulatory Node in the FGF21 Signaling Network in Adipocytes. <i>Cell Reports</i> , 2016 , 17, 29-36	10.6	59
37	DNMT1 is essential for mammary and cancer stem cell maintenance and tumorigenesis. <i>Nature Communications</i> , 2015 , 6, 6910	17.4	148
36	Global Phosphoproteomic Analysis of Human Skeletal Muscle Reveals a Network of Exercise-Regulated Kinases and AMPK Substrates. <i>Cell Metabolism</i> , 2015 , 22, 922-35	24.6	233
35	Knowledge-Based Analysis for Detecting Key Signaling Events from Time-Series Phosphoproteomics Data. <i>PLoS Computational Biology</i> , 2015 , 11, e1004403	5	20

34	Direction pathway analysis of large-scale proteomics data reveals novel features of the insulin action pathway. <i>Bioinformatics</i> , 2014 , 30, 808-14	7.2	20
33	Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. <i>EMBO Journal</i> , 2014 , 33, 878-89	13	109
32	Selective photodegradation of 1-methylimidazole-2-thiol by the magnetic and dual conductive imprinted photocatalysts based on TiO ₂ /Fe ₃ O ₄ /MWCNTs. <i>Chemical Engineering Journal</i> , 2014 , 240, 244-252	14.7	52
31	ISL1 regulates peroxisome proliferator-activated receptor β activation and early adipogenesis via bone morphogenetic protein 4-dependent and -independent mechanisms. <i>Molecular and Cellular Biology</i> , 2014 , 34, 3607-17	4.8	8
30	Histone-fold domain protein NF-Y promotes chromatin accessibility for cell type-specific master transcription factors. <i>Molecular Cell</i> , 2014 , 55, 708-22	17.6	108
29	Sample Subset Optimization Techniques for Imbalanced and Ensemble Learning Problems in Bioinformatics Applications. <i>IEEE Transactions on Cybernetics</i> , 2014 , 44, 445-55	10.2	51
28	Ensemble-Based Wrapper Methods for Feature Selection and Class Imbalance Learning. <i>Lecture Notes in Computer Science</i> , 2013 , 544-555	0.9	32
27	Dynamic adipocyte phosphoproteome reveals that Akt directly regulates mTORC2. <i>Cell Metabolism</i> , 2013 , 17, 1009-1020	24.6	269
26	Stability of Feature Selection Algorithms and Ensemble Feature Selection Methods in Bioinformatics 2013 , 333-352		15
25	Improving X!Tandem on peptide identification from mass spectrometry by self-boosted Percolator. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2012 , 9, 1273-80	3	14
24	Re-fraction: a machine learning approach for deterministic identification of protein homologues and splice variants in large-scale MS-based proteomics. <i>Journal of Proteome Research</i> , 2012 , 11, 3035-45	5.6	6
23	OCAP: an open comprehensive analysis pipeline for iTRAQ. <i>Bioinformatics</i> , 2012 , 28, 1404-5	7.2	8
22	Gene-gene interaction filtering with ensemble of filters. <i>BMC Bioinformatics</i> , 2011 , 12 Suppl 1, S10	3.6	32
21	Sample Subset Optimization for Classifying Imbalanced Biological Data. <i>Lecture Notes in Computer Science</i> , 2011 , 333-344	0.9	10
20	Genetic Algorithm-Based Multi-objective Optimisation for QoS-Aware Web Services Composition. <i>Lecture Notes in Computer Science</i> , 2010 , 549-554	0.9	2
19	A dynamic wavelet-based algorithm for pre-processing tandem mass spectrometry data. <i>Bioinformatics</i> , 2010 , 26, 2242-9	7.2	11
18	A clustering based hybrid system for biomarker selection and sample classification of mass spectrometry data. <i>Neurocomputing</i> , 2010 , 73, 2317-2331	5.4	10
17	A genetic ensemble approach for gene-gene interaction identification. <i>BMC Bioinformatics</i> , 2010 , 11, 524	3.6	17

16	A multi-filter enhanced genetic ensemble system for gene selection and sample classification of microarray data. <i>BMC Bioinformatics</i> , 2010 , 11 Suppl 1, S5	3.6	33
15	Hierarchical kernel mixture models for the prediction of AIDS disease progression using HIV structural gp120 profiles. <i>BMC Genomics</i> , 2010 , 11 Suppl 4, S22	4.5	5
14	A Review of Ensemble Methods in Bioinformatics. <i>Current Bioinformatics</i> , 2010 , 5, 296-308	4.7	284
13	An Agent-Based Hybrid System for Microarray Data Analysis. <i>IEEE Intelligent Systems</i> , 2009 , 24, 53-63	4.2	6
12	A particle swarm based hybrid system for imbalanced medical data sampling. <i>BMC Genomics</i> , 2009 , 10 Suppl 3, S34	4.5	33
11	Multiagent Framework for Bio-data Mining. <i>Lecture Notes in Computer Science</i> , 2009 , 200-207	0.9	2
10	Modeling and simulation of Parallel Current Mode controlled boost converter 2008 ,		1
9	A Clustering Based Hybrid System for Mass Spectrometry Data Analysis. <i>Lecture Notes in Computer Science</i> , 2008 , 98-109	0.9	3
8	Hybrid Methods to Select Informative Gene Sets in Microarray Data Classification 2007 , 810-814		4
7	hRUV: Hierarchical approach to removal of unwanted variation for large-scale metabolomics data		3
6	Evaluating stably expressed genes in single cells		8
5	scMerge: Integration of multiple single-cell transcriptomics datasets leveraging stable expression and pseudo-replication		8
4	scClassify: hierarchical classification of cells		8
3	A benchmark study of simulation methods for single-cell RNA sequencing data		1
2	Spatial analysis for highly multiplexed imaging data to identify tissue microenvironments		2
1	Cepo uncovers cell identity through differential stability		1