

Tak-Wah Lam

List of Publications by Citations

Source: <https://exaly.com/author-pdf/9420297/tak-wah-lam-publications-by-citations.pdf>

Version: 2024-04-27

This document has been generated based on the publications and citations recorded by exaly.com. For the latest version of this publication list, visit the link given above.

The third column is the impact factor (IF) of the journal, and the fourth column is the number of citations of the article.

128
papers

12,040
citations

22
h-index

109
g-index

151
ext. papers

16,198
ext. citations

3.1
avg, IF

5.95
L-index

#	Paper	IF	Citations
128	SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. <i>GigaScience</i> , 2012 , 1, 18	7.6	3152
127	SOAP2: an improved ultrafast tool for short read alignment. <i>Bioinformatics</i> , 2009 , 25, 1966-7	7.2	2784
126	MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. <i>Bioinformatics</i> , 2015 , 31, 1674-6	7.2	1973
125	The sequence and de novo assembly of the giant panda genome. <i>Nature</i> , 2010 , 463, 311-7	50.4	864
124	SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. <i>Bioinformatics</i> , 2014 , 30, 1660-6	7.2	621
123	MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. <i>Methods</i> , 2016 , 102, 3-11	4.6	529
122	Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. <i>GigaScience</i> , 2013 , 2, 10	7.6	461
121	SOAP3: ultra-fast GPU-based parallel alignment tool for short reads. <i>Bioinformatics</i> , 2012 , 28, 878-9	7.2	158
120	COPE: an accurate k-mer-based pair-end reads connection tool to facilitate genome assembly. <i>Bioinformatics</i> , 2012 , 28, 2870-4	7.2	118
119	SOAPSplICE: Genome-Wide ab initio Detection of Splice Junctions from RNA-Seq Data. <i>Frontiers in Genetics</i> , 2011 , 2, 46	4.5	82
118	SOAP3-dp: fast, accurate and sensitive GPU-based short read aligner. <i>PLoS ONE</i> , 2013 , 8, e65632	3.7	76
117	A multi-task convolutional deep neural network for variant calling in single molecule sequencing. <i>Nature Communications</i> , 2019 , 10, 998	17.4	63
116	Next generation sequencing has lower sequence coverage and poorer SNP-detection capability in the regulatory regions. <i>Scientific Reports</i> , 2011 , 1, 55	4.9	61
115	Scheduling for Speed Bounded Processors. <i>Lecture Notes in Computer Science</i> , 2008 , 409-420	0.9	55
114	Compressed indexes for dynamic text collections. <i>ACM Transactions on Algorithms</i> , 2007 , 3, 21	1.2	52
113	A Space and Time Efficient Algorithm for Constructing Compressed Suffix Arrays. <i>Algorithmica</i> , 2007 , 48, 23-36	0.9	45
112	Speed Scaling Functions for Flow Time Scheduling Based on Active Job Count. <i>Lecture Notes in Computer Science</i> , 2008 , 647-659	0.9	41

111	Exploring the limit of using a deep neural network on pileup data for germline variant calling. <i>Nature Machine Intelligence</i> , 2020 , 2, 220-227	22.5	40
110	An Even Faster and More Unifying Algorithm for Comparing Trees via Unbalanced Bipartite Matchings. <i>Journal of Algorithms</i> , 2001 , 40, 212-233		39
109	A Decomposition Theorem for Maximum Weight Bipartite Matchings. <i>SIAM Journal on Computing</i> , 2001 , 31, 18-26	1.1	32
108	Predicting RNA secondary structures with arbitrary pseudoknots by maximizing the number of stacking pairs. <i>Journal of Computational Biology</i> , 2003 , 10, 981-95	1.7	31
107	SOAPfusion: a robust and effective computational fusion discovery tool for RNA-seq reads. <i>Bioinformatics</i> , 2013 , 29, 2971-8	7.2	22
106	Optimizing throughput and energy in online deadline scheduling. <i>ACM Transactions on Algorithms</i> , 2009 , 6, 1-22	1.2	21
105	Deadline scheduling and power management for speed bounded processors. <i>Theoretical Computer Science</i> , 2010 , 411, 3587-3600	1.1	20
104	Compressed Index for Dictionary Matching 2008 ,		20
103	Competitive non-migratory scheduling for flow time and energy 2008 ,		19
102	Finding Connected Components in $O(\log n \log \log n)$ Time on the EREW PRAM. <i>Journal of Algorithms</i> , 1995 , 18, 378-402		19
101	From Peer-Reviewed to Peer-Reproduced in Scholarly Publishing: The Complementary Roles of Data Models and Workflows in Bioinformatics. <i>PLoS ONE</i> , 2015 , 10, e0127612	3.7	18
100	Succinct Text Indexing with Wildcards. <i>Lecture Notes in Computer Science</i> , 2009 , 39-50	0.9	18
99	Nonclairvoyant Speed Scaling for Flow and Energy. <i>Algorithmica</i> , 2011 , 61, 507-517	0.9	17
98	Reconstructing an ultrametric galled phylogenetic network from a distance matrix. <i>Journal of Bioinformatics and Computational Biology</i> , 2006 , 4, 807-32	1	17
97	FaSD-somatic: a fast and accurate somatic SNV detection algorithm for cancer genome sequencing data. <i>Bioinformatics</i> , 2014 , 30, 2498-500	7.2	15
96	Continuous Monitoring of Distributed Data Streams over a Time-Based Sliding Window. <i>Algorithmica</i> , 2012 , 62, 1088-1111	0.9	15
95	Space Efficient Indexes for String Matching with Don't Care 2007 , 846-857		15
94	Optimal Edge Ranking of Trees in Linear Time. <i>Algorithmica</i> , 2001 , 30, 12-33	0.9	14

93	A Linear Size Index for Approximate Pattern Matching. <i>Lecture Notes in Computer Science</i> , 2006 , 49-59	0.9	14
92	GLProbs: Aligning Multiple Sequences Adaptively. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2015 , 12, 67-78	3	13
91	Approximate String Matching Using Compressed Suffix Arrays. <i>Lecture Notes in Computer Science</i> , 2004 , 434-444	0.9	13
90	A Decomposition Theorem for MaximumWeight Bipartite Matchings with Applications to Evolutionary Trees. <i>Lecture Notes in Computer Science</i> , 1999 , 438-449	0.9	13
89	Compressed Indexes for Approximate String Matching. <i>Algorithmica</i> , 2010 , 58, 263-281	0.9	12
88	Sleep with Guilt and Work Faster to Minimize Flow Plus Energy. <i>Lecture Notes in Computer Science</i> , 2009 , 665-676	0.9	12
87	Structural alignment of RNA with complex pseudoknot structure. <i>Journal of Computational Biology</i> , 2011 , 18, 97-108	1.7	11
86	Nonmigratory Online Deadline Scheduling on Multiprocessors. <i>SIAM Journal on Computing</i> , 2005 , 34, 669-682	1.1	11
85	Results on communication complexity classes. <i>Journal of Computer and System Sciences</i> , 1992 , 44, 324-342		11
84	BALSA: integrated secondary analysis for whole-genome and whole-exome sequencing, accelerated by GPU. <i>PeerJ</i> , 2014 , 2, e421	3.1	11
83	Succinct Index for Dynamic Dictionary Matching. <i>Lecture Notes in Computer Science</i> , 2009 , 1034-1043	0.9	11
82	Clairvoyante: a multi-task convolutional deep neural network for variant calling in Single Molecule Sequencing		11
81	Serine peptidase inhibitor Kazal type 1 (SPINK1) as novel downstream effector of the cadherin-17/Eatenin axis in hepatocellular carcinoma. <i>Cellular Oncology (Dordrecht)</i> , 2017 , 40, 443-456	7.2	9
80	RENET: A Deep Learning Approach for Extracting Gene-Disease Associations from Literature. <i>Lecture Notes in Computer Science</i> , 2019 , 272-284	0.9	9
79	MegaGTA: a sensitive and accurate metagenomic gene-targeted assembler using iterative de Bruijn graphs. <i>BMC Bioinformatics</i> , 2017 , 18, 408	3.6	9
78	Improved multi-processor scheduling for flow time and energy. <i>Journal of Scheduling</i> , 2012 , 15, 105-116	1.6	9
77	A linear size index for approximate pattern matching. <i>Journal of Discrete Algorithms</i> , 2011 , 9, 358-364		9
76	General techniques for comparing unrooted evolutionary trees 1997 ,		9

75	Non-clairvoyant Speed Scaling for Weighted Flow Time. <i>Lecture Notes in Computer Science</i> , 2010 , 23-35	0.9	9
74	Improved Approximate String Matching Using Compressed Suffix Data Structures. <i>Lecture Notes in Computer Science</i> , 2005 , 339-348	0.9	9
73	Improved Approximate String Matching Using Compressed Suffix Data Structures. <i>Algorithmica</i> , 2008 , 51, 298-314	0.9	8
72	Cavity Matchings, Label Compressions, and Unrooted Evolutionary Trees. <i>SIAM Journal on Computing</i> , 2000 , 30, 602-624	1.1	8
71	MICA: A fast short-read aligner that takes full advantage of Many Integrated Core Architecture (MIC). <i>BMC Bioinformatics</i> , 2015 , 16 Suppl 7, S10	3.6	7
70	Performance guarantee for edf under overload. <i>Journal of Algorithms</i> , 2004 , 52, 193-206		7
69	Extra processors versus future information in optimal deadline scheduling 2002 ,		7
68	Potential utility of metagenomic sequencing for improving etiologic diagnosis of infective endocarditis. <i>Future Cardiology</i> , 2019 , 15, 411-424	1.3	7
67	Online Speed Scaling Based on Active Job Count to Minimize Flow Plus Energy. <i>Algorithmica</i> , 2013 , 65, 605-633	0.9	6
66	Nonmigratory Multiprocessor Scheduling for Response Time and Energy. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2008 , 19, 1527-1539	3.7	6
65	Trade-offs between communication and space. <i>Journal of Computer and System Sciences</i> , 1992 , 45, 296-315		6
64	Compressing Dictionary Matching Index via Sparsification Technique. <i>Algorithmica</i> , 2015 , 72, 515-538	0.9	5
63	Dictionary Matching with a Bounded Gap in Pattern or in Text. <i>Algorithmica</i> , 2018 , 80, 698-713	0.9	5
62	Sleep Management on Multiple Machines for Energy and Flow Time. <i>Lecture Notes in Computer Science</i> , 2011 , 219-231	0.9	5
61	RNASAlign: RNA structural alignment system. <i>Bioinformatics</i> , 2011 , 27, 2151-2	7.2	5
60	On-line stream merging in a general setting. <i>Theoretical Computer Science</i> , 2003 , 296, 27-46	1.1	5
59	Clair: Exploring the limit of using a deep neural network on pileup data for germline variant calling		5
58	BASE: a practical de novo assembler for large genomes using long NGS reads. <i>BMC Genomics</i> , 2016 , 17 Suppl 5, 499	4.5	5

57	Scheduling with Gaps: New Models and Algorithms. <i>Lecture Notes in Computer Science</i> , 2015 , 114-126	0.9	4
56	Scheduling for weighted flow time and energy with rejection penalty. <i>Theoretical Computer Science</i> , 2013 , 470, 93-104	1.1	4
55	Improving multiple sequence alignment by using better guide trees. <i>BMC Bioinformatics</i> , 2015 , 16 Suppl 5, S4	3.6	4
54	Extra Processors versus Future Information in Optimal Deadline Scheduling. <i>Theory of Computing Systems</i> , 2004 , 37, 323-341	0.6	4
53	Improving the efficiency of parallel minimum spanning tree algorithms. <i>Discrete Applied Mathematics</i> , 2003 , 126, 33-54	1	4
52	Automatic construction of online catalog topologies. <i>IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews</i> , 2002 , 32, 382-391		4
51	Predicting RNA secondary structures with arbitrary pseudoknots by maximizing the number of stacking pairs 2001 ,		4
50	Tradeoff between Energy and Throughput for Online Deadline Scheduling. <i>Lecture Notes in Computer Science</i> , 2011 , 59-70	0.9	4
49	database.bio: a web application for interpreting human variations. <i>Bioinformatics</i> , 2015 , 31, 4035-7	7.2	3
48	A simple and economical method for improving whole genome alignment. <i>BMC Genomics</i> , 2017 , 18, 362	4.5	3
47	Nonclairvoyant sleep management and flow-time scheduling on multiple processors 2013 ,		3
46	Approximating Frequent Items in Asynchronous Data Stream over a Sliding Window. <i>Algorithms</i> , 2011 , 4, 200-222	1.8	3
45	Cache-oblivious index for approximate string matching. <i>Theoretical Computer Science</i> , 2011 , 412, 3579-3588		3
44	Finding alternative splicing patterns with strong support from expressed sequences on individual exons/introns. <i>Journal of Bioinformatics and Computational Biology</i> , 2008 , 6, 1021-33	1	3
43	Skyhawk: An Artificial Neural Network-based discriminator for reviewing clinically significant genomic variants		3
42	Tradeoff between energy and throughput for online deadline scheduling. <i>Sustainable Computing: Informatics and Systems</i> , 2011 , 1, 189-195	3	2
41	Non-shared edges and nearest neighbor interchanges revisited. <i>Information Processing Letters</i> , 2004 , 91, 129-134	0.8	2
40	A MEMORY EFFICIENT ALGORITHM FOR STRUCTURAL ALIGNMENT OF RNAs WITH EMBEDDED SIMPLE PSEUDOKNOTS 2007 ,		2

39	Approximating Frequent Items in Asynchronous Data Stream over a Sliding Window. <i>Lecture Notes in Computer Science</i> , 2010 , 49-61	0.9	2
38	Edit Distance to Monotonicity in Sliding Windows. <i>Lecture Notes in Computer Science</i> , 2011 , 564-573	0.9	2
37	Approximating the Nearest Neighbor Interchange Distance for Evolutionary Trees with Non-uniform Degrees. <i>Lecture Notes in Computer Science</i> , 1999 , 61-70	0.9	2
36	Online Tracking of the Dominance Relationship of Distributed Multi-dimensional Data. <i>Lecture Notes in Computer Science</i> , 2011 , 178-189	0.9	2
35	CONNET: Accurate Genome Consensus in Assembling Nanopore Sequencing Data via Deep Learning. <i>IScience</i> , 2020 , 23, 101128	6.1	2
34	Building a Chinese pan-genome of 486 individuals. <i>Communications Biology</i> , 2021 , 4, 1016	6.7	2
33	Finding least-weight subsequences with fewer processors. <i>Lecture Notes in Computer Science</i> , 1990 , 318-327	3.7	2
32	Non-clairvoyant Weighted Flow Time Scheduling on Different Multi-processor Models. <i>Theory of Computing Systems</i> , 2015 , 56, 82-95	0.6	1
31	MegaPath: sensitive and rapid pathogen detection using metagenomic NGS data. <i>BMC Genomics</i> , 2020 , 21, 500	4.5	1
30	Non-clairvoyant weighted flow time scheduling with rejection penalty 2012 ,		1
29	Extra Unit-Speed Machines Are Almost as Powerful as Speedy Machines for Flow Time Scheduling. <i>SIAM Journal on Computing</i> , 2008 , 37, 1595-1612	1.1	1
28	SUBTREE TRANSFER DISTANCE FOR DEGREE-D PHYLOGENIES. <i>International Journal of Foundations of Computer Science</i> , 2004 , 15, 893-909	0.6	1
27	Competitive Deadline Scheduling via Additional or Faster Processors. <i>Journal of Scheduling</i> , 2003 , 6, 213-223	1.2	1
26	On-line Stream Merging with Max Span and Min Coverage. <i>Theory of Computing Systems</i> , 2005 , 38, 461-479	0.6	1
25	A Tighter Extra-Resource Analysis of Online Deadline Scheduling. <i>Journal of Combinatorial Optimization</i> , 2005 , 9, 157-165	0.9	1
24	APPROXIMATING THE NEAREST NEIGHBOR INTERCHARGE DISTANCE FOR NON-UNIFORM-DEGREE EVOLUTIONARY TREES. <i>International Journal of Foundations of Computer Science</i> , 2001 , 12, 533-550	0.6	1
23	Dictionary Matching with Uneven Gaps. <i>Lecture Notes in Computer Science</i> , 2015 , 247-260	0.9	1
22	Adjacent nucleotide dependence in ncRNA and order-1 SCFG for ncRNA identification. <i>PLoS ONE</i> , 2010 , 5, e12848	3.7	1

21	Online Flow Time Scheduling in the Presence of Preemption Overhead. <i>Lecture Notes in Computer Science</i> , 2012 , 85-97	0.9	1
20	Non-clairvoyant Weighted Flow Time Scheduling on Different Multi-processor Models. <i>Lecture Notes in Computer Science</i> , 2012 , 137-149	0.9	1
19	AC-DIAMOND: Accelerating Protein Alignment via Better SIMD Parallelization and Space-Efficient Indexing. <i>Lecture Notes in Computer Science</i> , 2016 , 426-433	0.9	1
18	Draft Genome Sequence of <i>Helicobacter cinaedi</i> , Compiled by Direct Whole-Genome Sequencing of a Blood Culture-Positive Isolate in Hong Kong. <i>Microbiology Resource Announcements</i> , 2018 , 7,	1.3	1
17	Transcriptome Analysis of Acute Phase Liver Graft Injury in Liver Transplantation. <i>Biomedicines</i> , 2018 , 6,	4.8	1
16	AC-DIAMOND v1: accelerating large-scale DNA-protein alignment. <i>Bioinformatics</i> , 2018 , 34, 3744-3746	7.2	1
15	Compressed Indexes for Approximate String Matching. <i>Lecture Notes in Computer Science</i> , 2006 , 208-219	0.9	1
14	ECNano: A cost-effective workflow for target enrichment sequencing and accurate variant calling on 4800 clinically significant genes using a single MinION flowcell.. <i>BMC Medical Genomics</i> , 2022 , 15, 43	3.7	1
13	Detecting structural variations with precise breakpoints using low-depth WGS data from a single oxford nanopore MinION flowcell.. <i>Scientific Reports</i> , 2022 , 12, 4519	4.9	1
12	PnpProbs: a better multiple sequence alignment tool by better handling of guide trees. <i>BMC Bioinformatics</i> , 2016 , 17 Suppl 8, 285	3.6	0
11	Identification of Cooperative Gene Regulation Among Transcription Factors, LncRNAs, and MicroRNAs in Diabetic Nephropathy Progression. <i>Frontiers in Genetics</i> , 2020 , 11, 1008	4.5	0
10	RENET2: high-performance full-text gene-disease relation extraction with iterative training data expansion. <i>NAR Genomics and Bioinformatics</i> , 2021 , 3, lqab062	3.7	0
9	Guest Editors Foreword. <i>Algorithmica</i> , 2015 , 73, 621-622	0.9	
8	Aggressive Online Deadline Scheduling. <i>Electronic Notes in Theoretical Computer Science</i> , 2004 , 91, 148-157	0.9	
7	HKG: an open genetic variant database of 205 Hong Kong cantonese exomes.. <i>NAR Genomics and Bioinformatics</i> , 2022 , 4, lqac005	3.7	
6	Cache-Oblivious Index for Approximate String Matching. <i>Lecture Notes in Computer Science</i> , 2007 , 40-51	0.9	
5	Unbalanced and Hierarchical Bipartite Matchings with Applications to Labeled Tree Comparison. <i>Lecture Notes in Computer Science</i> , 2000 , 479-490	0.9	
4	Improved Phylogeny Comparisons: Non-shared Edges, Nearest Neighbor Interchanges, and Subtree Transfers. <i>Lecture Notes in Computer Science</i> , 2000 , 527-538	0.9	

- 3 Energy-Efficient Due Date Scheduling. *Lecture Notes in Computer Science*, **2011**, 69-80 0.9
- 2 LCR_Finder: A de Novo Low Copy Repeat Finder for Human Genome. *Lecture Notes in Computer Science*, **2013**, 125-136 0.9
- 1 Scheduling with gaps: new models and algorithms. *Journal of Scheduling*, **2021**, 24, 381-403 1.6