

# Pengyi Yang

## List of Publications by Citations

**Source:** <https://exaly.com/author-pdf/7706880/pengyi-yang-publications-by-citations.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	A Review of Ensemble Methods in Bioinformatics. <i>Current Bioinformatics</i> , <b>2010</b> , 5, 296-308	4.7	284
86	Dynamic adipocyte phosphoproteome reveals that Akt directly regulates mTORC2. <i>Cell Metabolism</i> , <b>2013</b> , 17, 1009-1020	24.6	269
85	Global Phosphoproteomic Analysis of Human Skeletal Muscle Reveals a Network of Exercise-Regulated Kinases and AMPK Substrates. <i>Cell Metabolism</i> , <b>2015</b> , 22, 922-35	24.6	233
84	DNMT1 is essential for mammary and cancer stem cell maintenance and tumorigenesis. <i>Nature Communications</i> , <b>2015</b> , 6, 6910	17.4	148
83	Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. <i>EMBO Journal</i> , <b>2014</b> , 33, 878-89	13	109
82	Histone-fold domain protein NF-Y promotes chromatin accessibility for cell type-specific master transcription factors. <i>Molecular Cell</i> , <b>2014</b> , 55, 708-22	17.6	108
81	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> , <b>2019</b> , 116, 9775-9784	11.5	67
80	An integrative systems genetic analysis of mammalian lipid metabolism. <i>Nature</i> , <b>2019</b> , 567, 187-193	50.4	63
79	Mitochondrial CoQ deficiency is a common driver of mitochondrial oxidants and insulin resistance. <i>ELife</i> , <b>2018</b> , 7,	8.9	61
78	mTORC1 Is a Major Regulatory Node in the FGF21 Signaling Network in Adipocytes. <i>Cell Reports</i> , <b>2016</b> , 17, 29-36	10.6	59
77	Intragenic Enhancers Attenuate Host Gene Expression. <i>Molecular Cell</i> , <b>2017</b> , 68, 104-117.e6	17.6	54
76	Selective photodegradation of 1-methylimidazole-2-thiol by the magnetic and dual conductive imprinted photocatalysts based on TiO <sub>2</sub> /Fe <sub>3</sub> O <sub>4</sub> /MWCNTs. <i>Chemical Engineering Journal</i> , <b>2014</b> , 240, 244-252	14.7	52
75	Sample Subset Optimization Techniques for Imbalanced and Ensemble Learning Problems in Bioinformatics Applications. <i>IEEE Transactions on Cybernetics</i> , <b>2014</b> , 44, 445-55	10.2	51
74	Multi-omic Profiling Reveals Dynamics of the Phased Progression of Pluripotency. <i>Cell Systems</i> , <b>2019</b> , 8, 427-445.e10	10.6	50
73	Impact of similarity metrics on single-cell RNA-seq data clustering. <i>Briefings in Bioinformatics</i> , <b>2019</b> , 20, 2316-2326	13.4	47
72	Global redox proteome and phosphoproteome analysis reveals redox switch in Akt. <i>Nature Communications</i> , <b>2019</b> , 10, 5486	17.4	36
71	A particle swarm based hybrid system for imbalanced medical data sampling. <i>BMC Genomics</i> , <b>2009</b> , 10 Suppl 3, S34	4.5	33

70	A multi-filter enhanced genetic ensemble system for gene selection and sample classification of microarray data. <i>BMC Bioinformatics</i> , <b>2010</b> , 11 Suppl 1, S5	3.6	33
69	Ensemble-Based Wrapper Methods for Feature Selection and Class Imbalance Learning. <i>Lecture Notes in Computer Science</i> , <b>2013</b> , 544-555	0.9	32
68	Gene-gene interaction filtering with ensemble of filters. <i>BMC Bioinformatics</i> , <b>2011</b> , 12 Suppl 1, S10	3.6	32
67	Ensemble deep learning in bioinformatics. <i>Nature Machine Intelligence</i> , <b>2020</b> , 2, 500-508	22.5	32
66	NF-Y controls fidelity of transcription initiation at gene promoters through maintenance of the nucleosome-depleted region. <i>Nature Communications</i> , <b>2019</b> , 10, 3072	17.4	25
65	Evaluating stably expressed genes in single cells. <i>GigaScience</i> , <b>2019</b> , 8,	7.6	23
64	scClassify: sample size estimation and multiscale classification of cells using single and multiple reference. <i>Molecular Systems Biology</i> , <b>2020</b> , 16, e9389	12.2	22
63	AdaSampling for Positive-Unlabeled and Label Noise Learning With Bioinformatics Applications. <i>IEEE Transactions on Cybernetics</i> , <b>2019</b> , 49, 1932-1943	10.2	20
62	Direction pathway analysis of large-scale proteomics data reveals novel features of the insulin action pathway. <i>Bioinformatics</i> , <b>2014</b> , 30, 808-14	7.2	20
61	Knowledge-Based Analysis for Detecting Key Signaling Events from Time-Series Phosphoproteomics Data. <i>PLoS Computational Biology</i> , <b>2015</b> , 11, e1004403	5	20
60	CiteFuse enables multi-modal analysis of CITE-seq data. <i>Bioinformatics</i> , <b>2020</b> , 36, 4137-4143	7.2	19
59	CNOT3-Dependent mRNA Deadenylation Safeguards the Pluripotent State. <i>Stem Cell Reports</i> , <b>2016</b> , 7, 897-910	8	18
58	Autoencoder-based cluster ensembles for single-cell RNA-seq data analysis. <i>BMC Bioinformatics</i> , <b>2019</b> , 20, 660	3.6	18
57	Positive-unlabeled ensemble learning for kinase substrate prediction from dynamic phosphoproteomics data. <i>Bioinformatics</i> , <b>2016</b> , 32, 252-9	7.2	17
56	A genetic ensemble approach for gene-gene interaction identification. <i>BMC Bioinformatics</i> , <b>2010</b> , 11, 524	3.6	17
55	Stability of Feature Selection Algorithms and Ensemble Feature Selection Methods in Bioinformatics <b>2013</b> , 333-352		15
54	Improving X!Tandem on peptide identification from mass spectrometry by self-boosted Percolator. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , <b>2012</b> , 9, 1273-80	3	14
53	KinasePA: Phosphoproteomics data annotation using hypothesis driven kinase perturbation analysis. <i>Proteomics</i> , <b>2016</b> , 16, 1868-71	4.8	14

52	Data Independent Acquisition Proteomic Analysis Can Discriminate between Actinic Keratosis, Bowen's Disease, and Cutaneous Squamous Cell Carcinoma. <i>Journal of Investigative Dermatology</i> , <b>2020</b> , 140, 212-222.e11	4.3	13
51	Integrated Glycoproteomics Identifies a Role of N-Glycosylation and Galectin-1 on Myogenesis and Muscle Development. <i>Molecular and Cellular Proteomics</i> , <b>2020</b> , 20, 100030	7.6	13
50	MiR-93-5p is a novel predictor of coronary in-stent restenosis. <i>Heart Asia</i> , <b>2019</b> , 11, e011134	1.9	12
49	An improved Akt reporter reveals intra- and inter-cellular heterogeneity and oscillations in signal transduction. <i>Journal of Cell Science</i> , <b>2017</b> , 130, 2757-2766	5.3	12
48	scDC: single cell differential composition analysis. <i>BMC Bioinformatics</i> , <b>2019</b> , 20, 721	3.6	12
47	A dynamic wavelet-based algorithm for pre-processing tandem mass spectrometry data. <i>Bioinformatics</i> , <b>2010</b> , 26, 2242-9	7.2	11
46	A clustering based hybrid system for biomarker selection and sample classification of mass spectrometry data. <i>Neurocomputing</i> , <b>2010</b> , 73, 2317-2331	5.4	10
45	Positive unlabeled learning via wrapper-based adaptive sampling <b>2017</b> ,		10
44	Sample Subset Optimization for Classifying Imbalanced Biological Data. <i>Lecture Notes in Computer Science</i> , <b>2011</b> , 333-344	0.9	10
43	Integrative analysis identifies co-dependent gene expression regulation of BRG1 and CHD7 at distal regulatory sites in embryonic stem cells. <i>Bioinformatics</i> , <b>2017</b> , 33, 1916-1920	7.2	9
42	Unraveling Kinase Activation Dynamics Using Kinase-Substrate Relationships from Temporal Large-Scale Phosphoproteomics Studies. <i>PLoS ONE</i> , <b>2016</b> , 11, e0157763	3.7	9
41	scReClassify: post hoc cell type classification of single-cell rNA-seq data. <i>BMC Genomics</i> , <b>2019</b> , 20, 913	4.5	9
40	ISL1 regulates peroxisome proliferator-activated receptor $\beta$ activation and early adipogenesis via bone morphogenetic protein 4-dependent and -independent mechanisms. <i>Molecular and Cellular Biology</i> , <b>2014</b> , 34, 3607-17	4.8	8
39	OCAP: an open comprehensive analysis pipeline for iTRAQ. <i>Bioinformatics</i> , <b>2012</b> , 28, 1404-5	7.2	8
38	Evaluating stably expressed genes in single cells		8
37	scMerge: Integration of multiple single-cell transcriptomics datasets leveraging stable expression and pseudo-replication		8
36	scClassify: hierarchical classification of cells		8
35	Global phosphoproteomics reveals DYRK1A regulates CDK1 activity in glioblastoma cells. <i>Cell Death Discovery</i> , <b>2021</b> , 7, 81	6.9	8

34	Clonal evolution in liver cancer at single-cell and single-variant resolution. <i>Journal of Hematology and Oncology</i> , <b>2021</b> , 14, 22	22.4	8
33	Signaling Heterogeneity is Defined by Pathway Architecture and Intercellular Variability in Protein Expression. <i>IScience</i> , <b>2021</b> , 24, 102118	6.1	8
32	PhosR enables processing and functional analysis of phosphoproteomic data. <i>Cell Reports</i> , <b>2021</b> , 34, 108776	7.6	7
31	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> , <b>2012</b> , 11, 3035-45	5.6	6
30	An Agent-Based Hybrid System for Microarray Data Analysis. <i>IEEE Intelligent Systems</i> , <b>2009</b> , 24, 53-63	4.2	6
29	Transcriptional network dynamics during the progression of pluripotency revealed by integrative statistical learning. <i>Nucleic Acids Research</i> , <b>2020</b> , 48, 1828-1842	20.1	6
28	Hierarchical kernel mixture models for the prediction of AIDS disease progression using HIV structural gp120 profiles. <i>BMC Genomics</i> , <b>2010</b> , 11 Suppl 4, S22	4.5	5
27	Temporal ordering of omics and multiomic events inferred from time-series data. <i>Npj Systems Biology and Applications</i> , <b>2020</b> , 6, 22	5	5
26	Feature selection revisited in the single-cell era. <i>Genome Biology</i> , <b>2021</b> , 22, 321	18.3	4
25	Hybrid Methods to Select Informative Gene Sets in Microarray Data Classification <b>2007</b> , 810-814		4
24	NoisyGOA: Noisy GO annotations prediction using taxonomic and semantic similarity. <i>Computational Biology and Chemistry</i> , <b>2016</b> , 65, 203-211	3.6	4
23	Enhanced structure and function of human pluripotent stem cell-derived beta-cells cultured on extracellular matrix. <i>Stem Cells Translational Medicine</i> , <b>2021</b> , 10, 492-505	6.9	4
22	QCMAP: An Interactive Web-Tool for Performance Diagnosis and Prediction of LC-MS Systems. <i>Proteomics</i> , <b>2019</b> , 19, e1900068	4.8	3
21	A Clustering Based Hybrid System for Mass Spectrometry Data Analysis. <i>Lecture Notes in Computer Science</i> , <b>2008</b> , 98-109	0.9	3
20	Benchmarking clustering algorithms on estimating the number of cell types from single-cell RNA-sequencing data.. <i>Genome Biology</i> , <b>2022</b> , 23, 49	18.3	3
19	A benchmark study of simulation methods for single-cell RNA sequencing data. <i>Nature Communications</i> , <b>2021</b> , 12, 6911	17.4	3
18	Decoding the function of bivalent chromatin in development and cancer. <i>Genome Research</i> , <b>2021</b> ,	9.7	3
17	hRUV: Hierarchical approach to removal of unwanted variation for large-scale metabolomics data		3

16	An Uncertainty Visual Analytics Framework for fMRI Functional Connectivity. <i>Neuroinformatics</i> , <b>2019</b> , 17, 211-223	3.2	2
15	Genetic Algorithm-Based Multi-objective Optimisation for QoS-Aware Web Services Composition. <i>Lecture Notes in Computer Science</i> , <b>2010</b> , 549-554	0.9	2
14	Multiagent Framework for Bio-data Mining. <i>Lecture Notes in Computer Science</i> , <b>2009</b> , 200-207	0.9	2
13	Akt phosphorylates insulin receptor substrate to limit PI3K-mediated PIP3 synthesis. <i>ELife</i> , <b>2021</b> , 10,	8.9	2
12	Spatial analysis for highly multiplexed imaging data to identify tissue microenvironments		2
11	A hierarchical approach to removal of unwanted variation for large-scale metabolomics data. <i>Nature Communications</i> , <b>2021</b> , 12, 4992	17.4	2
10	Modeling and simulation of Parallel Current Mode controlled boost converter <b>2008</b> ,		1
9	Functional analysis of the stable phosphoproteome reveals cancer vulnerabilities.. <i>Bioinformatics</i> , <b>2022</b> ,	7.2	1
8	Protocol for the processing and downstream analysis of phosphoproteomic data with PhosR. <i>STAR Protocols</i> , <b>2021</b> , 2, 100585	1.4	1
7	A benchmark study of simulation methods for single-cell RNA sequencing data		1
6	Cepo uncovers cell identity through differential stability		1
5	Computational systems approach towards phosphoproteomics and their downstream regulation.. <i>Proteomics</i> , <b>2022</b> , e2200068	4.8	1
4	Defining cell identity beyond the premise of differential gene expression. <i>Cell Regeneration</i> , <b>2021</b> , 10, 20	2.5	0
3	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> , <b>2022</b> , 4, lqac023	3.7	0
2	Uncovering cell identity through differential stability with Cepo. <i>Nature Computational Science</i> , <b>2021</b> , 1, 784-790		0
1	Mouse organogenesis atlas at single-cell resolution.. <i>Cell</i> , <b>2022</b> , 185, 1625-1627	56.2	