S Cenk Sahinalp

List of Publications by Year
in descending order

Source: https://exaly.com/author-pdf/4813359/publications.pdf

Version: 2024-02-01

134 papers	18,972 citations	46918 47 h-index	18075 120 g-index
157 all docs	157 docs citations	157 times ranked	34939 citing authors

#	Article	IF	Citations
1	A map of human genome variation from population-scale sequencing. Nature, 2010, 467, 1061-1073.	13.7	7,209
2	PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics, 2010, 26, 1608-1615.	1.8	2,044
3	The evolutionary history of 2,658 cancers. Nature, 2020, 578, 122-128.	13.7	690
4	Personalized copy number and segmental duplication maps using next-generation sequencing. Nature Genetics, 2009, 41, 1061-1067.	9.4	656
5	Demographic history and rare allele sharing among human populations. Proceedings of the National Academy of Sciences of the United States of America, 2011, 108, 11983-11988.	3.3	589
6	deFuse: An Algorithm for Gene Fusion Discovery in Tumor RNA-Seq Data. PLoS Computational Biology, 2011, 7, e1001138.	1.5	477
7	Genomic hallmarks of localized, non-indolent prostate cancer. Nature, 2017, 541, 359-364.	13.7	462
8	Spatial genomic heterogeneity within localized, multifocal prostate cancer. Nature Genetics, 2015, 47, 736-745.	9.4	395
9	Widespread and Functional RNA Circularization in Localized Prostate Cancer. Cell, 2019, 176, 831-843.e22.	13.5	317
10	Comparative analysis of the small RNA transcriptomes of Pinus contorta and Oryza sativa. Genome Research, 2008, 18, 571-584.	2.4	305
11	Combinatorial algorithms for structural variation detection in high-throughput sequenced genomes. Genome Research, 2009, 19, 1270-1278.	2.4	266
12	Characterizing genetic intra-tumor heterogeneity across 2,658 human cancer genomes. Cell, 2021, 184, 2239-2254.e39.	13.5	260
13	mrsFAST: a cache-oblivious algorithm for short-read mapping. Nature Methods, 2010, 7, 576-577.	9.0	248
14	Applicability Domains for Classification Problems: Benchmarking of Distance to Models for Ames Mutagenicity Set. Journal of Chemical Information and Modeling, 2010, 50, 2094-2111.	2.5	202
15	Analysis of the androgen receptorâ€“regulated lncRNA landscape identifies a role for ARLNC1 in prostate cancer progression. Nature Genetics, 2018, 50, 814-824.	9.4	196
16	Next-generation VariationHunter: combinatorial algorithms for transposon insertion discovery. Bioinformatics, 2010, 26, i350-i357.	1.8	190
17	The structure and evolution of centromeric transition regions within the human genome. Nature, 2004, 430, 857-864.	13.7	179
18	The Proteogenomic Landscape of Curable Prostate Cancer. Cancer Cell, 2019, 35, 414-427.e6.	7.7	168

#	Article	IF	Citations
19	From sequence to molecular pathology, and a mechanism driving the neuroendocrine phenotype in prostate cancer. Journal of Pathology, 2012, 227, 286-297.	2.1	161
20	Biomolecular network motif counting and discovery by color coding. Bioinformatics, 2008, 24, i241-i249.	1.8	155
21	SRRM4 Drives Neuroendocrine Transdifferentiation of Prostate Adenocarcinoma Under Androgen Receptor Pathway Inhibition. European Urology, 2017, 71, 68-78.	0.9	136
22	SCALCE: boosting sequence compression algorithms using locally consistent encoding. Bioinformatics, 2012, 28, 3051-3057.	1.8	129
23	Analysis of Primate Genomic Variation Reveals a Repeat-Driven Expansion of the Human Genome. Genome Research, 2003, 13, 358-368.	2.4	127
24	Conifers have a unique small RNA silencing signature. Rna, 2008, 14, 1508-1515.	1.6	108
25	RNAâ€“RNA Interaction Prediction and Antisense RNA Target Search. Journal of Computational Biology, 2006, 13, 267-282.	0.8	106
26	Integrative inference of subclonal tumour evolution from single-cell and bulk sequencing data. Nature Communications, 2019, 10, 2750.	5.8	101
27	Detection and characterization of novel sequence insertions using paired-end next-generation sequencing. Bioinformatics, 2010, 26, 1277-1283.	1.8	98
28	Alu repeat discovery and characterization within human genomes. Genome Research, 2011, 21, 840-849.	2.4	94
29	Comparison of high-throughput sequencing data compression tools. Nature Methods, 2016, 13, 1005-1008.	9.0	91
30	BAP1 haploinsufficiency predicts a distinct immunogenic class of malignant peritoneal mesothelioma. Genome Medicine, 2019, 11, 8.	3.6	88
31	A partition function algorithm for interacting nucleic acid strands. Bioinformatics, 2009, 25, i365-i373.	1.8	85
32	Optimally discriminative subnetwork markers predict response to chemotherapy. Bioinformatics, 2011, 27, i205-i213.	1.8	81
33	Organization and Evolution of Primate Centromeric DNA from Whole-Genome Shotgun Sequence Data. PLoS Computational Biology, 2007, 3, e181.	1.5	80
34	HIT'nDRIVE: patient-specific multidriver gene prioritization for precision oncology. Genome Research, 2017, 27, 1573-1588.	2.4	78
35	Clonal Evolution and Heterogeneity of Osimertinib Acquired Resistance Mechanisms in EGFR Mutant Lung Cancer. Cell Reports Medicine, 2020, 1, 100007.	3.3	78
36	Not All Scale-Free Networks Are Born Equal: The Role of the Seed Graph in PPI Network Evolution. PLoS Computational Biology, 2007, 3, e118.	1.5	77

#	Article	IF	Citations
37	PRINCESS: Privacy-protecting Rare disease International Network Collaboration via Encryption through Software guard extensionS. Bioinformatics, 2017, 33, 871-878.	1.8	75
38	PhISCS: a combinatorial approach for subperfect tumor phylogeny reconstruction via integrative use of single-cell and bulk sequencing data. Genome Research, 2019, 29, 1860-1877.	2.4	73
39	nFuse: Discovery of complex genomic rearrangements in cancer using high-throughput sequencing. Genome Research, 2012, 22, 2250-2261.	2.4	67
40	Allelic decomposition and exact genotyping of highly polymorphic and structurally variant genes. Nature Communications, 2018, 9, 828.	5.8	67
41	Integrated genome and transcriptome sequencing identifies a novel form of hybrid and aggressive prostate cancer. Journal of Pathology, 2012, 227, 53-61.	2.1	63
42	Enabling Privacy-Preserving GWASs in Heterogeneous Human Populations. Cell Systems, 2016, 3, 54-61.	2.9	62
43	Simultaneous structural variation discovery among multiple paired-end sequenced genomes. Genome Research, 2011, 21, 2203-2212.	2.4	60
44	Stromal Gene Expression is Predictive for Metastatic Primary Prostate Cancer. European Urology, 2018, 73, 524-532.	0.9	60
45	Approximate nearest neighbors and sequence comparison with block operations. , 2000, , .		58
46	Mapping the Protein Interaction Network in Methicillin-Resistant Staphylococcus aureus. Journal of Proteome Research, 2011, 10, 1139-1150.	1.8	55
47	mrsFAST-Ultra: a compact, SNP-aware mapper for high performance sequencing applications. Nucleic Acids Research, 2014, 42, W494-W500.	6.5	54
48	The long noncoding RNA landscape of neuroendocrine prostate cancer and its clinical implications. GigaScience, 2018, 7, .	3.3	54
49	Comparative QSAR- and Fragments Distribution Analysis of Drugs, Druglikes, Metabolic Substances, and Antimicrobial Compounds. Journal of Chemical Information and Modeling, 2006, 46, 2167-2182.	2.5	50
50	CoLoRMap: Correcting Long Reads by Mapping short reads. Bioinformatics, 2016, 32, i545-i551.	1.8	49
51	SiNVICT: ultra-sensitive detection of single nucleotide variants and indels in circulating tumour DNA. Bioinformatics, 2017, 33, 26-34.	1.8	48
52	MechRNA: prediction of lncRNA mechanisms from RNAâ€“RNA and RNAâ€“protein interactions. Bioinformatics, 2018, 34, 3101-3110.	1.8	48
53	Ultra High-Dimensional Nonlinear Feature Selection for Big Biological Data. IEEE Transactions on Knowledge and Data Engineering, 2018, 30, 1352-1365.	4.0	48
54	Polyâ€gene fusion transcripts and chromothripsis in prostate cancer. Genes Chromosomes and Cancer, 2012, 51, 1144-1153.	1.5	46

#	Article	IF	Citations
55	Whole genome sequencing of Turkish genomes reveals functional private alleles and impact of genetic interactions with Europe, Asia and Africa. BMC Genomics, 2014, 15, 963.	1.2	46
56	Oblivious string embeddings and edit distance approximations. , 2006, , .		46
57	DeeZ: reference-based compression by local assembly. Nature Methods, 2014, 11, 1082-1084.	9.0	45
58	Fast prediction of RNA-RNA interaction. Algorithms for Molecular Biology, 2010, 5, 5.	0.3	44
59	Distance based algorithms for small biomolecule classification and structural similarity search. Bioinformatics, 2006, 22, e243-e251.	1.8	41
60	Comrad: detection of expressed rearrangements by integrated analysis of RNA-Seq and low coverage genome sequence data. Bioinformatics, 2011, 27, 1481-1488.	1.8	39
61	Cypiripi: exact genotyping of CYP2D6 using high-throughput sequencing data. Bioinformatics, 2015, 31, i27-i34.	1.8	37
62	Sketching algorithms for genomic data analysis and querying in a secure enclave. Nature Methods, 2020, 17, 295-301.	9.0	35
63	HITâ€™nDRIVE: Multi-driver Gene Prioritization Based on Hitting Time. Lecture Notes in Computer Science, 2014, , 293-306.	1.0	35
64	PRESAGE: PRivacy-preserving gEnetic testing via SoftwAre Guard Extension. BMC Medical Genomics, 2017, 10, 48.	0.7	32
65	CLIIQ: Accurate Comparative Detection and Quantification of Expressed Isoforms in a Population. Lecture Notes in Computer Science, 2012, , 178-189.	1.0	32
66	Protein-Protein Interaction Network Evaluation for Identifying Potential Drug Targets. Journal of Computational Biology, 2010, 17, 669-684.	0.8	30
67	ReMixT: clone-specific genomic structure estimation in cancer. Genome Biology, 2017, 18, 140.	3.8	29
68	biRNA: Fast RNA-RNA Binding Sites Prediction. Lecture Notes in Computer Science, 2009, , 25-36.	1.0	26
69	Divergent Origins and Concerted Expansion of Two Segmental Duplications on Chromosome 16. Journal of Heredity, 2001, 92, 462-468.	1.0	25
70	The Effect of Insertions and Deletions on Wirings in Protein-Protein Interaction Networks: A Large-Scale Study. Journal of Computational Biology, 2009, 16, 159-167.	0.8	25
71	A multi-labeled tree dissimilarity measure for comparing â€œclonal treesâ€•of tumor progression. Algorithms for Molecular Biology, 2019, 14, 17.	0.3	25
72	lordFAST: sensitive and Fast Alignment Search Tool for LOng noisy Read sequencing Data. Bioinformatics, 2019, 35, 20-27.	1.8	23

#	Article	IF	Citations
73	CYP2C8, CYP2C9, and CYP2C19 Characterization Using Next-Generation Sequencing and Haplotype Analysis. Journal of Molecular Diagnostics, 2022, 24, 337-350.	1.2	23
74	The Role of Unequal Crossover in Alpha-Satellite DNA Evolution: A Computational Analysis. Journal of Computational Biology, 2004, 11, 933-944.	0.8	20
75	Clonality Inference from Single Tumor Samples Using Low-Coverage Sequence Data. Journal of Computational Biology, 2017, 24, 515-523.	0.8	20
76	PhISCS-BnB: a fast branch and bound algorithm for the perfect tumor phylogeny reconstruction problem. Bioinformatics, 2020, 36, i169-i176.	1.8	19
77	Time and Space Efficient RNA-RNA Interaction Prediction via Sparse Folding. Lecture Notes in Computer Science, 2010, , 473-490.	1.0	18
78	Optimal spaced seeds for faster approximate string matching. Journal of Computer and System Sciences, 2007, 73, 1035-1044.	0.9	17
79	Sparsification of RNA structure prediction including pseudoknots. Algorithms for Molecular Biology, 2010, 5, 39.	0.3	17
80	ORMAN: Optimal resolution of ambiguous RNA-Seq multimappings in the presence of novel isoforms. Bioinformatics, 2014, 30, 644-651.	1.8	17
81	Structural variation and fusion detection using targeted sequencing data from circulating cell free DNA. Nucleic Acids Research, 2019, 47, e38-e38.	6.5	17
82	Optimal pooling for genome re-sequencing with ultra-high-throughput short-read technologies. Bioinformatics, 2008, 24, i32-i40.	1.8	16
83	Sensitive and fast mapping of di-base encoded reads. Bioinformatics, 2011, 27, 1915-1921.	1.8	16
84	Towards Improved Assessment of Functional Similarity in Large-Scale Screens: A Study on Indel Length. Journal of Computational Biology, 2010, 17, 1-20.	0.8	15
85	taveRNA: a web suite for RNA algorithms and applications. Nucleic Acids Research, 2007, 35, W325-W329.	6.5	14
86	Mutational Analysis of Gene Fusions Predicts Novel MHC Class Iâ€“Restricted T-Cell Epitopes and Immune Signatures in a Subset of Prostate Cancer. Clinical Cancer Research, 2017, 23, 7596-7607.	3.2	14
87	Optimal compressed representation of high throughput sequence data via light assembly. Nature Communications, 2018, 9, 566.	5.8	14
88	Permutation Editing and Matching via Embeddings. Lecture Notes in Computer Science, 2001, , 481-492.	1.0	14
89	Dissect: detection and characterization of novel structural alterations in transcribed sequences. Bioinformatics, 2012, 28, i179-i187.	1.8	13
90	Privacy-preserving genotype imputation in a trusted execution environment. Cell Systems, 2021, 12, 983-993.e7.	2.9	13

#	Article	IF	Citations
91	smyRNA: A Novel Ab Initio ncRNA Gene Finder. PLoS ONE, 2009, 4, e5433.	1.1	13
92	Barnacle: detecting and characterizing tandem duplications and fusions in transcriptome assemblies. BMC Genomics, 2013, 14, 550.	1.2	12
93	Tumor Phylogeny Topology Inference via Deep Learning. IScience, 2020, 23, 101655.	1.9	11
94	Sublinear Methods for Detecting Periodic Trends in Data Streams. Lecture Notes in Computer Science, 2004, , 16-28.	1.0	11
95	RNA Secondary Structure Prediction Via Energy Density Minimization. Lecture Notes in Computer Science, 2006, , 130-142.	1.0	9
96	Combinatorial Detection of Conserved Alteration Patterns for Identifying Cancer Subnetworks. GigaScience, 2019, 8, .	3.3	9
97	Identification of conserved evolutionary trajectories in tumors. Bioinformatics, 2020, 36, i427-i435.	1.8	9
98	Simple and Practical Sequence Nearest Neighbors with Block Operations. Lecture Notes in Computer Science, 2002, , 262-278.	1.0	9
99	Computational identification of micro-structural variations and their proteogenomic consequences in cancer. Bioinformatics, 2018, 34, 1672-1681.	1.8	8
100	Improved Duplication Models for Proteome Network Evolution. , 2005, , 119-137.		7
101	Novel approaches for small biomolecule classification and structural similarity search. SIGKDD Explorations: Newsletter of the Special Interest Group (SIG) on Knowledge Discovery & Data Mining, 2007, 9, 14-21.	3.2	7
102	Analytical Validation of a Computational Method for Pharmacogenetic Genotyping from Clinical Whole Exome Sequencing. Journal of Molecular Diagnostics, 2022, 24, 576-585.	1.2	7
103	On a parallel-algorithms method for string matching problems (overview). Lecture Notes in Computer Science, 1994, , 22-32.	1.0	6
104	Fast Prediction of RNA-RNA Interaction. Lecture Notes in Computer Science, 2009, , 261-272.	1.0	6
105	Manipulating multiple sequence alignments via MaM and WebMaM. Nucleic Acids Research, 2005, 33, W295-W298.	6.5	5
106	Protecting Genomic Data Privacy with Probabilistic Modeling. , 2018, , .		5
107	Mirroring co-evolving trees in the light of their topologies. Bioinformatics, 2012, 28, 1202-1208.	1.8	4
108	Graph Traversal Edit Distance and Extensions. Journal of Computational Biology, 2020, 27, 317-329.	0.8	4

#	Article	IF	Citations
109	Simultaneous Structural Variation Discovery in Multiple Paired-End Sequenced Genomes. Lecture Notes in Computer Science, 2011, , 104-105.	1.0	4
110	Hardness of String Similarity Search and Other Indexing Problems. Lecture Notes in Computer Science, 2004, , 1080-1098.	1.0	4
111	Quantifying Systemic Evolutionary Changes by Color Coding Confidence-Scored PPI Networks. Lecture Notes in Computer Science, 2009, , 37-48.	1.0	4
112	The Complexity of Gene Placement. Journal of Algorithms, 2001, 41, 225-243.	0.9	3
113	Periodicity testing with sublinear samples and space. ACM Transactions on Algorithms, 2010, 6, 1-14.	0.9	3
114	SubGraph2Vec: Highly-Vectorized Tree-like Subgraph Counting. , 2019, , .		3
115	Sparsification of RNA Structure Prediction Including Pseudoknots. Lecture Notes in Computer Science, 2010, , 40-51.	1.0	3
116	Robustness of Massively Parallel Sequencing Platforms. PLoS ONE, 2015, 10, e0138259.	1.1	3
117	Combinatorial Algorithms for Structural Variation Detection in High Throughput Sequenced Genomes. Lecture Notes in Computer Science, 2009, , 218-219.	1.0	3
118	Locally Consistent Parsing and Applications to Approximate String Comparisons. Lecture Notes in Computer Science, 2005, , 22-35.	1.0	2
119	Edit Distance Under Block Operations. , 2008, , 265-267.		2
120	GTED: Graph Traversal Edit Distance. Lecture Notes in Computer Science, 2018, , 37-53.	1.0	1
121	Abstract 378: The Cancer Genome Collaboratory. Cancer Research, 2017, 77, 378-378.	0.4	1
122	Genomic Data Compression. , 2019, , 783-783.		1
123	The Relation between Indel Length and Functional Divergence: A Formal Study. Lecture Notes in Computer Science, 2008, , 330-341.	1.0	1
124	Protecting Genomic Data Privacy with Probabilistic Modeling. Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing, 2019, 24, 403-414.	0.7	1
125	IDENTIFYING UNIFORMLY MUTATED SEGMENTS WITHIN REPEATS. Journal of Bioinformatics and Computational Biology, 2004, 02, 657-668.	0.3	0
126	Edit Distance Under Block Operations. , 2016, , 611-614.		0

#	Article	IF	Citations
127	Preface: Selected Articles from RECOMB 2017. Journal of Computational Biology, 2018, 25, 623-623.	0.8	0
128	COMPARATIVE QSAR ANALYSIS OF BACTERIAL, FUNGAL PLANT AND HUMAN METABOLITES. , 2006, , .		0
129	COMPUTATIONAL STUDIES OF NON-CODING RNAS â€“ Session Introduction. , 2009, , 54-56.		0
130	PERSONAL GENOMICS â€“ Session Introduction. , 2009, , 302-304.		0
131	Pair HMM Based Gap Statistics for Re-evaluation of Indels in Alignments with Affine Gap Penalties. Lecture Notes in Computer Science, 2010, , 350-361.	1.0	0
132	Proteome Network Emulating Models. , 2012, , 69-95.		0
133	Abstract B129: Clinical implications of inter- and intra- prostatic heterogeneity.. , 2013, , .		0
134	Not All Scale Free Networks Are Born Equal: The Role of the Seed Graph in PPI Network Emulation. Lecture Notes in Computer Science, 2006, , 1-13.	1.0	0