Matthew Stephens

List of Publications by Citations

101 papers	52,336 citations	51 h-index	110 g-index
110 ext. papers	61,208 ext. citations	14.1 avg, IF	7.93 L-index

#	Paper	IF	Citations
101	Inference of population structure using multilocus genotype data. *Genetics*, **2000**, 155, 945-59	4	22315
100	Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **2003**, 164, 1567-87	4	5901
99	A second generation human haplotype map of over 3.1 million SNPs. *Nature*, **2007**, 449, 851-61	50.4	3647
98	Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*, **2015**, 348, 648-60	33.3	3242
97	RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Research*, **2008**, 18, 1509-17	9.7	2051
96	A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *American Journal of Human Genetics*, **2006**, 78, 629-44	11	1493
95	Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, **2012**, 44, 821-4	36.3	1480
94	Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature Genetics*, **2012**, 44, 955-9	36.3	1292
93	Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature*, **2010**, 464, 768-72	50.4	993
92	fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*, **2014**, 197, 573-89	4	839
91	Genotype imputation with thousands of genomes. *G3: Genes, Genomes, Genetics*, **2011**, 1, 457-70	3.2	719
90	Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics*, **2003**, 165, 2213-33	4	652
89	Dealing with label switching in mixture models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **2000**, 62, 795-809	3.9	561
88	DNase I sensitivity QTLs are a major determinant of human expression variation. *Nature*, **2012**, 482, 390-4	50.4	479
87	High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS Genetics*, **2008**, 4, e1000214	6	456
86	Polygenic modeling with bayesian sparse linear mixed models. *PLoS Genetics*, **2013**, 9, e1003264	6	440
85	Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods*, **2014**, 11, 407-9	21.6	432

84	Imputation-based analysis of association studies: candidate regions and quantitative traits. *PLoS Genetics*, **2007**, 3, e114	6	386
83	Interpreting principal component analyses of spatial population genetic variation. *Nature Genetics*, **2008**, 40, 646-9	36.3	381
82	Bayesian statistical methods for genetic association studies. *Nature Reviews Genetics*, **2009**, 10, 681-90	30.1	339
81	Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, **2016**, 48, 94-100	36.3	226
80	Bayesian variable selection regression for genome-wide association studies and other large-scale problems. *Annals of Applied Statistics*, **2011**, 5,	2.1	220
79	False discovery rates: a new deal. *Biostatistics*, **2017**, 18, 275-294	3.7	201
78	Genome-wide association of lipid-lowering response to statins in combined study populations. *PLoS ONE*, **2010**, 5, e9763	3.7	185
77	A statistical framework for joint eQTL analysis in multiple tissues. *PLoS Genetics*, **2013**, 9, e1003486	6	175
76	A statin-dependent QTL for GATM expression is associated with statin-induced myopathy. *Nature*, **2013**, 502, 377-80	50.4	160
75	Polymorphisms of the HNF1A gene encoding hepatocyte nuclear factor-1 alpha are associated with C-reactive protein. *American Journal of Human Genetics*, **2008**, 82, 1193-201	11	155
74	Dissecting the regulatory architecture of gene expression QTLs. *Genome Biology*, **2012**, 13, R7	18.3	151
73	A unified framework for association analysis with multiple related phenotypes. *PLoS ONE*, **2013**, 8, e65245	3.7	146
72	Practical issues in imputation-based association mapping. *PLoS Genetics*, **2008**, 4, e1000279	6	144
71	Scalable Variational Inference for Bayesian Variable Selection in Regression, and Its Accuracy in Genetic Association Studies. *Bayesian Analysis*, **2012**, 7,	2.3	126
70	A multivariate genome-wide association analysis of 10 LDL subfractions, and their response to statin treatment, in 1868 Caucasians. *PLoS ONE*, **2015**, 10, e0120758	3.7	118
69	The impact of sex on gene expression across human tissues. *Science*, **2020**, 369,	33.3	100
68	Analysis of population structure: a unifying framework and novel methods based on sparse factor analysis. *PLoS Genetics*, **2010**, 6, e1001117	6	95
67	Flexible statistical methods for estimating and testing effects in genomic studies with multiple conditions. *Nature Genetics*, **2019**, 51, 187-195	36.3	94

66	A simple new approach to variable selection in regression, with application to genetic fine mapping. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **2020**, 82, 1273-1300	3.9	91
65	A Simple Model-Based Approach to Inferring and Visualizing Cancer Mutation Signatures. *PLoS Genetics*, **2015**, 11, e1005657	6	87
64	Thousands of novel translated open reading frames in humans inferred by ribosome footprint profiling. *ELife*, **2016**, 5,	8.9	82
63	The GTEx Consortium atlas of genetic regulatory effects across human tissues		81
62	Interactions between glucocorticoid treatment and cis-regulatory polymorphisms contribute to cellular response phenotypes. *PLoS Genetics*, **2011**, 7, e1002162	6	80
61	The contribution of RNA decay quantitative trait loci to inter-individual variation in steady-state gene expression levels. *PLoS Genetics*, **2012**, 8, e1003000	6	80
60	BAYESIAN LARGE-SCALE MULTIPLE REGRESSION WITH SUMMARY STATISTICS FROM GENOME-WIDE ASSOCIATION STUDIES. *Annals of Applied Statistics*, **2017**, 11, 1561-1592	2.1	73
59	The genetic architecture of gene expression levels in wild baboons. *ELife*, **2015**, 4,	8.9	73
58	Cell type-specific genetic regulation of gene expression across human tissues. *Science*, **2020**, 369,	33.3	68
57	Visualizing the structure of RNA-seq expression data using grade of membership models. *PLoS Genetics*, **2017**, 13, e1006599	6	66
56	Genome-wide association study of d-amphetamine response in healthy volunteers identifies putative associations, including cadherin 13 (CDH13). *PLoS ONE*, **2012**, 7, e42646	3.7	65
55	Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. *Nature Genetics*, **2020**, 52, 740-747	36.3	64
54	Genotyping Polyploids from Messy Sequencing Data. *Genetics*, **2018**, 210, 789-807	4	64
53	Genetic analyses support the contribution of mRNA N-methyladenosine (mA) modification to human disease heritability. *Nature Genetics*, **2020**, 52, 939-949	36.3	52
52	USING LINEAR PREDICTORS TO IMPUTE ALLELE FREQUENCIES FROM SUMMARY OR POOLED GENOTYPE DATA. *Annals of Applied Statistics*, **2010**, 4, 1158-1182	2.1	52
51	Promoter shape varies across populations and affects promoter evolution and expression noise. *Nature Genetics*, **2017**, 49, 550-558	36.3	51
50	Epigenetic modifications are associated with inter-species gene expression variation in primates. *Genome Biology*, **2014**, 15, 547	18.3	49
49	Large-scale genome-wide enrichment analyses identify new trait-associated genes and pathways across 31 human phenotypes. *Nature Communications*, **2018**, 9, 4361	17.4	48

48	Integrated enrichment analysis of variants and pathways in genome-wide association studies indicates central role for IL-2 signaling genes in type 1 diabetes, and cytokine signaling genes in Crohn's disease. *PLoS Genetics*, **2013**, 9, e1003770	6	47
47	An estimate of the average number of recessive lethal mutations carried by humans. *Genetics*, **2015** , 199, 1243-54	4	45
46	BAYESIAN METHODS FOR GENETIC ASSOCIATION ANALYSIS WITH HETEROGENEOUS SUBGROUPS: FROM META-ANALYSES TO GENE-ENVIRONMENT INTERACTIONS. *Annals of Applied Statistics*, **2014** , 8, 176-203	2.1	39
45	Exploiting the GTEx resources to decipher the mechanisms at GWAS loci. *Genome Biology*, **2021**, 22, 49	18.3	38
44	Discovery and characterization of variance QTLs in human induced pluripotent stem cells. *PLoS Genetics*, **2019**, 15, e1008045	6	31
43	Estimating Time to the Common Ancestor for a Beneficial Allele. *Molecular Biology and Evolution*, **2018**, 35, 1003-1017	8.3	31
42	New evidence for hybrid zones of forest and savanna elephants in Central and West Africa. *Molecular Ecology*, **2015**, 24, 6134-47	5.7	30
41	A simple new approach to variable selection in regression, with application to genetic fine-mapping		30
40	Estimating recent migration and population-size surfaces. *PLoS Genetics*, **2019**, 15, e1007908	6	29
39	msCentipede: Modeling Heterogeneity across Genomic Sites and Replicates Improves Accuracy in the Inference of Transcription Factor Binding. *PLoS ONE*, **2015**, 10, e0138030	3.7	28
38	Separating measurement and expression models clarifies confusion in single-cell RNA sequencing analysis. *Nature Genetics*, **2021**, 53, 770-777	36.3	27
37	Characterizing and inferring quantitative cell cycle phase in single-cell RNA-seq data analysis. *Genome Research*, **2020**, 30, 611-621	9.7	26
36	Detailed modeling of positive selection improves detection of cancer driver genes. *Nature Communications*, **2019**, 10, 3399	17.4	24
35	Creating and sharing reproducible research code the workflowr way. *F1000Research*, **2019**, 8, 1749	3.6	23
34	A new sequence logo plot to highlight enrichment and depletion. *BMC Bioinformatics*, **2018**, 19, 473	3.6	23
33	Exploiting the GTEx resources to decipher the mechanisms at GWAS loci		21
32	Regional influences on community structure across the tropical-temperate divide. *Nature Communications*, **2019**, 10, 2646	17.4	20
31	Genetic, functional and molecular features of glucocorticoid receptor binding. *PLoS ONE*, **2013**, 8, e61654	3.7	20

30	Population-scale tissue transcriptomics maps long non-coding RNAs to complex disease. *Cell*, **2021**, 184, 2633-2648.e19	56.2	20
29	WAVELET-BASED GENETIC ASSOCIATION ANALYSIS OF FUNCTIONAL PHENOTYPES ARISING FROM HIGH-THROUGHPUT SEQUENCING ASSAYS. *Annals of Applied Statistics*, **2015**, 9, 655-686	2.1	18
28	A vast resource of allelic expression data spanning human tissues. *Genome Biology*, **2020**, 21, 234	18.3	18
27	Exon-specific QTLs skew the inferred distribution of expression QTLs detected using gene expression array data. *PLoS ONE*, **2012**, 7, e30629	3.7	17
26	Accurate genomic prediction of Coffea canephora in multiple environments using whole-genome statistical models. *Heredity*, **2019**, 122, 261-275	3.6	15
25	Silencing of transposable elements may not be a major driver of regulatory evolution in primate iPSCs. *ELife*, **2018**, 7,	8.9	15
24	Separating measurement and expression models clarifies confusion in single-cell RNA sequencing analysis		11
23	False Discovery Rates: A New Deal		10
22	Dynamic effects of genetic variation on gene expression revealed following hypoxic stress in cardiomyocytes. *ELife*, **2021**, 10,	8.9	8
21	Empirical Bayes shrinkage and false discovery rate estimation, allowing for unwanted variation. *Biostatistics*, **2020**, 21, 15-32	3.7	5
20	Bayesian multivariate reanalysis of large genetic studies identifies many new associations. *PLoS Genetics*, **2019**, 15, e1008431	6	5
19	A large-scale genome-wide enrichment analysis identifies new trait-associated genes, pathways and tissues across 31 human phenotypes*		5
18	Mangravite et al. reply. *Nature*, **2014**, 513, E3	50.4	4
17	Flexible statistical methods for estimating and testing effects in genomic studies with multiple conditions		4
16	Bayesian large-scale multiple regression with summary statistics from genome-wide association studies		4
15	A Fast Algorithm for Maximum Likelihood Estimation of Mixture Proportions Using Sequential Quadratic Programming. *Journal of Computational and Graphical Statistics*, **2020**, 29, 261-273	1.4	4
14	Inference and visualization of DNA damage patterns using a grade of membership model. *Bioinformatics*, **2019**, 35, 1292-1298	7.2	4
13	Variance adaptive shrinkage (vash): flexible empirical Bayes estimation of variances. *Bioinformatics*, **2016**, 32, 3428-3434	7.2	3

List of Publications

12	Characterizing and inferring quantitative cell cycle phase in single-cell RNA-seq data analysis		3
11	Dispersal syndromes drive the formation of biogeographical regions, illustrated by the case of Wallace's Line. *Global Ecology and Biogeography*, **2021**, 30, 685-696	6.1	3
10	Visualizing the Structure of RNA-seq Expression Data using Grade of Membership Models		2
9	Fine-mapping from summary data with the Sum of Single Effects model		1
8	Dynamic effects of genetic variation on gene expression revealed following hypoxic stress in cardiomyocytes		1
7	Genotyping Polyploids from Messy Sequencing Data		1
6	Bayesian multivariate reanalysis of large genetic studies identifies many new associations **2019**, 15, e1008431		
5	Bayesian multivariate reanalysis of large genetic studies identifies many new associations **2019**, 15, e1008431		
4	Bayesian multivariate reanalysis of large genetic studies identifies many new associations **2019**, 15, e1008431		
3	Bayesian multivariate reanalysis of large genetic studies identifies many new associations **2019**, 15, e1008431		
2	Bayesian multivariate reanalysis of large genetic studies identifies many new associations **2019**, 15, e1008431		
1	Bayesian multivariate reanalysis of large genetic studies identifies many new associations **2019**, 15, e1008431		