Lior S Pachter

List of Publications by Citations

194 papers	66,331 citations	56 h-index	224 g-index
224 ext. papers	81,759 ext. citations	13.1 avg, IF	8.01 L-index

#	Paper	IF	Citations
194	Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, **2010**, 28, 511-5	44.5	10225
193	TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, **2009**, 25, 1105-11	7.2	9121
192	Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, **2012**, 7, 562-78	18.8	8342
191	Initial sequencing and comparative analysis of the mouse genome. *Nature*, **2002**, 420, 520-62	50.4	5376
190	Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, **2007**, 447, 799-816	50.4	4121
189	Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, **2016**, 34, 525-7	44.5	3742
188	Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature Biotechnology* , **2013**, 31, 46-53	44.5	2465
187	Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, **2004**, 432, 695-716	50.4	2143
186	Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *New England Journal of Medicine*, **2015**, 372, 2481-98	59.2	1828
185	The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science*, **2004**, 306, 636-40	33.3	1692
184	Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature*, **2004** , 428, 493-521	50.4	1689
183	Evolution of genes and genomes on the Drosophila phylogeny. *Nature*, **2007**, 450, 203-18	50.4	1586
182	VISTA: computational tools for comparative genomics. *Nucleic Acids Research*, **2004**, 32, W273-9	20.1	1441
181	Disordered microbial communities in asthmatic airways. *PLoS ONE*, **2010**, 5, e8578	3.7	1085
180	Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biology*, **2011**, 12, R22	18.3	804
179	Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics*, **2011**, 27, 2325-9	7.2	716
178	Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature Methods*, **2017**, 14, 687-690	21.6	706

177	Streaming fragment assignment for real-time analysis of sequencing experiments. *Nature Methods*, **2013**, 10, 71-3	21.6	649
176	Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures. *Nature*, **2007**, 450, 219-32	50.4	506
175	Population genomics: whole-genome analysis of polymorphism and divergence in Drosophila simulans. *PLoS Biology*, **2007**, 5, e310	9.7	479
174	Phylogenetic shadowing of primate sequences to find functional regions of the human genome. *Science*, **2003**, 299, 1391-4	33.3	431
173	A genome-wide map of conserved microRNA targets in C. elegans. *Current Biology*, **2006**, 16, 460-71	6.3	353
172	AVID: A global alignment program. *Genome Research*, **2003**, 13, 97-102	9.7	328
171	rVista for comparative sequence-based discovery of functional transcription factor binding sites. *Genome Research*, **2002**, 12, 832-9	9.7	282
170	Multiplexed RNA structure characterization with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Proceedings of the National Academy of Sciences of the United States of America*, **2011**, 108, 11063-8	11.5	280
169	Fast statistical alignment. *PLoS Computational Biology*, **2009**, 5, e1000392	5	252
168	Active conservation of noncoding sequences revealed by three-way species comparisons. *Genome Research*, **2000**, 10, 1304-6	9.7	240
167	Human and mouse gene structure: comparative analysis and application to exon prediction. *Genome Research*, **2000**, 10, 950-8	9.7	225
166	Bioinformatics for whole-genome shotgun sequencing of microbial communities. *PLoS Computational Biology*, **2005**, 1, 106-12	5	221
165	MAVID: constrained ancestral alignment of multiple sequences. *Genome Research*, **2004**, 14, 693-9	9.7	209
164	Viral population estimation using pyrosequencing. *PLoS Computational Biology*, **2008**, 4, e1000074	5	174
163	Identification and correction of systematic error in high-throughput sequence data. *BMC Bioinformatics*, **2011**, 12, 451	3.6	173
162	Strategies and tools for whole-genome alignments. *Genome Research*, **2003**, 13, 73-80	9.7	165
161	Analyses of deep mammalian sequence alignments and constraint predictions for 1% of the human genome. *Genome Research*, **2007**, 17, 760-74	9.7	163
160	Binding site turnover produces pervasive quantitative changes in transcription factor binding between closely related Drosophila species. *PLoS Biology*, **2010**, 8, e1000343	9.7	154

159	Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. *Science*, **2015**, 350, 1251-5	33.3	139
158	SLAM: cross-species gene finding and alignment with a generalized pair hidden Markov model. *Genome Research*, **2003**, 13, 496-502	9.7	112
157	A dynamic intron retention program enriched in RNA processing genes regulates gene expression during terminal erythropoiesis. *Nucleic Acids Research*, **2016**, 44, 838-51	20.1	111
156	Exon-level microarray analyses identify alternative splicing programs in breast cancer. *Molecular Cancer Research*, **2010**, 8, 961-74	6.6	101
155	Mapping and identification of essential gene functions on the X chromosome of Drosophila. *EMBO Reports*, **2002**, 3, 34-8	6.5	99
154	Genome methylation in D. melanogaster is found at specific short motifs and is independent of DNMT2 activity. *Genome Research*, **2014**, 24, 821-30	9.7	96
153	Modeling and automation of sequencing-based characterization of RNA structure. *Proceedings of the National Academy of Sciences of the United States of America*, **2011**, 108, 11069-74	11.5	95
152	Multimodal Analysis of Cell Types in a Hypothalamic Node Controlling Social Behavior. *Cell*, **2019**, 179, 713-728.e17	56.2	84
151	Fast and accurate single-cell RNA-seq analysis by clustering of transcript-compatibility counts. *Genome Biology*, **2016**, 17, 112	18.3	76
150	Human intestinal tissue with adult stem cell properties derived from pluripotent stem cells. *Stem Cell Reports*, **2014**, 2, 838-52	8	72
149	A dynamic alternative splicing program regulates gene expression during terminal erythropoiesis. *Nucleic Acids Research*, **2014**, 42, 4031-42	20.1	65
148	Multiple alignment by sequence annealing. *Bioinformatics*, **2007**, 23, e24-9	7.2	60
147	Tropical geometry of statistical models. *Proceedings of the National Academy of Sciences of the United States of America*, **2004**, 101, 16132-7	11.5	60
146	CGAL: computing genome assembly likelihoods. *Genome Biology*, **2013**, 14, R8	18.3	59
145	Why Neighbor-Joining Works. *Algorithmica*, **2009**, 54, 1-24	0.9	59
144	SHAPE-Seq: High-Throughput RNA Structure Analysis. *Current Protocols in Chemical Biology*, **2012**, 4, 275-97	1.8	58
143	Gene-level differential analysis at transcript-level resolution. *Genome Biology*, **2018**, 19, 53	18.3	57
142	Applications of generalized pair hidden Markov models to alignment and gene finding problems. *Journal of Computational Biology*, **2002**, 9, 389-99	1.7	57

141	Modular and efficient pre-processing of single-cell RNA-seq		57
140	A curated database reveals trends in single-cell transcriptomics. *Database: the Journal of Biological Databases and Curation*, **2020**, 2020,	5	56
139	Single-cell analysis at the threshold. *Nature Biotechnology*, **2016**, 34, 1111-1118	44.5	55
138	The barcode, UMI, set format and BUStools. *Bioinformatics*, **2019**, 35, 4472-4473	7.2	54
137	A discriminative learning approach to differential expression analysis for single-cell RNA-seq. *Nature Methods*, **2019**, 16, 163-166	21.6	53
136	Modular, efficient and constant-memory single-cell RNA-seq preprocessing. *Nature Biotechnology*, **2021**, 39, 813-818	44.5	53
135	Reference based annotation with GeneMapper. *Genome Biology*, **2006**, 7, R29	18.3	50
134	Association mapping from sequencing reads using -mers. *ELife*, **2018**, 7,	8.9	48
133	Analysis of epistatic interactions and fitness landscapes using a new geometric approach. *BMC Evolutionary Biology*, **2007**, 7, 60	3	47
132	Development of a low bias method for characterizing viral populations using next generation sequencing technology. *PLoS ONE*, **2010**, 5, e13564	3.7	45
131	A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature*, **2021**, 598, 86-102	50.4	44
130	HMM sampling and applications to gene finding and alternative splicing. *Bioinformatics*, **2003**, 19 Suppl 2, ii36-41	7.2	43
129	Intraspecies sequence comparisons for annotating genomes. *Genome Research*, **2004**, 14, 2406-11	9.7	42
128	Evolution at the nucleotide level: the problem of multiple whole-genome alignment. *Human Molecular Genetics*, **2006**, 15 Spec No 1, R51-6	5.6	41
127	Parametric inference for biological sequence analysis. *Proceedings of the National Academy of Sciences of the United States of America*, **2004**, 101, 16138-43	11.5	40
126	MAVID multiple alignment server. *Nucleic Acids Research*, **2003**, 31, 3525-6	20.1	40
125	A diverse epigenetic landscape at human exons with implication for expression. *Nucleic Acids Research*, **2015**, 43, 3498-508	20.1	38
124	Shape-based peak identification for ChIP-Seq. *BMC Bioinformatics*, **2011**, 12, 15	3.6	38

123	Convex Rank Tests and Semigraphoids. *SIAM Journal on Discrete Mathematics*, **2009**, 23, 1117-1134	0.7	38
122	Identification of transposable elements using multiple alignments of related genomes. *Genome Research*, **2006**, 16, 260-70	9.7	38
121	Interpretable factor models of single-cell RNA-seq via variational autoencoders. *Bioinformatics*, **2020**, 36, 3418-3421	7.2	37
120	Multiple-sequence functional annotation and the generalized hidden Markov phylogeny. *Bioinformatics*, **2004**, 20, 1850-60	7.2	37
119	Reconstructing trees from subtree weights. *Applied Mathematics Letters*, **2004**, 17, 615-621	3.5	37
118	Constrained Optimization for UAV Task Assignment **2004**,		37
117	Optimization of air vehicles operations using mixed-integer linear programming. *Journal of the Operational Research Society*, **2007**, 58, 516-527	2	36
116	Highly multiplexed single-cell RNA-seq by DNA oligonucleotide tagging of cellular proteins. *Nature Biotechnology*, **2020**, 38, 35-38	44.5	35
115	Accurate design of translational output by a neural network model of ribosome distribution. *Nature Structural and Molecular Biology*, **2018**, 25, 577-582	17.6	35
114	Forcing matchings on square grids. *Discrete Mathematics*, **1998**, 190, 287-294	0.7	34
113	Phyloepigenomic comparison of great apes reveals a correlation between somatic and germline methylation states. *Genome Research*, **2011**, 21, 2049-57	9.7	32
112	Comparison of pattern detection methods in microarray time series of the segmentation clock. *PLoS ONE*, **2008**, 3, e2856	3.7	32
111	Specific alignment of structured RNA: stochastic grammars and sequence annealing. *Bioinformatics*, **2008**, 24, 2677-83	7.2	31
110	Differential analysis of RNA-Seq incorporating quantification uncertainty		31
109	UAV Task Assignment with Timing Constraints via Mixed-Integer Linear Programming **2004**,		30
108	RNA Velocity: Molecular Kinetics from Single-Cell RNA-Seq. *Molecular Cell*, **2018**, 72, 7-9	17.6	30
107	RNA structure characterization from chemical mapping experiments **2011**,		29
106	Parametric alignment of Drosophila genomes. *PLoS Computational Biology*, **2006**, 2, e73	5	28

105	Swab-Seq: A high-throughput platform for massively scaled up SARS-CoV-2 testing **2021**,		28
104	On the optimality of the neighbor-joining algorithm. *Algorithms for Molecular Biology*, **2008**, 3, 5	1.8	25
103	Combining statistical alignment and phylogenetic footprinting to detect regulatory elements. *Bioinformatics*, **2008**, 24, 1236-42	7.2	25
102	The Mathematics of Phylogenomics. *SIAM Review*, **2007**, 49, 3-31	7.4	25
101	Accurate identification of novel human genes through simultaneous gene prediction in human, mouse, and rat. *Genome Research*, **2004**, 14, 661-4	9.7	25
100	Principles of open source bioinstrumentation applied to the poseidon syringe pump system. *Scientific Reports*, **2019**, 9, 12385	4.9	24
99	The NIH BD2K center for big data in translational genomics. *Journal of the American Medical Informatics Association: JAMIA*, **2015**, 22, 1143-7	8.6	24
98	Rational experiment design for sequencing-based RNA structure mapping. *Rna*, **2014**, 20, 1864-77	5.8	24
97	A transcriptomic and epigenomic cell atlas of the mouse primary motor cortex. *Nature*, **2021**, 598, 103-110	50.4	23
96	An integrated transcriptomic and epigenomic atlas of mouse primary motor cortex cell types		23
95	Protein velocity and acceleration from single-cell multiomics experiments. *Genome Biology*, **2020**, 21, 39	18.3	19
94	Beyond pairwise distances: neighbor-joining with phylogenetic diversity estimates. *Molecular Biology and Evolution*, **2006**, 23, 491-8	8.3	19
93	Fragment assignment in the cloud with eXpress-D. *BMC Bioinformatics*, **2013**, 14, 358	3.6	18
92	Updating RNA-Seq analyses after re-annotation. *Bioinformatics*, **2013**, 29, 1631-7	7.2	18
91	RNA-Seq and find: entering the RNA deep field. *Genome Medicine*, **2011**, 3, 74	14.4	18
90	Toward the human genotope. *Bulletin of Mathematical Biology*, **2007**, 69, 2723-35	2.1	18
89	scvi-tools: a library for deep probabilistic analysis of single-cell omics data		18
88	Estimating intrinsic and extrinsic noise from single-cell gene expression measurements. *Statistical Applications in Genetics and Molecular Biology*, **2016**, 15, 447-471	1.2	17

87	Identification of evolutionary hotspots in the rodent genomes. *Genome Research*, **2004**, 14, 574-9	9.7	17
86	A curated database reveals trends in single-cell transcriptomics		17
85	A dictionary-based approach for gene annotation. *Journal of Computational Biology*, **1999**, 6, 419-30	1.7	16
84	Finding Convex Sets Among Points in the Plane. *Discrete and Computational Geometry*, **1998**, 19, 405-410	0.6	15
83	Subtree power analysis and species selection for comparative genomics. *Proceedings of the National Academy of Sciences of the United States of America*, **2005**, 102, 7900-5	11.5	15
82	The neighbor-net algorithm. *Advances in Applied Mathematics*, **2011**, 47, 240-258	0.8	14
81	Visualization of multiple genome annotations and alignments with the K-BROWSER. *Genome Research*, **2004**, 14, 716-20	9.7	14
80	Pregnancy-Induced Changes in Systemic Gene Expression among Healthy Women and Women with Rheumatoid Arthritis. *PLoS ONE*, **2015**, 10, e0145204	3.7	14
79	Massively scaled-up testing for SARS-CoV-2 RNA via next-generation sequencing of pooled and barcoded nasal and saliva samples. *Nature Biomedical Engineering*, **2021**, 5, 657-665	19	14
78	PROBer Provides a General Toolkit for Analyzing Sequencing-Based Toeprinting Assays. *Cell Systems*, **2017**, 4, 568-574.e7	10.6	13
77	Forcing numbers of stop signs. *Theoretical Computer Science*, **2003**, 303, 409-416	1.1	13
76	Picking alignments from (Steiner) trees. *Journal of Computational Biology*, **2003**, 10, 509-20	1.7	13
75	Museum of spatial transcriptomics.. *Nature Methods*, **2022**,	21.6	13
74	A multimodal cell census and atlas of the mammalian primary motor cortex		12
73	The Specious Art of Single-Cell Genomics		12
72	Barcode identification for single cell genomics. *BMC Bioinformatics*, **2019**, 20, 32	3.6	11
71	Fusion detection and quantification by pseudoalignment		11
70	The Lair: a resource for exploratory analysis of published RNA-Seq data. *BMC Bioinformatics*, **2016**, 17, 490	3.6	10

69	Combinatorics of least-squares trees. *Proceedings of the National Academy of Sciences of the United States of America*, **2008**, 105, 13206-11	11.5	10
68	Odd-paired is a pioneer-like factor that coordinates with Zelda to control gene expression in embryos. *ELife*, **2020**, 9,	8.9	10
67	Quantifying the tradeoff between sequencing depth and cell number in single-cell RNA-seq		10
66	Museum of Spatial Transcriptomics		10
65	MetMap enables genome-scale Methyltyping for determining methylation states in populations. *PLoS Computational Biology*, **2010**, 6, e1000888	5	9
64	Coverage statistics for sequence census methods. *BMC Bioinformatics*, **2010**, 11, 430	3.6	9
63	The computational challenges of applying comparative-based computational methods to whole genomes. *Briefings in Bioinformatics*, **2002**, 3, 18-22	13.4	9
62	A Python library for probabilistic analysis of single-cell omics data.. *Nature Biotechnology*, **2022**, 40, 163-166	44.5	9
61	Transcriptomic response of Drosophila melanogaster pupae developed in hypergravity. *Genomics*, **2016**, 108, 158-167	4.3	9
60	Expression reflects population structure. *PLoS Genetics*, **2018**, 14, e1007841	6	9
59	SLAM web server for comparative gene finding and alignment. *Nucleic Acids Research*, **2003**, 31, 3507-9	20.1	8
58	Decrease inACE2mRNA expression in aged mouse lung		8
57	Special function methods for bursty models of transcription. *Physical Review E*, **2020**, 102, 022409	2.4	8
56	The cyclohedron test for finding periodic genes in time course expression studies. *Statistical Applications in Genetics and Molecular Biology*, **2007**, 6, Article 21	1.2	7
55	Combinatorial Approaches and Conjectures for 2-Divisibility Problems Concerning Domino Tilings of Polyominoes. *Electronic Journal of Combinatorics*, **1997**, 4,	1.1	7
54	A direct comparison of genome alignment and transcriptome pseudoalignment		7
53	Phylogenetic Profiling of Insertions and Deletions in Vertebrate Genomes. *Lecture Notes in Computer Science*, **2006**, 265-280	0.9	7
52	RefShannon: A genome-guided transcriptome assembler using sparse flow decomposition. *PLoS ONE*, **2020**, 15, e0232946	3.7	6

51	Pregnancy-induced gene expression changes in vivo among women with rheumatoid arthritis: a pilot study. *Arthritis Research and Therapy*, **2017**, 19, 104	5.7	6
50	Constructing status injective graphs. *Discrete Applied Mathematics*, **1997**, 80, 107-113	1	6
49	Large multiple organism gene finding by collapsed Gibbs sampling. *Journal of Computational Biology*, **2005**, 12, 599-608	1.7	6
48	Whole-animal multiplexed single-cell RNA-seq reveals transcriptional shifts across medusa cell types. *Science Advances*, **2021**, 7, eabh1683	14.3	6
47	Markedly heterogeneous COVID-19 testing plans among US colleges and universities		6
46	Identification of transcriptional signatures for cell types from single-cell RNA-Seq		6
45	Quantifying uniformity of mapped reads. *Bioinformatics*, **2012**, 28, 2680-2	7.2	5
44	Applications of generalized pair hidden Markov models to alignment and gene finding problems **2001**,		5
43	Isoform cell-type specificity in the mouse primary motor cortex. *Nature*, **2021**, 598, 195-199	50.4	5
42	Reliable and accurate diagnostics from highly multiplexed sequencing assays. *Scientific Reports*, **2020**, 10, 21759	4.9	5
41	Whole Animal Multiplexed Single-Cell RNA-Seq Reveals Plasticity of Clytia Medusa Cell Types		5
40	Controlling for conservation in genome-wide DNA methylation studies. *BMC Genomics*, **2015**, 16, 420	4.5	4
39	Zika infection of neural progenitor cells perturbs transcription in neurodevelopmental pathways. *PLoS ONE*, **2017**, 12, e0175744	3.7	4
38	Affine and Projective Tree Metric Theorems. *Annals of Combinatorics*, **2013**, 17, 205-228	0.7	4
37	Human and mouse gene structure **2000**,		4
36	Fast and accurate diagnostics from highly multiplexed sequencing assays		4
35	Intrinsic and extrinsic noise are distinguishable in a synthesis export degradation model of mRNA production		4
34	Interpretable factor models of single-cell RNA-seq via variational autoencoders		4

33	Benchmarking of lightweight-mapping based single-cell RNA-seq pre-processing		4
32	Normalization of single-cell RNA-seq counts by log(x + 1)* or log(1 + x). *Bioinformatics*, **2021**,	7.2	4
31	Patterns of gene duplication and intron loss in the ENCODE regions suggest a confounding factor. *Genomics*, **2007**, 90, 44-8	4.3	3
30	The Barcode, UMI, Set format and BUStools		3
29	Barcode identification for single cell genomics		3
28	Association Mapping from Sequencing Reads Using K-mers		3
27	Design principles for open source bioinstrumentation: the poseidon syringe pump system as an example		3
26	Structural variation among wild and industrial strains of Penicillium chrysogenum. *PLoS ONE*, **2014**, 9, e96784	3.7	2
25	Tracing the most parsimonious indel history. *Journal of Computational Biology*, **2011**, 18, 967-86	1.7	2
24	Pair hidden Markov models **2005**,		2
23	Reference-free Association Mapping from Sequencing Reads Using k-mers. *Bio-protocol*, **2020**, 10, e3815	0.9	2
22	Isoform cell type specificity in the mouse primary motor cortex		2
21	Normalization of single-cell RNA-seq counts by log(x+1)* or log(1+x)*		2
20	Compositional Data Analysis is necessary for simulating and analyzing RNA-Seq data		2
19	RNA velocity and protein acceleration from single-cell multiomics experiments		2
18	Analytical solutions of the chemical master equation with bursty production and isomerization reactions		2
17	Low-cost, scalable, and automated fluid sampling for fluidics applications. *HardwareX*, **2021**, 10, e00201	2.7	2
16	Factor analysis for survival time prediction with informative censoring and diverse covariates. *Statistics in Medicine*, **2019**, 38, 3719-3732	2.3	1

15	Identification and correction of systematic error in high-throughput sequence data. *Nature Precedings*, **2011**,		1
14	Multiple organism gene finding by collapsed gibbs sampling **2004**,		1
13	Computation **2005**, 43-84		1
12	Picking alignments from (steiner) trees **2002**,		1
11	Prober: A general toolkit for analyzing sequencing-based □oeprinting□assays		1
10	Addressing the pooled amplification paradox with unique molecular identifiers in single-cell RNA-seq		1
9	Gene-level differential analysis at transcript-level resolution		1
8	BUTTERFLY: addressing the pooled amplification paradox with unique molecular identifiers in single-cell RNA-seq. *Genome Biology*, **2021**, 22, 174	18.3	1
7	Direct simulation of a stochastically driven multi-step birth-death process		1
6	Interpretable and tractable models of transcriptional noise for the rational design of single-molecule quantification experiments		1
5	Determining coding CpG islands by identifying regions significant for pattern statistics on Markov chains. *Statistical Applications in Genetics and Molecular Biology*, **2011**, 10,	1.2	0
4	SWALO: scaffolding with assembly likelihood optimization. *Nucleic Acids Research*, **2021**, 49, e117	20.1	0
3	Deterministic column subset selection for single-cell RNA-Seq. *PLoS ONE*, **2019**, 14, e0210571	3.7	
2	Transcript Abundance Estimation and the Laminar Packing Problem. *Lecture Notes in Computer Science*, **2019**, 203-211	0.9	
1	Exploring the genetic basis of variation in gene predictions with a synthetic association study. *PLoS ONE*, **2010**, 5, e11645	3.7	