Jeffrey T Leek

List of Publications by Citations

Source: https://exaly.com/author-pdf/3771243/jeffrey-t-leek-publications-by-citations.pdf
Version: 2024-04-23

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.

92 papers	12,606 citations	37 h-index	106 g-index
106 ext. papers	17,617 ext. citations	12.9 avg, IF	6.82 L-index

#	Paper	IF	Citations
92	Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols*, **2016**, 11, 1650-67	18.8	2264
91	The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*, **2012**, 28, 882-3	7.2	1996
90	Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature Reviews Genetics*, **2010**, 11, 733-9	30.1	1232
89	Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genetics*, **2007**, 3, 1724-35	6	1165
88	Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature*, **2011**, 478, 519-23	50.4	541
87	Significance analysis of time course microarray experiments. *Proceedings of the National Academy of Sciences of the United States of America*, **2005**, 102, 12837-42	11.5	469
86	Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. *International Journal of Epidemiology*, **2012**, 41, 200-9	7.8	430
85	Ballgown bridges the gap between transcriptome assembly and expression analysis. *Nature Biotechnology*, **2015**, 33, 243-6	44.5	413
84	svaseq: removing batch effects and other unwanted noise from sequencing data. *Nucleic Acids Research*, **2014**, 42,	20.1	270
83	EDGE: extraction and analysis of differential gene expression. *Bioinformatics*, **2006**, 22, 507-8	7.2	250
82	Systems-level dynamic analyses of fate change in murine embryonic stem cells. *Nature*, **2009**, 462, 358-62	50.4	237
81	Cloud-scale RNA-sequencing differential expression analysis with Myrna. *Genome Biology*, **2010**, 11, R83	18.3	227
80	A general framework for multiple testing dependence. *Proceedings of the National Academy of Sciences of the United States of America*, **2008**, 105, 18718-23	11.5	225
79	Reproducible RNA-seq analysis using recount2. *Nature Biotechnology*, **2017**, 35, 319-321	44.5	211
78	Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. *Nature Neuroscience*, **2018**, 21, 1117-1125	25.5	176
77	Polyester: simulating RNA-seq datasets with differential transcript expression. *Bioinformatics*, **2015**, 31, 2778-84	7.2	160
76	Sequencing technology does not eliminate biological variability. *Nature Biotechnology*, **2011**, 29, 572-3	44.5	143

75	What Should Researchers Expect When They Replicate Studies? A Statistical View of Replicability in Psychological Science. *Perspectives on Psychological Science*, **2016**, 11, 539-44	9.8	118
74	ReCount: a multi-experiment resource of analysis-ready RNA-seq gene count datasets. *BMC Bioinformatics*, **2011**, 12, 449	3.6	115
73	Opinion: Reproducible research can still be wrong: adopting a prevention approach. *Proceedings of the National Academy of Sciences of the United States of America*, **2015**, 112, 1645-6	11.5	111
72	The optimal discovery procedure for large-scale significance testing, with applications to comparative microarray experiments. *Biostatistics*, **2007**, 8, 414-32	3.7	111
71	Developmental regulation of human cortex transcription and its clinical relevance at single base resolution. *Nature Neuroscience*, **2015**, 18, 154-161	25.5	110
70	Statistics: P values are just the tip of the iceberg. *Nature*, **2015**, 520, 612	50.4	108
69	On the design and analysis of gene expression studies in human populations. *Nature Genetics*, **2007**, 39, 807-8; author reply 808-9	36.3	101
68	An estimate of the science-wise false discovery rate and application to the top medical literature. *Biostatistics*, **2014**, 15, 1-12	3.7	86
67	Transparency and reproducibility in artificial intelligence. *Nature*, **2020**, 586, E14-E16	50.4	85
66	Significance analysis and statistical dissection of variably methylated regions. *Biostatistics*, **2012**, 13, 166-78	3.7	75
65	Inflammatory molecular signature associated with infectious agents in psychosis. *Schizophrenia Bulletin*, **2014**, 40, 963-72	1.3	71
64	Human splicing diversity and the extent of unannotated splice junctions across human RNA-seq samples on the Sequence Read Archive. *Genome Biology*, **2016**, 17, 266	18.3	65
63	Statistics. What is the question?. *Science*, **2015**, 347, 1314-5	33.3	59
62	Test set bias affects reproducibility of gene signatures. *Bioinformatics*, **2015**, 31, 2318-23	7.2	59
61	qSVA framework for RNA quality correction in differential expression analysis. *Proceedings of the National Academy of Sciences of the United States of America*, **2017**, 114, 7130-7135	11.5	58
60	Preserving biological heterogeneity with a permuted surrogate variable analysis for genomics batch correction. *Bioinformatics*, **2014**, 30, 2757-63	7.2	55
59	Evolution of cellular morpho-phenotypes in cancer metastasis. *Scientific Reports*, **2015**, 5, 18437	4.9	45
58	Dissecting inflammatory complications in critically injured patients by within-patient gene expression changes: a longitudinal clinical genomics study. *PLoS Medicine*, **2011**, 8, e1001093	11.6	44

57	Differential expression analysis of RNA-seq data at single-base resolution. *Biostatistics*, **2014**, 15, 413-26	3.7	42
56	Addressing confounding artifacts in reconstruction of gene co-expression networks. *Genome Biology*, **2019**, 20, 94	18.3	37
55	A simple and reproducible breast cancer prognostic test. *BMC Genomics*, **2013**, 14, 336	4.5	35
54	Removing batch effects for prediction problems with frozen surrogate variable analysis. *PeerJ*, **2014**, 2, e561	3.1	34
53	A direct approach to estimating false discovery rates conditional on covariates. *PeerJ*, **2018**, 6, e6035	3.1	34
52	Rail-RNA: scalable analysis of RNA-seq splicing and coverage. *Bioinformatics*, **2017**, 33, 4033-4040	7.2	33
51	Flexible expressed region analysis for RNA-seq with derfinder. *Nucleic Acids Research*, **2017**, 45, e9	20.1	32
50	Practical impacts of genomic data "cleaning" on biological discovery using surrogate variable analysis. *BMC Bioinformatics*, **2015**, 16, 372	3.6	32
49	Gene expression anti-profiles as a basis for accurate universal cancer signatures. *BMC Bioinformatics*, **2012**, 13, 272	3.6	32
48	Improving the value of public RNA-seq expression data by phenotype prediction. *Nucleic Acids Research*, **2018**, 46, e54	20.1	31
47	BatchQC: interactive software for evaluating sample and batch effects in genomic data. *Bioinformatics*, **2016**, 32, 3836-3838	7.2	31
46	Asymptotic conditional singular value decomposition for high-dimensional genomic data. *Biometrics*, **2011**, 67, 344-52	1.8	31
45	Is Most Published Research Really False?. *Annual Review of Statistics and Its Application*, **2017**, 4, 109-122	7.6	27
44	The tspair package for finding top scoring pair classifiers in R. *Bioinformatics*, **2009**, 25, 1203-4	7.2	26
43	A statistical definition for reproducibility and replicability		25
42	Cooperation between referees and authors increases peer review accuracy. *PLoS ONE*, **2011**, 6, e26895	3.7	24
41	The practical effect of batch on genomic prediction. *Statistical Applications in Genetics and Molecular Biology*, **2012**, 11, Article 10	1.2	18
40	A visual tool for defining reproducibility and replicability. *Nature Human Behaviour*, **2019**, 3, 650-652	12.8	15

39	Recounting the FANTOM CAGE-Associated Transcriptome. *Genome Research*, **2020**, 30, 1073-1081	9.7	13
38	A computationally efficient modular optimal discovery procedure. *Bioinformatics*, **2011**, 27, 509-15	7.2	12
37	A statistical approach to selecting and confirming validation targets in -omics experiments. *BMC Bioinformatics*, **2012**, 13, 150	3.6	10
36	Methods for correcting inference based on outcomes predicted by machine learning. *Proceedings of the National Academy of Sciences of the United States of America*, **2020**, 117, 30266-30275	11.5	9
35	How to share data for collaboration. *American Statistician*, **2018**, 72, 53-57	5	8
34	A randomized trial in a massive online open course shows people don't know what a statistically significant relationship looks like, but they can learn. *PeerJ*, **2014**, 2, e589	3.1	8
33	The Democratization of Data Science Education. *American Statistician*, **2020**, 74, 1-7	5	8
32	The Joint Null Criterion for Multiple Hypothesis Tests. *Statistical Applications in Genetics and Molecular Biology*, **2011**, 10,	1.2	7
31	Developmental and genetic regulation of the human cortex transcriptome in schizophrenia		7
30	SVAw - a web-based application tool for automated surrogate variable analysis of gene expression studies. *Source Code for Biology and Medicine*, **2013**, 8, 8	1.9	6
29	Gene set bagging for estimating the probability a statistically significant result will replicate. *BMC Bioinformatics*, **2013**, 14, 360	3.6	6
28	Inverting the model of genomics data sharing with the NHGRI Genomic Data Science Analysis, Visualization, and Informatics Lab-space.. *Cell Genomics*, **2022**, 2, 100085-100085		6
27	Flexible expressed region analysis for RNA-seq with derfinder		6
26	Rail-dbGaP: analyzing dbGaP-protected data in the cloud with Amazon Elastic MapReduce. *Bioinformatics*, **2016**, 32, 2551-3	7.2	5
25	A decision-theory approach to interpretable set analysis for high-dimensional data. *Biometrics*, **2013**, 69, 614-23	1.8	5
24	Personalized medicine: Keep a way open for tailored treatments. *Nature*, **2012**, 484, 318	50.4	5
23	Rail-RNA: Scalable analysis of RNA-seq splicing and coverage		5
22	recount: A large-scale resource of analysis-ready RNA-seq expression data		5

21	Inverting the model of genomics data sharing with the NHGRI Genomic Data Science Analysis, Visualization, and Informatics Lab-space (AnVIL)		5
20	Transcriptional profile of platelets and iPSC-derived megakaryocytes from whole-genome and RNA sequencing. *Blood*, **2021**, 137, 959-968	2.2	5
19	regionReport: Interactive reports for region-based analyses. *F1000Research*, **2015**, 4, 105	3.6	4
18	regionReport: Interactive reports for region-level and feature-level genomic analyses. *F1000Research*, **2015**, 4, 105	3.6	4
17	recount3: summaries and queries for large-scale RNA-seq expression and splicing. *Genome Biology*, **2021**, 22, 323	18.3	3
16	A direct approach to estimating false discovery rates conditional on covariates		3
15	Strategies for cellular deconvolution in human brain RNA sequencing data		3
14	recount-brain: a curated repository of human brain RNA-seq datasets metadata		3
13	Sequestration: inadvertently killing biomedical research to score political points. *Genome Biology*, **2013**, 14, 109	18.3	2
12	Capturing Heterogeneity in Gene Expression Studies by "Surrogate Variable Analysis". *PLoS Genetics*, **2005**, preprint, e161	6	2
11	Is most published research really false?		2
10	A framework for RNA quality correction in differential expression analysis		2
9	Improving the value of public RNA-seq expression data by phenotype prediction		2
8	Addressing confounding artifacts in reconstruction of gene co-expression networks		2
7	RNA-seq transcript quantification from reduced-representation data in recount2		2
6	Gene and protein expression in human megakaryocytes derived from induced pluripotent stem cells. *Journal of Thrombosis and Haemostasis*, **2021**, 19, 1783-1799	15.4	2
5	Post-prediction inference		1
4	Genomic and clinical predictors for improving estimator precision in randomized trials of breast cancer treatments. *Contemporary Clinical Trials Communications*, **2016**, 3, 48-54	1.8	0

3	Strategies for cellular deconvolution in human brain RNA sequencing data. *F1000Research*,10, 750	3.6	0
2	Discussion of "Visualizing statistical models: Removing the blindfold" *Statistical Analysis and Data Mining*, **2015**, 8, 240-241	1.4	
1	Measurement, Summary, and Methodological Variation in RNA-sequencing **2014**, 115-128		