# Vineet Bafna

## List of Publications by Year in Descending Order



| | | | |
|---|---|---|---|
| 184 papers | 38,431 citations | 56 h-index | 196 g-index |
| 211 ext. papers | 46,241 ext. citations | 9.7 avg, IF | 6.49 L-index |

| # | Paper | IF | Citations |
|---|---|---|---|
| 184 | Optical genome mapping identifies rare structural variations as predisposition factors associated with severe COVID-19.. *IScience*, **2022**, 103760 | 6.1 | 1 |
| 183 | Mapping clustered mutations in cancer reveals APOBEC3 mutagenesis of ecDNA.. *Nature*, **2022**, | 50.4 | 5 |
| 182 | Uncertainty Quantification Using Subsampling for Assembly-Free Estimates of Genomic Distance and Phylogenetic Relationships. *Lecture Notes in Computer Science*, **2022**, 366-368 | 0.9 | |
| 181 | FastViFi: Fast and accurate detection of (Hybrid) Viral DNA and RNA.. *NAR Genomics and Bioinformatics*, **2022**, 4, lqac032 | 3.7 | |
| 180 | Estimating repeat spectra and genome length from low-coverage genome skims with RESPECT. *PLoS Computational Biology*, **2021**, 17, e1009449 | 5 | 4 |
| 179 | Extrachromosomal DNA: An Emerging Hallmark in Human Cancer. *Annual Review of Pathology: Mechanisms of Disease*, **2021**, | 34 | 5 |
| 178 | ecDNA hubs drive cooperative intermolecular oncogene expression. *Nature*, **2021**, | 50.4 | 18 |
| 177 | Variable number tandem repeats mediate the expression of proximal genes. *Nature Communications*, **2021**, 12, 2075 | 17.4 | 10 |
| 176 | FaNDOM: Fast nested distance-based seeding of optical maps. *Patterns*, **2021**, 2, 100248 | 5.1 | 4 |
| 175 | CONSULT: accurate contamination removal using locality-sensitive hashing. *NAR Genomics and Bioinformatics*, **2021**, 3, lqab071 | 3.7 | 2 |
| 174 | Extrachromosomal DNA (ecDNA) in cancer pathogenesis. *Current Opinion in Genetics and Development*, **2021**, 66, 78-82 | 4.9 | 9 |
| 173 | Multiple mechanisms drive genomic adaptation to extreme O levels in Drosophila melanogaster. *Nature Communications*, **2021**, 12, 997 | 17.4 | 3 |
| 172 | Extrachromosomal DNA in HPV-Mediated Oropharyngeal Cancer Drives Diverse Oncogene Transcription. *Clinical Cancer Research*, **2021**, | 12.9 | 3 |
| 171 | Plasticity of extrachromosomal and intrachromosomal BRAF amplifications in overcoming targeted therapy dosage challenges.. *Cancer Discovery*, **2021**, | 24.4 | 1 |
| 170 | The impact of contaminants on the accuracy of genome skimming and the effectiveness of exclusion read filters. *Molecular Ecology Resources*, **2020**, 20, 649 | 8.4 | 6 |
| 169 | Longitudinal assessment of tumor development using cancer avatars derived from genetically engineered pluripotent stem cells. *Nature Communications*, **2020**, 11, 550 | 17.4 | 23 |
| 168 | AmpliconReconstructor integrates NGS and optical mapping to resolve the complex structures of focal amplifications. *Nature Communications*, **2020**, 11, 4374 | 17.4 | 14 |

| | | | |
|---|---|---|---|
| 167 | Extrachromosomal DNA is associated with oncogene amplification and poor outcome across multiple cancers. *Nature Genetics*, **2020**, 52, 891-897 | 36.3 | 79 |
| 166 | Exploring the landscape of focal amplifications in cancer using AmpliconArchitect. *Nature Communications*, **2019**, 10, 392 | 17.4 | 59 |
| 165 | NAD metabolic dependency in cancer is shaped by gene amplification and enhancer remodelling. *Nature*, **2019**, 569, 570-575 | 50.4 | 81 |
| 164 | Proteogenomic Annotation of Chinese Hamsters Reveals Extensive Novel Translation Events and Endogenous Retroviral Elements. *Journal of Proteome Research*, **2019**, 18, 2433-2445 | 5.6 | 12 |
| 163 | Extrachromosomal oncogene amplification in tumour pathogenesis and evolution. *Nature Reviews Cancer*, **2019**, 19, 283-288 | 31.3 | 108 |
| 162 | Skmer: assembly-free and alignment-free sample identification using genome skims. *Genome Biology*, **2019**, 20, 34 | 18.3 | 37 |
| 161 | Computing the Statistical Significance of Overlap between Genome Annotations with iStat. *Cell Systems*, **2019**, 8, 523-529.e4 | 10.6 | 2 |
| 160 | EcSeg: Semantic Segmentation of Metaphase Images Containing Extrachromosomal DNA. *IScience*, **2019**, 21, 428-435 | 6.1 | 12 |
| 159 | Circular ecDNA promotes accessible chromatin and high oncogene expression. *Nature*, **2019**, 575, 699-703 | 50.4 | 134 |
| 158 | Novel insight into the genetic basis of high-altitude pulmonary hypertension in Kyrgyz highlanders. *European Journal of Human Genetics*, **2019**, 27, 150-159 | 5.3 | 7 |
| 157 | Identifying the favored mutation in a positive selective sweep. *Nature Methods*, **2018**, 15, 279-282 | 21.6 | 33 |
| 156 | MHC class I loaded ligands from breast cancer cell lines: A potential HLA-I-typed antigen collection. *Journal of Proteomics*, **2018**, 176, 13-23 | 3.9 | 17 |
| 155 | ViFi: accurate detection of viral integration and mRNA fusion reveals indiscriminate and unregulated transcription in proximal genomic regions in cervical cancer. *Nucleic Acids Research*, **2018**, 46, 3309-3325 | 20.1 | 23 |
| 154 | ProteoStorm: An Ultrafast Metaproteomics Database Search Framework. *Cell Systems*, **2018**, 7, 463-467.e6 | 10.6 | 18 |
| 153 | Targeted genotyping of variable number tandem repeats with adVNTR. *Genome Research*, **2018**, 28, 1709-1719 | 9.7 | 26 |
| 152 | Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. *Nature*, **2017**, 543, 122-125 | 50.4 | 260 |
| 151 | Clear: Composition of Likelihoods for Evolve and Resequence Experiments. *Genetics*, **2017**, 206, 1011-1023 | 4 | 17 |
| 150 | HapCUT2: robust and accurate haplotype assembly for diverse sequencing technologies. *Genome Research*, **2017**, 27, 801-812 | 9.7 | 161 |

| | | | |
|---|---|---|---|
| 149 | Ultraaccurate genome sequencing and haplotyping of single human cells. *Proceedings of the National Academy of Sciences of the United States of America*, **2017**, 114, 12512-12517 | 11.5 | 27 |
| 148 | High-altitude adaptation in humans: from genomics to integrative physiology. *Journal of Molecular Medicine*, **2017**, 95, 1269-1282 | 5.5 | 43 |
| 147 | The Antibody Repertoire of Colorectal Cancer. *Molecular and Cellular Proteomics*, **2017**, 16, 2111-2124 | 7.6 | 3 |
| 146 | New Insights into the Genetic Basis of Monge's Disease and Adaptation to High-Altitude. *Molecular Biology and Evolution*, **2017**, 34, 3154-3168 | 8.3 | 19 |
| 145 | Senp1 drives hypoxia-induced polycythemia via GATA1 and Bcl-xL in subjects with Monge's disease. *Journal of Experimental Medicine*, **2016**, 213, 2729-2744 | 16.6 | 22 |
| 144 | Integrated Proteogenomic Characterization of Human High-Grade Serous Ovarian Cancer. *Cell*, **2016**, 166, 755-765 | 56.2 | 544 |
| 143 | Diversity, Productivity, and Stability of an Industrial Microbial Ecosystem. *Applied and Environmental Microbiology*, **2016**, 82, 2494-2505 | 4.8 | 34 |
| 142 | InPhaDel: integrative shotgun and proximity-ligation sequencing to phase deletions with single nucleotide polymorphisms. *Nucleic Acids Research*, **2016**, 44, e111 | 20.1 | 1 |
| 141 | Endothelin receptor B, a candidate gene from human studies at high altitude, improves cardiac tolerance to hypoxia in genetically engineered heterozygote mice. *Proceedings of the National Academy of Sciences of the United States of America*, **2015**, 112, 10425-30 | 11.5 | 35 |
| 140 | Advanced Proteogenomic Analysis Reveals Multiple Peptide Mutations and Complex Immunoglobulin Peptides in Colon Cancer. *Journal of Proteome Research*, **2015**, 14, 3555-67 | 5.6 | 22 |
| 139 | Haplotype Allele Frequency (HAF) Score: Predicting Carriers of Ongoing Selective Sweeps Without Knowledge of the Adaptive Allele. *Lecture Notes in Computer Science*, **2015**, 276-280 | 0.9 | 1 |
| 138 | A global reference for human genetic variation. *Nature*, **2015**, 526, 68-74 | 50.4 | 8599 |
| 137 | Reconstructing breakage fusion bridge architectures using noisy copy numbers. *Journal of Computational Biology*, **2015**, 22, 577-94 | 1.7 | 7 |
| 136 | Next-Generation Sequencing of Plasmodium vivax Patient Samples Shows Evidence of Direct Evolution in Drug-Resistance Genes. *ACS Infectious Diseases*, **2015**, 1, 367-79 | 5.5 | 22 |
| 135 | Genetic simulation tools for post-genome wide association studies of complex diseases. *Genetic Epidemiology*, **2015**, 39, 11-19 | 2.6 | 15 |
| 134 | Predicting Carriers of Ongoing Selective Sweeps without Knowledge of the Favored Allele. *PLoS Genetics*, **2015**, 11, e1005527 | 6 | 13 |
| 133 | The TGFβ1 Promoter SNP C-509T and Food Sensitization Promote Esophageal Remodeling in Pediatric Eosinophilic Esophagitis. *PLoS ONE*, **2015**, 10, e0144651 | 3.7 | 20 |
| 132 | Using Genome Query Language to uncover genetic variation. *Bioinformatics*, **2014**, 30, 1-8 | 7.2 | 82 |

| | | | |
|---|---|---|---|
| 131 | An automated proteogenomic method uses mass spectrometry to reveal novel genes in Zea mays. *Molecular and Cellular Proteomics*, **2014**, 13, 157-67 | 7.6 | 69 |
| 130 | Annotation of the zebrafish genome through an integrated transcriptomic and proteomic analysis. *Molecular and Cellular Proteomics*, **2014**, 13, 3184-98 | 7.6 | 40 |
| 129 | Proteogenomic database construction driven from large scale RNA-seq data. *Journal of Proteome Research*, **2014**, 13, 21-8 | 5.6 | 90 |
| 128 | The genetic basis of chronic mountain sickness. *Physiology*, **2014**, 29, 403-12 | 9.8 | 21 |
| 127 | Proteogenomic strategies for identification of aberrant cancer peptides using large-scale next-generation sequencing data. *Proteomics*, **2014**, 14, 2719-30 | 4.8 | 47 |
| 126 | Amplification and thrifty single-molecule sequencing of recurrent somatic structural variations. *Genome Research*, **2014**, 24, 318-28 | 9.7 | 15 |
| 125 | The elusive evidence for chromothripsis. *Nucleic Acids Research*, **2014**, 42, 8231-42 | 20.1 | 30 |
| 124 | Inferring gene ontologies from pairwise similarity data. *Bioinformatics*, **2014**, 30, i34-42 | 7.2 | 57 |
| 123 | Whole genome sequencing of Ethiopian highlanders reveals conserved hypoxia tolerance genes. *Genome Biology*, **2014**, 15, R36 | 18.3 | 59 |
| 122 | Reconstructing Breakage Fusion Bridge Architectures Using Noisy Copy Numbers. *Lecture Notes in Computer Science*, **2014**, 400-417 | 0.9 | 1 |
| 121 | Evaluating genome architecture of a complex region via generalized bipartite matching. *BMC Bioinformatics*, **2013**, 14 Suppl 5, S13 | 3.6 | 1 |
| 120 | Whole-genome sequencing uncovers the genetic basis of chronic mountain sickness in Andean highlanders. *American Journal of Human Genetics*, **2013**, 93, 452-62 | 11 | 90 |
| 119 | Learning natural selection from the site frequency spectrum. *Genetics*, **2013**, 195, 181-93 | 4 | 78 |
| 118 | Virmid: accurate detection of somatic mutations with sample impurity inference. *Genome Biology*, **2013**, 14, R90 | 18.3 | 50 |
| 117 | On the design of clone-based haplotyping. *Genome Biology*, **2013**, 14, R100 | 18.3 | 17 |
| 116 | An algorithmic approach for breakage-fusion-bridge detection in tumor genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **2013**, 110, 5546-51 | 11.5 | 39 |
| 115 | Reprever: resolving low-copy duplicated sequences using template driven assembly. *Nucleic Acids Research*, **2013**, 41, e128 | 20.1 | 5 |
| 114 | Wessim: a whole-exome sequencing simulator based on in silico exome capture. *Bioinformatics*, **2013**, 29, 1076-7 | 7.2 | 26 |

| | | | |
|---|---|---|---|
| 113 | Abstractions for Genomics. *Communications of the ACM*, **2013**, 56, 83-93 | 2.5 | 9 |
| 112 | Cerulean: A Hybrid Assembly Using High Throughput Short and Long Reads. *Lecture Notes in Computer Science*, **2013**, 349-363 | 0.9 | 21 |
| 111 | Learning Natural Selection from the Site Frequency Spectrum. *Lecture Notes in Computer Science*, **2013**, 230-233 | 0.9 | 0 |
| 110 | An integrated map of genetic variation from 1,092 human genomes. *Nature*, **2012**, 491, 56-65 | 50.4 | 6049 |
| 109 | De novo somatic mutations in components of the PI3K-AKT3-mTOR pathway cause hemimegalencephaly. *Nature Genetics*, **2012**, 44, 941-5 | 36.3 | 505 |
| 108 | Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. *Genome Research*, **2012**, 22, 246-58 | 9.7 | 385 |
| 107 | Protein identification using top-down. *Molecular and Cellular Proteomics*, **2012**, 11, M111.008524 | 7.6 | 112 |
| 106 | Accurate mass spectrometry based protein quantification via shared peptides. *Journal of Computational Biology*, **2012**, 19, 337-48 | 1.7 | 31 |
| 105 | Combinatorics of the breakage-fusion-bridge mechanism. *Journal of Computational Biology*, **2012**, 19, 662-78 | 1.7 | 14 |
| 104 | Speeding up tandem mass spectral identification using indexes. *Bioinformatics*, **2012**, 28, 1692-7 | 7.2 | 10 |
| 103 | Exome sequencing can improve diagnosis and alter patient management. *Science Translational Medicine*, **2012**, 4, 138ra78 | 17.5 | 191 |
| 102 | iDASH: integrating data for analysis, anonymization, and sharing. *Journal of the American Medical Informatics Association: JAMIA*, **2012**, 19, 196-201 | 8.6 | 114 |
| 101 | Sample reproducibility of genetic association using different multimarker TDTs in genome-wide association studies: characterization and a new approach. *PLoS ONE*, **2012**, 7, e29613 | 3.7 | 4 |
| 100 | Automated querying and identification of novel peptides using MALDI mass spectrometric imaging. *Journal of Proteome Research*, **2011**, 10, 1915-28 | 5.6 | 28 |
| 99 | AMASS: algorithm for MSI analysis by semi-supervised segmentation. *Journal of Proteome Research*, **2011**, 10, 4734-43 | 5.6 | 22 |
| 98 | Tests of selection in pooled case-control data: an empirical study. *Frontiers in Genetics*, **2011**, 2, 83 | 4.5 | 3 |
| 97 | Strobe sequence design for haplotype assembly. *BMC Bioinformatics*, **2011**, 12 Suppl 1, S24 | 3.6 | 15 |
| 96 | Resurrection of a clinical antibody: template proteogenomic de novo proteomic sequencing and reverse engineering of an anti-lymphotoxin-β antibody. *Proteomics*, **2011**, 11, 395-405 | 4.8 | 25 |

| | | | |
|---|---|---|---|
| 95 | TCLUST: a fast method for clustering genome-scale expression data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **2011**, 8, 808-18 | 3 | 12 |
| 94 | On the Approximability of Reachability-Preserving Network Orientations. *Internet Mathematics*, **2011**, 7, 209-232 | 0 | 5 |
| 93 | Practical 4Sphosphopantetheine active site discovery from proteomic samples. *Journal of Proteome Research*, **2011**, 10, 320-9 | 5.6 | 14 |
| 92 | Compressing genomic sequence fragments using SlimGene. *Journal of Computational Biology*, **2011**, 18, 401-13 | 1.7 | 57 |
| 91 | Experimental selection of hypoxia-tolerant Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, **2011**, 108, 2349-54 | 11.5 | 85 |
| 90 | Sensitive gene fusion detection using ambiguously mapping RNA-Seq read pairs. *Bioinformatics*, **2011**, 27, 1068-75 | 7.2 | 48 |
| 89 | Deconvolution and database search of complex tandem mass spectra of intact proteins: a combinatorial approach. *Molecular and Cellular Proteomics*, **2010**, 9, 2772-82 | 7.6 | 130 |
| 88 | Template proteogenomics: sequencing whole proteins using an imperfect database. *Molecular and Cellular Proteomics*, **2010**, 9, 1260-70 | 7.6 | 40 |
| 87 | Protein-protein interaction network evaluation for identifying potential drug targets. *Journal of Computational Biology*, **2010**, 17, 669-84 | 1.7 | 24 |
| 86 | A covering method for detecting genetic associations between rare variants and common phenotypes. *PLoS Computational Biology*, **2010**, 6, e1000954 | 5 | 73 |
| 85 | RAPID detection of gene-gene interactions in genome-wide association studies. *Bioinformatics*, **2010**, 26, 2856-62 | 7.2 | 28 |
| 84 | Population sequencing of two endocannabinoid metabolic genes identifies rare and common regulatory variants associated with extreme obesity and metabolite level. *Genome Biology*, **2010**, 11, R118 | 18.3 | 23 |
| 83 | Deciphering the genetic basis of common diseases by integrated functional annotation of common and rare variants. *Genome Biology*, **2010**, 11, | 18.3 | 78 |
| 82 | Expansion of the mycobacterial "PUPylome". *Molecular BioSystems*, **2010**, 6, 376-85 | | 75 |
| 81 | Optimizing PCR assays for DNA-based cancer diagnostics. *Journal of Computational Biology*, **2010**, 17, 369-81 | 1.7 | 4 |
| 80 | Designing deep sequencing experiments: detecting structural variation and estimating transcript abundance. *BMC Genomics*, **2010**, 11, 385 | 4.5 | 20 |
| 79 | Construction of a medicinal leech transcriptome database and its application to the identification of leech homologs of neural and innate immune genes. *BMC Genomics*, **2010**, 11, 407 | 4.5 | 44 |
| 78 | Proteogenomics to discover the full coding content of genomes: a computational perspective. *Journal of Proteomics*, **2010**, 73, 2124-35 | 3.9 | 132 |

| | | | |
|---|---|---|---|
| 77 | Compressing Genomic Sequence Fragments Using SlimGene. *Lecture Notes in Computer Science*, **2010**, 310-324 | 0.9 | 7 |
| 76 | Fast and accurate alignment of multiple protein networks. *Journal of Computational Biology*, **2009**, 16, 989-99 | 1.7 | 47 |
| 75 | Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding. *Genome Research*, **2009**, 19, 1527-41 | 9.7 | 401 |
| 74 | Shared Peptides in Mass Spectrometry Based Protein Quantification. *Lecture Notes in Computer Science*, **2009**, 356-371 | 0.9 | 4 |
| 73 | Optimizing PCR Assays for DNA Based Cancer Diagnostics. *Lecture Notes in Computer Science*, **2009**, 220-235 | 0.9 | 23 |
| 72 | Accurate annotation of peptide modifications through unrestrictive database search. *Journal of Proteome Research*, **2008**, 7, 170-81 | 5.6 | 47 |
| 71 | Phosphorylation-specific MS/MS scoring for rapid and accurate phosphoproteome analysis. *Journal of Proteome Research*, **2008**, 7, 3373-81 | 5.6 | 45 |
| 70 | HapCUT: an efficient and accurate algorithm for the haplotype assembly problem. *Bioinformatics*, **2008**, 24, i153-9 | 7.2 | 198 |
| 69 | A multidimensional chromatography technology for in-depth phosphoproteome analysis. *Molecular and Cellular Proteomics*, **2008**, 7, 1389-96 | 7.6 | 392 |
| 68 | QNet: a tool for querying protein interaction networks. *Journal of Computational Biology*, **2008**, 15, 913-25 | 1.7 | 71 |
| 67 | Discovery and revision of Arabidopsis genes by proteogenomics. *Proceedings of the National Academy of Sciences of the United States of America*, **2008**, 105, 21034-8 | 11.5 | 226 |
| 66 | Evaluation of paired-end sequencing strategies for detection of genome rearrangements in cancer. *PLoS Computational Biology*, **2008**, 4, e1000051 | 5 | 63 |
| 65 | Comparative proteogenomics: combining mass spectrometry and comparative genomics to analyze multiple genomes. *Genome Research*, **2008**, 18, 1133-42 | 9.7 | 90 |
| 64 | An MCMC algorithm for haplotype assembly from whole-genome sequence data. *Genome Research*, **2008**, 18, 1336-46 | 9.7 | 88 |
| 63 | Structural alignment of pseudoknotted RNA. *Journal of Computational Biology*, **2008**, 15, 489-504 | 1.7 | 18 |
| 62 | Fast and Accurate Alignment of Multiple Protein Networks **2008**, 246-256 | | 28 |
| 61 | An Algorithm for Orienting Graphs Based on Cause-Effect Pairs and Its Applications to Orienting Protein Networks. *Lecture Notes in Computer Science*, **2008**, 222-232 | 0.9 | 23 |
| 60 | Whole proteome analysis of post-translational modifications: applications of mass-spectrometry for proteogenomic annotation. *Genome Research*, **2007**, 17, 1362-77 | 9.7 | 155 |

| | | | |
|---|---|---|---|
| 59 | The Sorcerer II Global Ocean Sampling expedition: expanding the universe of protein families. *PLoS Biology*, **2007**, 5, e16 | 9.7 | 638 |
| 58 | Evidence for large inversion polymorphisms in the human genome from HapMap data. *Genome Research*, **2007**, 17, 219-30 | 9.7 | 51 |
| 57 | The diploid genome sequence of an individual human. *PLoS Biology*, **2007**, 5, e254 | 9.7 | 1249 |
| 56 | A decomposition theory for phylogenetic networks and incompatible characters. *Journal of Computational Biology*, **2007**, 14, 1247-72 | 1.7 | 36 |
| 55 | Optimization of primer design for the detection of variable genomic lesions in cancer. *Bioinformatics*, **2007**, 23, 2807-15 | 7.2 | 16 |
| 54 | Improving gene annotation using peptide mass spectrometry. *Genome Research*, **2007**, 17, 231-9 | 9.7 | 144 |
| 53 | QNet: A Tool for Querying Protein Interaction Networks. *Lecture Notes in Computer Science*, **2007**, 1-15 | 0.9 | 26 |
| 52 | A sequence-based filtering method for ncRNA identification and its application to searching for riboswitch elements. *Bioinformatics*, **2006**, 22, e557-65 | 7.2 | 23 |
| 51 | Structural Alignment of Pseudoknotted RNA. *Lecture Notes in Computer Science*, **2006**, 143-158 | 0.9 | 2 |
| 50 | Consensus folding of unaligned RNA sequences revisited. *Journal of Computational Biology*, **2006**, 13, 283-95 | 1.7 | 23 |
| 49 | Inference about recombination from haplotype data: lower bounds and recombination hotspots. *Journal of Computational Biology*, **2006**, 13, 501-21 | 1.7 | 13 |
| 48 | Age-related changes in human crystallins determined from comparative analysis of post-translational modifications in young and aged lens: does deamidation contribute to crystallin insolubility?. *Journal of Proteome Research*, **2006**, 5, 2554-66 | 5.6 | 224 |
| 47 | Unrestrictive identification of post-translational modifications through peptide mass spectrometry. *Nature Protocols*, **2006**, 1, 67-72 | 18.8 | 32 |
| 46 | Searching genomes for noncoding RNA using FastR. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **2005**, 2, 366-79 | 3 | 37 |
| 45 | Identification of post-translational modifications via blind search of mass-spectra **2005**, 157-66 | | 13 |
| 44 | Peptide sequence tags for fast database search in mass-spectrometry. *Journal of Proteome Research*, **2005**, 4, 1287-95 | 5.6 | 103 |
| 43 | InsPecT: identification of posttranslationally modified peptides from tandem mass spectra. *Analytical Chemistry*, **2005**, 77, 4626-39 | 7.8 | 502 |
| 42 | Identification of post-translational modifications by blind search of mass spectra. *Nature Biotechnology*, **2005**, 23, 1562-7 | 44.5 | 231 |

| | | | |
|---|---|---|---|
| 41 | Polynomial and APX-hard cases of the individual haplotyping problem. *Theoretical Computer Science* , **2005**, 335, 109-125 | 1.1 | 36 |
| 40 | Improved Recombination Lower Bounds for Haplotype Data. *Lecture Notes in Computer Science*, **2005**, 569-584 | 0.9 | 7 |
| 39 | Orthologous repeats and mammalian phylogenetic inference. *Genome Research*, **2005**, 15, 998-1006 | 9.7 | 32 |
| 38 | A note on efficient computation of haplotypes via perfect phylogeny. *Journal of Computational Biology*, **2004**, 11, 858-66 | 1.7 | 25 |
| 37 | Optimal haplotype block-free selection of tagging SNPs for genome-wide association studies. *Genome Research*, **2004**, 14, 1633-40 | 9.7 | 84 |
| 36 | FastR: fast database search tool for non-coding RNA **2004**, 52-61 | | 13 |
| 35 | The number of recombination events in a sample history: conflict graph and lower bounds. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **2004**, 1, 78-90 | 3 | 31 |
| 34 | Shotgun protein sequencing by tandem mass spectra assembly. *Analytical Chemistry*, **2004**, 76, 7221-33 | 7.8 | 44 |
| 33 | A Survey of Computational Methods for Determining Haplotypes. *Lecture Notes in Computer Science*, **2004**, 26-47 | 0.9 | 55 |
| 32 | The dog genome: survey sequencing and comparative analysis. *Science*, **2003**, 301, 1898-903 | 33.3 | 422 |
| 31 | Haplotyping as perfect phylogeny: a direct approach. *Journal of Computational Biology*, **2003**, 10, 323-40 | 1.7 | 91 |
| 30 | Robustness of inference of haplotype block structure. *Journal of Computational Biology*, **2003**, 10, 13-9 | 1.7 | 51 |
| 29 | Haplotypes and informative SNP selection algorithms **2003**, | | 33 |
| 28 | Combinatorial Problems Arising in SNP and Haplotype Analysis. *Lecture Notes in Computer Science*, **2003**, 26-47 | 0.9 | 11 |
| 27 | Practical Algorithms and Fixed-Parameter Tractability for the Single Individual SNP Haplotyping Problem. *Lecture Notes in Computer Science*, **2002**, 29-43 | 0.9 | 38 |
| 26 | The sequence of the human genome. *Science*, **2001**, 291, 1304-51 | 33.3 | 10609 |
| 25 | SCOPE: a probabilistic model for scoring tandem mass spectra against a peptide database. *Bioinformatics*, **2001**, 17 Suppl 1, S13-21 | 7.2 | 134 |
| 24 | SNPs Problems, Complexity, and Algorithms. *Lecture Notes in Computer Science*, **2001**, 182-193 | 0.9 | 80 |

| | | | |
|---|---|---|---|
| 23 | Ligand-receptor pairing via tree comparison. *Journal of Computational Biology*, **2000**, 7, 59-70 | 1.7 | 10 |
| 22 | A Polynomial-Time Approximation Scheme for Minimum Routing Cost Spanning Trees. *SIAM Journal on Computing*, **2000**, 29, 761-778 | 1.1 | 78 |
| 21 | A 2-Approximation Algorithm for the Undirected Feedback Vertex Set Problem. *SIAM Journal on Discrete Mathematics*, **1999**, 12, 289-297 | 0.7 | 215 |
| 20 | Sorting by Transpositions. *SIAM Journal on Discrete Mathematics*, **1998**, 11, 224-240 | 0.7 | 258 |
| 19 | On the Approximability of Numerical Taxonomy (Fitting Distances by Tree Metrics). *SIAM Journal on Computing*, **1998**, 28, 1073-1085 | 1.1 | 52 |
| 18 | Detecting non-adjoining correlations with signals in DNA **1998**, | | 12 |
| 17 | Human beta-defensin 2 is a salt-sensitive peptide antibiotic expressed in human lung. *Journal of Clinical Investigation*, **1998**, 102, 874-80 | 15.9 | 449 |
| 16 | Approximation algorithms for multiple sequence alignment. *Theoretical Computer Science*, **1997**, 182, 233-244 | 1.1 | 29 |
| 15 | Genome Rearrangements and Sorting by Reversals. *SIAM Journal on Computing*, **1996**, 25, 272-289 | 1.1 | 252 |
| 14 | Nonoverlapping local alignments (weighted independent sets of axis-parallel rectangles). *Discrete Applied Mathematics*, **1996**, 71, 41-53 | 1 | 45 |
| 13 | Sorting by Reversals: Genome Rearrangements in Plant Organelles and Evolutionary History of X Chromosome. *Molecular Biology and Evolution*, **1995**, | 8.3 | 1 |
| 12 | Constant ratio approximations of the weighted feedback vertex set problem for undirected graphs. *Lecture Notes in Computer Science*, **1995**, 142-151 | 0.9 | 21 |
| 11 | Computing similarity between RNA strings. *Lecture Notes in Computer Science*, **1995**, 1-16 | 0.9 | 30 |
| 10 | Parallel implementation of logic languages. *Lecture Notes in Computer Science*, **1990**, 154-165 | 0.9 | |
| 9 | Targeted profiling of human extrachromosomal DNA by CRISPR-CATCH | | 1 |
| 8 | Frequent extrachromosomal oncogene amplification drives aggressive tumors | | 1 |
| 7 | AmpliconReconstructor: Integrated analysis of NGS and optical mapping resolves the complex structures of focal amplifications in cancer | | 1 |
| 6 | Fine-mapping the Favored Mutation in a Positive Selective Sweep | | 4 |

| | | |
|---|---|---|
| 5 | Variable Number Tandem Repeats mediate the expression of proximal genes | 2 |
| 4 | EcDNA hubs drive cooperative intermolecular oncogene expression | 3 |
| 3 | Assembly-free and alignment-free sample identification using genome skims | 1 |
| 2 | Principles of ecDNA random inheritance drive rapid genome change and therapy resistance in human cancers | 4 |
| 1 | Estimating repeat spectra and genome length from low-coverage genome skims with RESPECT | 2 |