# Manolis Kellis

## List of Publications by Year in Descending Order



| | | | |
|---|---|---|---|
| 284 papers | 93,859 citations | 107 h-index | 306 g-index |
| 334 ext. papers | 118,058 ext. citations | 21.3 avg, IF | 8.81 L-index |

| # | Paper | IF | Citations |
|---|---|---|---|
| 284 | Single-cell dissection of the human brain vasculature.. *Nature*, **2022**, | 50.4 | 8 |
| 283 | Immune genes outside immune cells for multiple sclerosis.. *Neuron*, **2022**, 110, 1090-1092 | 13.9 | |
| 282 | MEF2 is a key regulator of cognitive potential and confers resilience to neurodegeneration. *Science Translational Medicine*, **2021**, 13, eabd7695 | 17.5 | 2 |
| 281 | Unannotated proteins expand the MHC-I-restricted immunopeptidome in cancer. *Nature Biotechnology*, **2021**, | 44.5 | 13 |
| 280 | disrupts intracellular lipid homeostasis in human iPSC-derived glia. *Science Translational Medicine*, **2021**, 13, | 17.5 | 35 |
| 279 | SARS-CoV-2 gene content and COVID-19 mutation impact by comparing 44 Sarbecovirus genomes. *Nature Communications*, **2021**, 12, 2642 | 17.4 | 52 |
| 278 | NEBULA is a fast negative binomial mixed model for differential or co-expression analysis of large-scale multi-subject single-cell data. *Communications Biology*, **2021**, 4, 629 | 6.7 | 6 |
| 277 | Evolution of delayed resistance to immunotherapy in a melanoma responder. *Nature Medicine*, **2021**, 27, 985-992 | 50.5 | 11 |
| 276 | Conflicting and ambiguous names of overlapping ORFs in the SARS-CoV-2 genome: A homology-based resolution. *Virology*, **2021**, 558, 145-151 | 3.6 | 15 |
| 275 | High-throughput 5' UTR engineering for enhanced protein production in non-viral gene therapies. *Nature Communications*, **2021**, 12, 4138 | 17.4 | 8 |
| 274 | Genetic drivers of mA methylation in human brain, lung, heart and muscle. *Nature Genetics*, **2021**, 53, 1156-1165 | 36.3 | 9 |
| 273 | GENCODE 2021. *Nucleic Acids Research*, **2021**, 49, D916-D923 | 20.1 | 82 |
| 272 | Genus-Wide Characterization of Bumblebee Genomes Provides Insights into Their Evolution and Variation in Ecological and Behavioral Traits. *Molecular Biology and Evolution*, **2021**, 38, 486-501 | 8.3 | 12 |
| 271 | Exome-wide age-of-onset analysis reveals exonic variants in ERN1 and SPPL2C associated with Alzheimer's disease. *Translational Psychiatry*, **2021**, 11, 146 | 8.6 | 1 |
| 270 | Regulatory genomic circuitry of human disease loci by integrative epigenomics. *Nature*, **2021**, 590, 300-307 | 50.4 | 43 |
| 269 | CoCoA-diff: counterfactual inference for single-cell gene expression analysis. *Genome Biology*, **2021**, 22, 228 | 18.3 | 0 |
| 268 | Distinct metabolic programs established in the thymus control effector functions of γδ T cell subsets in tumor microenvironments. *Nature Immunology*, **2021**, 22, 179-192 | 19.1 | 26 |

| | | | |
|---|---|---|---|
| 267 | Evolution of enhanced innate immune evasion by SARS-CoV-2.. *Nature*, **2021**, | 50.4 | 30 |
| 266 | Genomic RNA Elements Drive Phase Separation of the SARS-CoV-2 Nucleocapsid. *Molecular Cell*, **2020**, 80, 1078-1091.e6 | 17.6 | 98 |
| 265 | Inferring multimodal latent topics from electronic health records. *Nature Communications*, **2020**, 11, 2536 | 17.4 | 13 |
| 264 | Reconstruction of the human blood-brain barrier in vitro reveals a pathogenic mechanism of APOE4 in pericytes. *Nature Medicine*, **2020**, 26, 952-963 | 50.5 | 83 |
| 263 | Few SINEs of life: Alu elements have little evidence for biological relevance despite elevated translation. *NAR Genomics and Bioinformatics*, **2020**, 2, lqz023 | 3.7 | 6 |
| 262 | Evidence for a novel overlapping coding sequence in POLG initiated at a CUG start codon. *BMC Genetics*, **2020**, 21, 25 | 2.6 | 18 |
| 261 | Molecular Transducers of Physical Activity Consortium (MoTrPAC): Mapping the Dynamic Responses to Exercise. *Cell*, **2020**, 181, 1464-1474 | 56.2 | 51 |
| 260 | Genome-wide In Vivo CNS Screening Identifies Genes that Modify CNS Neuronal Survival and mHTT Toxicity. *Neuron*, **2020**, 106, 76-89.e8 | 13.9 | 27 |
| 259 | Analyses of non-coding somatic drivers in 2,658´cancer whole genomes. *Nature*, **2020**, 578, 102-111 | 50.4 | 220 |
| 258 | Pan-cancer analysis of whole genomes. *Nature*, **2020**, 578, 82-93 | 50.4 | 840 |
| 257 | Interleukin-6 deficiency exacerbates Huntington's disease model phenotypes. *Molecular Neurodegeneration*, **2020**, 15, 29 | 19 | 5 |
| 256 | SARS-CoV-2 gene content and COVID-19 mutation impact by comparing 44 Sarbecovirus genomes **2020**, | | 3 |
| 255 | SARS-CoV-2 gene content and COVID-19 mutation impact by comparing 44 Sarbecovirus genomes **2020**, | | 14 |
| 254 | Improved haplotype inference by exploiting long-range linking and allelic imbalance in RNA-seq datasets. *Nature Communications*, **2020**, 11, 4662 | 17.4 | 8 |
| 253 | Evidence for secondary-variant genetic burden and non-random distribution across biological modules in a recessive ciliopathy. *Nature Genetics*, **2020**, 52, 1145-1150 | 36.3 | 12 |
| 252 | Mapping the epigenomic and transcriptomic interplay during memory formation and recall in the hippocampal engram ensemble. *Nature Neuroscience*, **2020**, 23, 1606-1617 | 25.5 | 29 |
| 251 | Cell Type-Specific Transcriptomics Reveals that Mutant Huntingtin Leads to Mitochondrial RNA Release and Neuronal Innate Immune Activation. *Neuron*, **2020**, 107, 891-908.e8 | 13.9 | 46 |
| 250 | Conserved Epigenetic Regulatory Logic Infers Genes Governing Cell Identity. *Cell Systems*, **2020**, 11, 625-639.e13 | 10.5 | 13 |

| | | | |
|---|---|---|---|
| 249 | Plasma-derived extracellular vesicle analysis and deconvolution enable prediction and tracking of melanoma checkpoint blockade outcome. *Science Advances*, **2020**, 6, | 14.3 | 14 |
| 248 | Perspectives on ENCODE. *Nature*, **2020**, 583, 693-698 | 50.4 | 61 |
| 247 | Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature*, **2020**, 583, 699-710 | 50.4 | 360 |
| 246 | Translation Initiation Site Profiling Reveals Widespread Synthesis of Non-AUG-Initiated Protein Isoforms in Yeast. *Cell Systems*, **2020**, 11, 145-160.e5 | 10.6 | 16 |
| 245 | Analysis of Genetically Regulated Gene Expression Identifies a Prefrontal PTSD Gene, SNRNP35, Specific to Military Cohorts. *Cell Reports*, **2020**, 31, 107716 | 10.6 | 21 |
| 244 | A multiresolution framework to characterize single-cell state landscapes. *Nature Communications*, **2020**, 11, 5399 | 17.4 | 11 |
| 243 | A Quantitative Proteome Map of the Human Body. *Cell*, **2020**, 183, 269-283.e19 | 56.2 | 73 |
| 242 | A vast resource of allelic expression data spanning human tissues. *Genome Biology*, **2020**, 21, 234 | 18.3 | 18 |
| 241 | Spectral Alignment of Graphs. *IEEE Transactions on Network Science and Engineering*, **2020**, 7, 1182-1197 | 4.9 | 10 |
| 240 | Elucidation of Codon Usage Signatures across the Domains of Life. *Molecular Biology and Evolution*, **2019**, 36, 2328-2339 | 8.3 | 23 |
| 239 | Single-cell transcriptomic analysis of Alzheimer's disease. *Nature*, **2019**, 570, 332-337 | 50.4 | 682 |
| 238 | Challenges in IBD Research: Environmental Triggers. *Inflammatory Bowel Diseases*, **2019**, 25, S13-S23 | 4.5 | 35 |
| 237 | A gene expression atlas of embryonic neurogenesis in reveals complex spatiotemporal regulation of lncRNAs. *Development (Cambridge)*, **2019**, 146, | 6.6 | 18 |
| 236 | Network Infusion to Infer Information Sources in Networks. *IEEE Transactions on Network Science and Engineering*, **2019**, 6, 402-417 | 4.9 | 7 |
| 235 | ncdDetect2: improved models of the site-specific mutation rate in cancer and driver detection with robust significance evaluation. *Bioinformatics*, **2019**, 35, 189-199 | 7.2 | 3 |
| 234 | Joint profiling of DNA methylation and chromatin architecture in single cells. *Nature Methods*, **2019**, 16, 991-993 | 21.6 | 75 |
| 233 | Integrative construction of regulatory region networks in 127 human reference epigenomes by matrix factorization. *Nucleic Acids Research*, **2019**, 47, 7235-7246 | 20.1 | 2 |
| 232 | A high-throughput screening and computation platform for identifying synthetic promoters with enhanced cell-state specificity (SPECS). *Nature Communications*, **2019**, 10, 2880 | 17.4 | 18 |

| | | | |
|---|---|---|---|
| 231 | Single-cell transcriptomic atlas of the human retina identifies cell types associated with age-related macular degeneration. *Nature Communications*, **2019**, 10, 4902 | 17.4 | 100 |
| 230 | Rate of brain aging and are synergistic risk factors for Alzheimer's disease. *Life Science Alliance*, **2019**, 2, | 5.8 | 5 |
| 229 | Human Primordial Germ Cells Are Specified from Lineage-Primed Progenitors. *Cell Reports*, **2019**, 29, 4568-4582.e5 | 10.6 | 44 |
| 228 | Reconstruction of Cell-type-Specific Interactomes at Single-Cell Resolution. *Cell Systems*, **2019**, 9, 559-568.e4 | 10.1 | 24 |
| 227 | Discovery of high-confidence human protein-coding genes and exons by whole-genome PhyloCSF helps elucidate 118 GWAS loci. *Genome Research*, **2019**, 29, 2073-2087 | 9.7 | 26 |
| 226 | An AR-ERG transcriptional signature defined by long-range chromatin interactomes in prostate cancer cells. *Genome Research*, **2019**, 29, 223-235 | 9.7 | 24 |
| 225 | Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nature Genetics*, **2019**, 51, 237-244 | 36.3 | 516 |
| 224 | GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Research*, **2019**, 47, D766-D773 | 20.1 | 1140 |
| 223 | Pan-cancer screen for mutations in non-coding elements with conservation and cancer specificity reveals correlations with expression and survival. *Npj Genomic Medicine*, **2018**, 3, 1 | 6.2 | 38 |
| 222 | Stop codon readthrough generates a C-terminally extended variant of the human vitamin D receptor with reduced calcitriol response. *Journal of Biological Chemistry*, **2018**, 293, 4434-4444 | 5.4 | 30 |
| 221 | N-methyladenosine RNA modification regulates embryonic neural stem cell self-renewal through histone modifications. *Nature Neuroscience*, **2018**, 21, 195-206 | 25.5 | 185 |
| 220 | RANGER-DTL 2.0: rigorous reconstruction of gene-family evolution by duplication, transfer and loss. *Bioinformatics*, **2018**, 34, 3214-3216 | 7.2 | 27 |
| 219 | Analyses of mRNA structure dynamics identify embryonic gene regulatory programs. *Nature Structural and Molecular Biology*, **2018**, 25, 677-686 | 17.6 | 61 |
| 218 | Loose ends: almost one in five human genes still have unresolved coding status. *Nucleic Acids Research*, **2018**, 46, 7070-7084 | 20.1 | 36 |
| 217 | Deep coverage whole genome sequences and plasma lipoprotein(a) in individuals of European and African ancestries. *Nature Communications*, **2018**, 9, 2606 | 17.4 | 53 |
| 216 | Chromatin Accessibility Impacts Transcriptional Reprogramming in Oocytes. *Cell Reports*, **2018**, 24, 304-316 | 10.6 | 29 |
| 215 | Allele-specific epigenome maps reveal sequence-dependent stochastic switching at regulatory loci. *Science*, **2018**, 361, | 33.3 | 48 |
| 214 | Deep-coverage whole genome sequences and blood lipids among 16,324 individuals. *Nature Communications*, **2018**, 9, 3391 | 17.4 | 90 |

| | | | |
|---|---|---|---|
| 213 | Abstract 4282: Deconvolution of plasma-derived exosomes for tracking and prediction of immunotherapy across multiple tissues **2018**, | | 4 |
| 212 | Phylogenetic analysis of longitudinal melanoma samples to reveal convergent evolution and markers of immunotherapy resistance.. *Journal of Clinical Oncology*, **2018**, 36, 9581-9581 | 2.2 | |
| 211 | Target site specificity and in vivo complexity of the mammalian arginylome. *Scientific Reports*, **2018**, 8, 16177 | 4.9 | 16 |
| 210 | High-resolution genome-wide functional dissection of transcriptional regulatory regions and nucleotides in human. *Nature Communications*, **2018**, 9, 5380 | 17.4 | 52 |
| 209 | Loss of LDAH associated with prostate cancer and hearing loss. *Human Molecular Genetics*, **2018**, 27, 4194-4203 | 5.6 | 9 |
| 208 | Using an atlas of gene regulation across 44 human tissues to inform complex disease- and trait-associated variation. *Nature Genetics*, **2018**, 50, 956-967 | 36.3 | 239 |
| 207 | Predicting gene expression in massively parallel reporter assays: A comparative study. *Human Mutation*, **2017**, 38, 1240-1250 | 4.7 | 23 |
| 206 | Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. *Nature Genetics*, **2017**, 49, 834-841 | 36.3 | 257 |
| 205 | Network Maximal Correlation. *IEEE Transactions on Network Science and Engineering*, **2017**, 4, 229-247 | 4.9 | 3 |
| 204 | Multi-scale chromatin state annotation using a hierarchical hidden Markov model. *Nature Communications*, **2017**, 8, 15011 | 17.4 | 24 |
| 203 | Dynamic landscape and regulation of RNA editing in mammals. *Nature*, **2017**, 550, 249-254 | 50.4 | 286 |
| 202 | Landscape of X chromosome inactivation across human tissues. *Nature*, **2017**, 550, 244-248 | 50.4 | 417 |
| 201 | Genetic effects on gene expression across human tissues. *Nature*, **2017**, 550, 204-213 | 50.4 | 2086 |
| 200 | Evidence of reduced recombination rate in human regulatory domains. *Genome Biology*, **2017**, 18, 193 | 18.3 | 23 |
| 199 | Chromatin-state discovery and genome annotation with ChromHMM. *Nature Protocols*, **2017**, 12, 2478-2492 | 9.8 | 257 |
| 198 | Joint Bayesian inference of risk variants and tissue-specific epigenomic enrichments across multiple complex human diseases. *Nucleic Acids Research*, **2016**, 44, e144 | 20.1 | 58 |
| 197 | 52 Genetic Loci Influencing Myocardial Mass. *Journal of the American College of Cardiology*, **2016**, 68, 1435-1448 | 15.1 | 76 |
| 196 | Evolutionary Dynamics of Abundant Stop Codon Readthrough. *Molecular Biology and Evolution*, **2016**, 33, 3108-3132 | 8.3 | 32 |

| | | | |
|---|---|---|---|
| 195 | Soft X-Ray Tomography Reveals Gradual Chromatin Compaction and Reorganization during Neurogenesis In Vivo. *Cell Reports*, **2016**, 17, 2125-2136 | 10.6 | 59 |
| 194 | Survey of variation in human transcription factors reveals prevalent DNA binding changes. *Science*, **2016**, 351, 1450-1454 | 33.3 | 78 |
| 193 | Improved Identification and Analysis of Small Open Reading Frame Encoded Polypeptides. *Analytical Chemistry*, **2016**, 88, 3967-75 | 7.8 | 73 |
| 192 | MicroRNA and gene expression changes in unruptured human cerebral aneurysms. *Journal of Neurosurgery*, **2016**, 125, 1390-1399 | 3.2 | 29 |
| 191 | Tissue-specific regulatory circuits reveal variable modular perturbations across complex diseases. *Nature Methods*, **2016**, 13, 366-70 | 21.6 | 201 |
| 190 | Discovery of Genetic Variation on Chromosome 5q22 Associated with Mortality in Heart Failure. *PLoS Genetics*, **2016**, 12, e1006034 | 6 | 26 |
| 189 | Discovery and validation of sub-threshold genome-wide association study loci using epigenomic signatures. *ELife*, **2016**, 5, | 8.9 | 64 |
| 188 | SwiSpot: modeling riboswitches by spotting out switching sequences. *Bioinformatics*, **2016**, 32, 3252-3259 | 7.2 | 7 |
| 187 | HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Research*, **2016**, 44, D877-81 | 20.1 | 522 |
| 186 | Genome-scale high-resolution mapping of activating and repressive nucleotides in regulatory regions. *Nature Biotechnology*, **2016**, 34, 1180-1190 | 44.5 | 85 |
| 185 | Deep learning for regulatory genomics. *Nature Biotechnology*, **2015**, 33, 825-6 | 44.5 | 103 |
| 184 | Sharing and Specificity of Co-expression Networks across 35 Human Tissues. *PLoS Computational Biology*, **2015**, 11, e1004220 | 5 | 104 |
| 183 | Context influences on TALE-DNA binding revealed by quantitative profiling. *Nature Communications*, **2015**, 6, 7440 | 17.4 | 22 |
| 182 | FRESCo: finding regions of excess synonymous constraint in diverse viruses. *Genome Biology*, **2015**, 16, 38 | 18.3 | 30 |
| 181 | Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*, **2015**, 348, 648-60 | 33.3 | 3242 |
| 180 | Human genomics. Effect of predicted protein-truncating genetic variants on the human transcriptome. *Science*, **2015**, 348, 666-9 | 33.3 | 170 |
| 179 | A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreservation and Biobanking*, **2015**, 13, 311-9 | 2.1 | 432 |
| 178 | Systematic chromatin state comparison of epigenomes associated with diverse properties including sex and tissue type. *Nature Communications*, **2015**, 6, 7973 | 17.4 | 42 |

| | | | |
|---|---|---|---|
| 177 | FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *New England Journal of Medicine*, **2015**, 373, 895-907 | 59.2 | 821 |
| 176 | PRC2 Is Required to Maintain Expression of the Maternal Gtl2-Rian-Mirg Locus by Preventing De Novo DNA Methylation in Mouse Embryonic Stem Cells. *Cell Reports*, **2015**, 12, 1456-70 | 10.6 | 46 |
| 175 | Mosquito genomics. Highly evolvable malaria vectors: the genomes of 16 Anopheles mosquitoes. *Science*, **2015**, 347, 1258522 | 33.3 | 372 |
| 174 | Alzheimer's loci: epigenetic associations and interaction with genetic factors. *Annals of Clinical and Translational Neurology*, **2015**, 2, 636-47 | 5.3 | 42 |
| 173 | Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. *Nature Genetics*, **2015**, 47, 381-6 | 36.3 | 414 |
| 172 | Activity-Induced DNA Breaks Govern the Expression of Neuronal Early-Response Genes. *Cell*, **2015**, 161, 1592-605 | 56.2 | 386 |
| 171 | Improved gene tree error correction in the presence of horizontal gene transfer. *Bioinformatics*, **2015**, 31, 1211-8 | 7.2 | 37 |
| 170 | BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. *Molecular Cell*, **2015**, 57, 636-647 | 17.6 | 245 |
| 169 | Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues. *Nature Biotechnology*, **2015**, 33, 364-76 | 44.5 | 252 |
| 168 | Intermediate DNA methylation is a conserved signature of genome regulation. *Nature Communications*, **2015**, 6, 6363 | 17.4 | 67 |
| 167 | Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. *Nature*, **2015**, 518, 365-9 | 50.4 | 390 |
| 166 | Integrative analysis of 111 reference human epigenomes. *Nature*, **2015**, 518, 317-30 | 50.4 | 3849 |
| 165 | The Discovery of Human sORF-Encoded Polypeptides (SEPs) in Cell Lines and Tissue. *FASEB Journal*, **2015**, 29, 567.21 | 0.9 | |
| 164 | Discovery of human sORF-encoded polypeptides (SEPs) in cell lines and tissue. *Journal of Proteome Research*, **2014**, 13, 1757-65 | 5.6 | 111 |
| 163 | Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science*, **2014**, 343, 1246980 | 33.3 | 309 |
| 162 | Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature* , **2014**, 505, 701-5 | 50.4 | 560 |
| 161 | Evolutionary dynamics and tissue specificity of human long noncoding RNAs in six mammals. *Genome Research*, **2014**, 24, 616-28 | 9.7 | 250 |
| 160 | RECOMB/ISCB systems biology, regulatory genomics, and DREAM 2013 special issue. *Journal of Computational Biology*, **2014**, 21, 371-2 | 1.7 | 1 |

| | | | |
|---|---|---|---|
| 159 | A comparative encyclopedia of DNA elements in the mouse genome. *Nature*, **2014**, 515, 355-64 | 50.4 | 1026 |
| 158 | Comparative analysis of regulatory information and circuits across distant species. *Nature*, **2014**, 512, 453-6 | 50.4 | 135 |
| 157 | Comparative analysis of metazoan chromatin organization. *Nature*, **2014**, 512, 449-52 | 50.4 | 265 |
| 156 | Reply to Brunet and Doolittle: Both selected effect and causal role elements can influence human biology and disease. *Proceedings of the National Academy of Sciences of the United States of America* , **2014**, 111, E3366 | 11.5 | 22 |
| 155 | Comparative validation of the D. melanogaster modENCODE transcriptome annotation. *Genome Research*, **2014**, 24, 1209-23 | 9.7 | 95 |
| 154 | CHD8 regulates neurodevelopmental pathways associated with autism spectrum disorder in neural progenitors. *Proceedings of the National Academy of Sciences of the United States of America*, **2014**, 111, E4468-77 | 11.5 | 210 |
| 153 | Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nature Neuroscience*, **2014**, 17, 1156-63 | 25.5 | 579 |
| 152 | Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization. *Nature Genetics*, **2014**, 46, 826-36 | 36.3 | 199 |
| 151 | Evidence of efficient stop codon readthrough in four mammalian genes. *Nucleic Acids Research*, **2014**, 42, 8928-38 | 20.1 | 130 |
| 150 | Defining functional DNA elements in the human genome. *Proceedings of the National Academy of Sciences of the United States of America*, **2014**, 111, 6131-8 | 11.5 | 490 |
| 149 | Pareto-optimal phylogenetic tree reconciliation. *Bioinformatics*, **2014**, 30, i87-95 | 7.2 | 41 |
| 148 | Most parsimonious reconciliation in the presence of gene duplication, loss, and deep coalescence using labeled coalescent trees. *Genome Research*, **2014**, 24, 475-86 | 9.7 | 46 |
| 147 | Core and region-enriched networks of behaviorally regulated genes and the singing genome. *Science*, **2014**, 346, 1256780 | 33.3 | 81 |
| 146 | Diverse patterns of genomic targeting by transcriptional regulators in Drosophila melanogaster. *Genome Research*, **2014**, 24, 1224-35 | 9.7 | 27 |
| 145 | Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. *Nucleic Acids Research*, **2014**, 42, 2976-87 | 20.1 | 302 |
| 144 | The NF-**B** genomic landscape in lymphoblastoid B cells. *Cell Reports*, **2014**, 8, 1595-606 | 10.6 | 112 |
| 143 | Distinct and predictive histone lysine acetylation patterns at promoters, enhancers, and gene bodies. *G3: Genes, Genomes, Genetics*, **2014**, 4, 2051-63 | 3.2 | 27 |
| 142 | Energy-based RNA consensus secondary structure prediction in multiple sequence alignments. *Methods in Molecular Biology*, **2014**, 1097, 125-41 | 1.4 | 2 |

| | | | |
|---|---|---|---|
| 141 | Genomic evidence for ameiotic evolution in the bdelloid rotifer Adineta vaga. *Nature*, **2013**, 500, 453-7 | 50.4 | 274 |
| 140 | Network deconvolution as a general method to distinguish direct dependencies in networks. *Nature Biotechnology*, **2013**, 31, 726-33 | 44.5 | 174 |
| 139 | Extensive variation in chromatin states across humans. *Science*, **2013**, 342, 750-2 | 33.3 | 276 |
| 138 | Spatial expression of transcription factors in Drosophila embryonic organ development. *Genome Biology*, **2013**, 14, R140 | 18.3 | 87 |
| 137 | Reconciliation revisited: handling multiple optima when reconciling with duplication, transfer, and loss. *Journal of Computational Biology*, **2013**, 20, 738-54 | 1.7 | 45 |
| 136 | Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Research*, **2013**, 41, 827-41 | 10.1 | 383 |
| 135 | Long noncoding RNAs regulate adipogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, **2013**, 110, 3387-92 | 11.5 | 315 |
| 134 | Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Research*, **2013**, 23, 270-80 | 9.7 | 301 |
| 133 | The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Developmental Cell*, **2013**, 24, 206-14 | 10.2 | 718 |
| 132 | Integrating and mining the chromatin landscape of cell-type specificity using self-organizing maps. *Genome Research*, **2013**, 23, 2136-48 | 9.7 | 39 |
| 131 | Systematic dissection of regulatory motifs in 2000 predicted human enhancers using a massively parallel reporter assay. *Genome Research*, **2013**, 23, 800-11 | 9.7 | 191 |
| 130 | Interplay between chromatin state, regulator binding, and regulatory motifs in six human cell types. *Genome Research*, **2013**, 23, 1142-54 | 9.7 | 72 |
| 129 | Arboretum: reconstruction and analysis of the evolutionary history of condition-specific transcriptional modules. *Genome Research*, **2013**, 23, 1039-50 | 9.7 | 40 |
| 128 | TreeFix: statistically informed gene tree error correction using species trees. *Systematic Biology*, **2013**, 62, 110-20 | 8.4 | 80 |
| 127 | The Genotype-Tissue Expression (GTEx) project. *Nature Genetics*, **2013**, 45, 580-5 | 36.3 | 4179 |
| 126 | Response to comment on "Evidence of abundant purifying selection in humans for recently acquired regulatory functions". *Science*, **2013**, 340, 682 | 33.3 | 12 |
| 125 | Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *ELife* , **2013**, 2, e01749 | 8.9 | 516 |
| 124 | RFECS: a random-forest based algorithm for enhancer identification from chromatin state. *PLoS Computational Biology*, **2013**, 9, e1002968 | 5 | 160 |

| | | | |
|---|---|---|---|
| 123 | Heterologous stop codon readthrough of metazoan readthrough candidates in yeast. *PLoS ONE*, **2013**, 8, e59450 | 3.7 | 6 |
| 122 | Evolutionary principles of modular gene regulation in yeasts. *ELife*, **2013**, 2, e00603 | 8.9 | 50 |
| 121 | Reconciliation Revisited: Handling Multiple Optima When Reconciling with Duplication, Transfer, and Loss. *Lecture Notes in Computer Science*, **2013**, 1-13 | 0.9 | 4 |
| 120 | Interpreting noncoding genetic variation in complex traits and human disease. *Nature Biotechnology*, **2012**, 30, 1095-106 | 44.5 | 347 |
| 119 | Analysis of variation at transcription factor binding sites in Drosophila and humans. *Genome Biology* , **2012**, 13, R49 | 18.3 | 71 |
| 118 | Linking DNA methyltransferases to epigenetic marks and nucleosome structure genome-wide in human tumor cells. *Cell Reports*, **2012**, 2, 1411-24 | 10.6 | 86 |
| 117 | An integrated encyclopedia of DNA elements in the human genome. *Nature*, **2012**, 489, 57-74 | 50.4 | 11449 |
| 116 | ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Research*, **2012**, 22, 1813-31 | 9.7 | 1211 |
| 115 | GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Research*, **2012**, 22, 1760-74 | 9.7 | 3142 |
| 114 | Wisdom of crowds for robust gene network inference. *Nature Methods*, **2012**, 9, 796-804 | 21.6 | 1097 |
| 113 | Evidence of abundant purifying selection in humans for recently acquired regulatory functions. *Science*, **2012**, 337, 1675-8 | 33.3 | 151 |
| 112 | Computational analysis of noncoding RNAs. *Wiley Interdisciplinary Reviews RNA*, **2012**, 3, 759-78 | 9.3 | 44 |
| 111 | High depth, whole-genome sequencing of cholera isolates from Haiti and the Dominican Republic. *BMC Genomics*, **2012**, 13, 468 | 4.5 | 7 |
| 110 | Disruption of a large intergenic noncoding RNA in subjects with neurodevelopmental disabilities. *American Journal of Human Genetics*, **2012**, 91, 1128-34 | 11 | 53 |
| 109 | Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nature Biotechnology*, **2012**, 30, 271-7 | 44.5 | 431 |
| 108 | ChromHMM: automating chromatin-state discovery and characterization. *Nature Methods*, **2012**, 9, 215-6 | 21.6 | 1413 |
| 107 | Predictive regulatory models in Drosophila melanogaster by integrative inference of transcriptional networks. *Genome Research*, **2012**, 22, 1334-49 | 9.7 | 79 |
| 106 | Common variants at 9p21 and 8q22 are associated with increased susceptibility to optic nerve degeneration in glaucoma. *PLoS Genetics*, **2012**, 8, e1002654 | 6 | 227 |

| | | | |
|---|---|---|---|
| 105 | Efficient algorithms for the reconciliation problem with gene duplication, horizontal transfer and loss. *Bioinformatics*, **2012**, 28, i283-91 | 7.2 | 138 |
| 104 | Unified modeling of gene duplication, loss, and coalescence using a locus tree. *Genome Research*, **2012**, 22, 755-65 | 9.7 | 117 |
| 103 | Preface: RECOMB Systems Biology, Regulatory Genomics, and DREAM 2011 special issue. *Journal of Computational Biology*, **2012**, 19, 101 | 1.7 | 2 |
| 102 | RNA folding with soft constraints: reconciliation of probing data and thermodynamic secondary structure prediction. *Nucleic Acids Research*, **2012**, 40, 4261-72 | 20.1 | 84 |
| 101 | Evolution at the subgene level: domain rearrangements in the Drosophila phylogeny. *Molecular Biology and Evolution*, **2012**, 29, 689-705 | 8.3 | 34 |
| 100 | HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Research*, **2012**, 40, D930-4 | 20.1 | 1630 |
| 99 | Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nature Genetics*, **2011**, 43, 621-9 | 36.3 | 902 |
| 98 | Comparative functional genomics of the fission yeasts. *Science*, **2011**, 332, 930-6 | 33.3 | 364 |
| 97 | An epigenetic signature for monoallelic olfactory receptor expression. *Cell*, **2011**, 145, 555-70 | 56.2 | 218 |
| 96 | Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells. *Cell*, **2011**, 147, 1628-39 | 56.2 | 265 |
| 95 | A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biology*, **2011**, 9, e1001046 | 9.7 | 1060 |
| 94 | Error and error mitigation in low-coverage genome assemblies. *PLoS ONE*, **2011**, 6, e17034 | 3.7 | 29 |
| 93 | Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. *Nature*, **2011**, 471, 480-5 | 50.4 | 641 |
| 92 | Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*, **2011**, 473, 43-9 | 50.4 | 2153 |
| 91 | A cis-regulatory map of the Drosophila genome. *Nature*, **2011**, 471, 527-31 | 50.4 | 390 |
| 90 | A high-resolution map of human evolutionary constraint using 29 mammals. *Nature*, **2011**, 478, 476-82 | 50.4 | 802 |
| 89 | Phylogenomic Approach to the Evolutionary Dynamics of Gene Duplication in Birds **2011**, 253-268 | | 2 |
| 88 | PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics*, **2011**, 27, i275-82 | 7.2 | 637 |

| | | | |
|---|---|---|---|
| 87 | Three periods of regulatory innovation during vertebrate evolution. *Science*, **2011**, 333, 1019-24 | 33.3 | 92 |
| 86 | Locating protein-coding sequences under selection for additional, overlapping functions in 29 mammalian genomes. *Genome Research*, **2011**, 21, 1916-28 | 9.7 | 71 |
| 85 | New families of human regulatory RNA structures identified by comparative analysis of vertebrate genomes. *Genome Research*, **2011**, 21, 1929-43 | 9.7 | 76 |
| 84 | SubMAP: aligning metabolic pathways with subnetwork mappings. *Journal of Computational Biology*, **2011**, 18, 219-35 | 1.7 | 51 |
| 83 | A Bayesian approach for fast and accurate gene tree reconstruction. *Molecular Biology and Evolution*, **2011**, 28, 273-90 | 8.3 | 95 |
| 82 | Dynamics of the epigenetic landscape during erythroid differentiation after GATA1 restoration. *Genome Research*, **2011**, 21, 1659-71 | 9.7 | 100 |
| 81 | Evidence of abundant stop codon readthrough in Drosophila and other metazoa. *Genome Research*, **2011**, 21, 2096-113 | 9.7 | 139 |
| 80 | Discovery and Characterization of Chromatin States for Systematic Annotation of the Human Genome. *Lecture Notes in Computer Science*, **2011**, 53-53 | 0.9 | 2 |
| 79 | Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nature Biotechnology*, **2010**, 28, 817-25 | 44.5 | 757 |
| 78 | The NIH Roadmap Epigenomics Mapping Consortium. *Nature Biotechnology*, **2010**, 28, 1045-8 | 44.5 | 1284 |
| 77 | PhyloCSF: a comparative genomics method to distinguish protein-coding and non-coding regions. *Nature Precedings*, **2010**, | | 1 |
| 76 | A comprehensive map of insulator elements for the Drosophila genome. *PLoS Genetics*, **2010**, 6, e1000814 | 64 | 255 |
| 75 | Motif Discovery in Physiological Datasets: A Methodology for Inferring Predictive Elements. *ACM Transactions on Knowledge Discovery From Data*, **2010**, 4, 2 | 4 | 17 |
| 74 | Identification of functional elements and regulatory circuits by Drosophila modENCODE. *Science*, **2010**, 330, 1787-97 | 33.3 | 892 |
| 73 | The Tasmanian devil transcriptome reveals Schwann cell origins of a clonally transmissible cancer. *Science*, **2010**, 327, 84-7 | 33.3 | 188 |
| 72 | Optimization of parameters for coverage of low molecular weight proteins. *Analytical and Bioanalytical Chemistry*, **2010**, 398, 2867-81 | 4.4 | 36 |
| 71 | The consensus coding sequence (CCDS) project: Identifying a common protein-coding gene set for the human and mouse genomes. *Genome Research*, **2009**, 19, 1316-23 | 9.7 | 415 |
| 70 | Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*, **2009**, 458, 223-7 | 50.4 | 3230 |

| | | | |
|---|---|---|---|
| 69 | Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*, **2009**, 459, 108-12 | 50.4 | 1832 |
| 68 | Evolution of pathogenicity and sexual reproduction in eight Candida genomes. *Nature*, **2009**, 459, 657-62 | 50.4 | 764 |
| 67 | Unlocking the secrets of the genome. *Nature*, **2009**, 459, 927-30 | 50.4 | 620 |
| 66 | Genome analysis of the platypus reveals unique signatures of evolution. *Nature*, **2008**, 453, 175-83 | 50.4 | 545 |
| 65 | An endogenous small interfering RNA pathway in Drosophila. *Nature*, **2008**, 453, 798-802 | 50.4 | 542 |
| 64 | A single Hox locus in Drosophila produces functional microRNAs from opposite DNA strands. *Genes and Development*, **2008**, 22, 8-13 | 12.6 | 188 |
| 63 | The evolutionary dynamics of the Saccharomyces cerevisiae protein interaction network after duplication. *Proceedings of the National Academy of Sciences of the United States of America*, **2008**, 105, 950-4 | 11.5 | 46 |
| 62 | Performance and scalability of discriminative metrics for comparative gene identification in 12 Drosophila genomes. *PLoS Computational Biology*, **2008**, 4, e1000067 | 5 | 41 |
| 61 | Conservation of small RNA pathways in platypus. *Genome Research*, **2008**, 18, 995-1004 | 9.7 | 34 |
| 60 | Network Motif Discovery Using Subgraph Enumeration and Symmetry-Breaking **2007**, 92-106 | | 124 |
| 59 | RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. *Nature Genetics*, **2007**, 39, 1512-6 | 36.3 | 589 |
| 58 | Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures. *Nature*, **2007**, 450, 219-32 | 50.4 | 506 |
| 57 | Evolution of genes and genomes on the Drosophila phylogeny. *Nature*, **2007**, 450, 203-18 | 50.4 | 1586 |
| 56 | Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. *Proceedings of the National Academy of Sciences of the United States of America*, **2007**, 104, 7145-50 | 11.5 | 242 |
| 55 | Systematic discovery and characterization of fly microRNAs using 12 Drosophila genomes. *Genome Research*, **2007**, 17, 1865-79 | 9.7 | 158 |
| 54 | Revisiting the protein-coding gene catalog of Drosophila melanogaster using 12 fly genomes. *Genome Research*, **2007**, 17, 1823-36 | 9.7 | 121 |
| 53 | Distinguishing protein-coding and noncoding genes in the human genome. *Proceedings of the National Academy of Sciences of the United States of America*, **2007**, 104, 19428-33 | 11.5 | 440 |
| 52 | Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the Drosophila embryo. *Genes and Development*, **2007**, 21, 385-90 | 12.6 | 232 |

| | | | |
|---|---|---|---|
| 51 | Reliable prediction of regulator targets using 12 Drosophila genomes. *Genome Research*, **2007**, 17, 1919-31 | 9.7 | 123 |
| 50 | Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell*, **2007**, 128, 1089-103 | 56.2 | 1802 |
| 49 | Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs. *Genome Research*, **2007**, 17, 1850-64 | 9.7 | 462 |
| 48 | Accurate gene-tree reconstruction by learning gene- and species-specific substitution rates across multiple complete genomes. *Genome Research*, **2007**, 17, 1932-42 | 9.7 | 61 |
| 47 | Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature*, **2005**, 434, 338-45 | 50.4 | 1551 |
| 46 | Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*, **2005**, 438, 803-19 | 50.4 | 1809 |
| 45 | Transcriptional regulatory code of a eukaryotic genome. *Nature*, **2004**, 431, 99-104 | 50.4 | 1732 |
| 44 | Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature*, **2004**, 431, 946-57 | 50.4 | 1543 |
| 43 | The changing face of genomics. *Genome Biology*, **2004**, 5, 324 | 18.3 | |
| 42 | Methods in comparative genomics: genome correspondence, gene identification and regulatory motif discovery. *Journal of Computational Biology*, **2004**, 11, 319-55 | 1.7 | 72 |
| 41 | Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae. *Nature*, **2004**, 428, 617-24 | 50.4 | 1158 |
| 40 | Position specific variation in the rate of evolution in transcription factor binding sites. *BMC Evolutionary Biology*, **2003**, 3, 19 | 3 | 114 |
| 39 | The genome sequence of the filamentous fungus Neurospora crassa. *Nature*, **2003**, 422, 859-68 | 50.4 | 1323 |
| 38 | Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature*, **2003**, 423, 241-54 | 50.4 | 1468 |
| 37 | Phylogenetically and spatially conserved word pairs associated with gene-expression changes in yeasts. *Genome Biology*, **2003**, 4, R43 | 18.3 | 40 |
| 36 | Deep-coverage whole genome sequences and blood lipids among 16,324 individuals | | 4 |
| 35 | Deep coverage whole genome sequences and plasma lipoprotein(a) in individuals of European and African ancestries | | 1 |
| 34 | HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants | | 1 |

| | | |
|---|---|---|
| 33 | Phylogenetic Identification and Functional Characterization of Orthologs and Paralogs across Human, Mouse, Fly, and Worm | 7 |
| 32 | Functional enrichments of disease variants across thousands of independent loci in eight diseases | 3 |
| 31 | RiVIERA-beta: Joint Bayesian inference of risk variants and tissue-specific epigenomic enrichments across multiple complex human diseases | 1 |
| 30 | A probabilistic framework to dissect functional cell-type-specific regulatory elements and risk loci underlying the genetics of complex traits | 2 |
| 29 | Distant regulatory effects of genetic variation in multiple human tissues | 6 |
| 28 | Local genetic effects on gene expression across 44 human tissues | 22 |
| 27 | Convergence of dispersed regulatory mutations predicts driver genes in prostate cancer | 4 |
| 26 | Multi-tissue polygenic models for transcriptome-wide association studies | 12 |
| 25 | Modeling prediction error improves power of transcriptome-wide association studies | 6 |
| 24 | Best practices for genome-wide RNA structure analysis: combination of mutational profiles and drop-off information | 7 |
| 23 | High-resolution genome-wide functional dissection of transcriptional regulatory regions in human | 9 |
| 22 | Thousands of novel unannotated proteins expand the MHC I immunopeptidome in cancer | 15 |
| 21 | Correlating predicted epigenetic marks with expression data to find interactions between SNPs and genes | 1 |
| 20 | High-Throughput 5′UTR Engineering for Enhanced Protein Production in Non-Viral Gene Therapies | 2 |
| 19 | Single-cell dissection of schizophrenia reveals neurodevelopmental-synaptic axis and transcriptional resilience | 5 |
| 18 | Mathematical analysis of Córdoba calcifediol trial suggests strong role for Vitamin D in reducing ICU admissions of hospitalized COVID-19 patients | 5 |
| 17 | A Bayesian approach to mediation analysis predicts 206 causal target genes in Alzheimer's disease | 7 |
| 16 | Discovery and characterization of coding and non-coding driver mutations in more than 2,500 whole cancer genomes | 12 |

# List of Publications

| | | |
|---|---|---|
| 15 | Nearly all new protein-coding predictions in the CHESS database are not protein-coding | 8 |
| 14 | Evidence for secondary-variant genetic burden and non-random distribution across biological modules in a recessive ciliopathy | 2 |
| 13 | Simultaneous profiling of DNA methylation and chromatin architecture in mixed populations and in single cells | 1 |
| 12 | Analysis of Genetically Regulated Gene Expression identifies a trauma type specific PTSD gene, SNRNP35 | 1 |
| 11 | Conserved epigenetic regulatory logic infers genes governing cell identity | 2 |
| 10 | A multiresolution framework to characterize single-cell state landscapes | 2 |
| 9 | Integrative analysis of 10,000 epigenomic maps across 800 samples for regulatory genomics and disease dissection | 4 |
| 8 | Evidence of a recombination rate valley in human regulatory domains | 1 |
| 7 | Genome-wide regulatory model from MPRA data predicts functional regions, eQTLs, and GWAS hits | 1 |
| 6 | Pan-cancer screen for mutations in non-coding elements with conservation and cancer specificity reveals correlations with expression and survival | 1 |
| 5 | Structural variant selection for high-altitude adaptation using single-molecule long-read sequencing | 1 |
| 4 | Single-cell dissection of the human cerebrovasculature in health and disease | 4 |
| 3 | Single-cell dissection of live human hearts in ischemic heart disease and heart failure reveals cell-type-specific driver genes and pathways | 2 |
| 2 | Single-cell profiling of the human primary motor cortex in ALS and FTLD | 1 |
| 1 | Single-cell deconvolution of 3,000 post-mortem brain samples for eQTL and GWAS dissection in mental disorders | 4 |