

# Tak-Wah Lam

## List of Publications by Year in Descending Order

**Source:** <https://exaly.com/author-pdf/9420297/tak-wah-lam-publications-by-year.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	HKG: an open genetic variant database of 205 Hong Kong cantonese exomes.. <i>NAR Genomics and Bioinformatics</i> , <b>2022</b> , 4, lqac005	3.7	
127	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> , <b>2022</b> , 15, 43	3.7	1
126	Detecting structural variations with precise breakpoints using low-depth WGS data from a single oxford nanopore MinION flowcell.. <i>Scientific Reports</i> , <b>2022</b> , 12, 4519	4.9	1
125	RENET2: high-performance full-text gene-disease relation extraction with iterative training data expansion. <i>NAR Genomics and Bioinformatics</i> , <b>2021</b> , 3, lqab062	3.7	0
124	Scheduling with gaps: new models and algorithms. <i>Journal of Scheduling</i> , <b>2021</b> , 24, 381-403	1.6	
123	Building a Chinese pan-genome of 486 individuals. <i>Communications Biology</i> , <b>2021</b> , 4, 1016	6.7	2
122	MegaPath: sensitive and rapid pathogen detection using metagenomic NGS data. <i>BMC Genomics</i> , <b>2020</b> , 21, 500	4.5	1
121	Exploring the limit of using a deep neural network on pileup data for germline variant calling. <i>Nature Machine Intelligence</i> , <b>2020</b> , 2, 220-227	22.5	40
120	Identification of Cooperative Gene Regulation Among Transcription Factors, LncRNAs, and MicroRNAs in Diabetic Nephropathy Progression. <i>Frontiers in Genetics</i> , <b>2020</b> , 11, 1008	4.5	0
119	CONNET: Accurate Genome Consensus in Assembling Nanopore Sequencing Data via Deep Learning. <i>IScience</i> , <b>2020</b> , 23, 101128	6.1	2
118	RENET: A Deep Learning Approach for Extracting Gene-Disease Associations from Literature. <i>Lecture Notes in Computer Science</i> , <b>2019</b> , 272-284	0.9	9
117	A multi-task convolutional deep neural network for variant calling in single molecule sequencing. <i>Nature Communications</i> , <b>2019</b> , 10, 998	17.4	63
116	Potential utility of metagenomic sequencing for improving etiologic diagnosis of infective endocarditis. <i>Future Cardiology</i> , <b>2019</b> , 15, 411-424	1.3	7
115	Dictionary Matching with a Bounded Gap in Pattern or in Text. <i>Algorithmica</i> , <b>2018</b> , 80, 698-713	0.9	5
114	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> , <b>2018</b> , 7,	1.3	1
113	Transcriptome Analysis of Acute Phase Liver Graft Injury in Liver Transplantation. <i>Biomedicines</i> , <b>2018</b> , 6,	4.8	1
112	AC-DIAMOND v1: accelerating large-scale DNA-protein alignment. <i>Bioinformatics</i> , <b>2018</b> , 34, 3744-3746	7.2	1

111	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> , <b>2017</b> , 40, 443-456	7.2	9
110	MegaGTA: a sensitive and accurate metagenomic gene-targeted assembler using iterative de Bruijn graphs. <i>BMC Bioinformatics</i> , <b>2017</b> , 18, 408	3.6	9
109	A simple and economical method for improving whole genome alignment. <i>BMC Genomics</i> , <b>2017</b> , 18, 362	4.5	3
108	PnpProbs: a better multiple sequence alignment tool by better handling of guide trees. <i>BMC Bioinformatics</i> , <b>2016</b> , 17 Suppl 8, 285	3.6	0
107	MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. <i>Methods</i> , <b>2016</b> , 102, 3-11	4.6	529
106	AC-DIAMOND: Accelerating Protein Alignment via Better SIMD Parallelization and Space-Efficient Indexing. <i>Lecture Notes in Computer Science</i> , <b>2016</b> , 426-433	0.9	1
105	BASE: a practical de novo assembler for large genomes using long NGS reads. <i>BMC Genomics</i> , <b>2016</b> , 17 Suppl 5, 499	4.5	5
104	Scheduling with Gaps: New Models and Algorithms. <i>Lecture Notes in Computer Science</i> , <b>2015</b> , 114-126	0.9	4
103	GLProbs: Aligning Multiple Sequences Adaptively. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , <b>2015</b> , 12, 67-78	3	13
102	database.bio: a web application for interpreting human variations. <i>Bioinformatics</i> , <b>2015</b> , 31, 4035-7	7.2	3
101	Non-clairvoyant Weighted Flow Time Scheduling on Different Multi-processor Models. <i>Theory of Computing Systems</i> , <b>2015</b> , 56, 82-95	0.6	1
100	Compressing Dictionary Matching Index via Sparsification Technique. <i>Algorithmica</i> , <b>2015</b> , 72, 515-538	0.9	5
99	Guest Editors Foreword. <i>Algorithmica</i> , <b>2015</b> , 73, 621-622	0.9	
98	MICA: A fast short-read aligner that takes full advantage of Many Integrated Core Architecture (MIC). <i>BMC Bioinformatics</i> , <b>2015</b> , 16 Suppl 7, S10	3.6	7
97	Improving multiple sequence alignment by using better guide trees. <i>BMC Bioinformatics</i> , <b>2015</b> , 16 Suppl 5, S4	3.6	4
96	MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. <i>Bioinformatics</i> , <b>2015</b> , 31, 1674-6	7.2	1973
95	From Peer-Reviewed to Peer-Reproduced in Scholarly Publishing: The Complementary Roles of Data Models and Workflows in Bioinformatics. <i>PLoS ONE</i> , <b>2015</b> , 10, e0127612	3.7	18
94	Dictionary Matching with Uneven Gaps. <i>Lecture Notes in Computer Science</i> , <b>2015</b> , 247-260	0.9	1

93	FaSD-somatic: a fast and accurate somatic SNV detection algorithm for cancer genome sequencing data. <i>Bioinformatics</i> , <b>2014</b> , 30, 2498-500	7.2	15
92	SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. <i>Bioinformatics</i> , <b>2014</b> , 30, 1660-6	7.2	621
91	BALSA: integrated secondary analysis for whole-genome and whole-exome sequencing, accelerated by GPU. <i>PeerJ</i> , <b>2014</b> , 2, e421	3.1	11
90	Online Speed Scaling Based on Active Job Count to Minimize Flow Plus Energy. <i>Algorithmica</i> , <b>2013</b> , 65, 605-633	0.9	6
89	Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. <i>GigaScience</i> , <b>2013</b> , 2, 10	7.6	461
88	Scheduling for weighted flow time and energy with rejection penalty. <i>Theoretical Computer Science</i> , <b>2013</b> , 470, 93-104	1.1	4
87	SOAPfusion: a robust and effective computational fusion discovery tool for RNA-seq reads. <i>Bioinformatics</i> , <b>2013</b> , 29, 2971-8	7.2	22
86	Nonclairvoyant sleep management and flow-time scheduling on multiple processors <b>2013</b> ,		3
85	SOAP3-dp: fast, accurate and sensitive GPU-based short read aligner. <i>PLoS ONE</i> , <b>2013</b> , 8, e65632	3.7	76
84	LCR_Finder: A de Novo Low Copy Repeat Finder for Human Genome. <i>Lecture Notes in Computer Science</i> , <b>2013</b> , 125-136	0.9	
83	Improved multi-processor scheduling for flow time and energy. <i>Journal of Scheduling</i> , <b>2012</b> , 15, 105-116	1.6	9
82	Continuous Monitoring of Distributed Data Streams over a Time-Based Sliding Window. <i>Algorithmica</i> , <b>2012</b> , 62, 1088-1111	0.9	15
81	COPE: an accurate k-mer-based pair-end reads connection tool to facilitate genome assembly. <i>Bioinformatics</i> , <b>2012</b> , 28, 2870-4	7.2	118
80	SOAP3: ultra-fast GPU-based parallel alignment tool for short reads. <i>Bioinformatics</i> , <b>2012</b> , 28, 878-9	7.2	158
79	SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. <i>GigaScience</i> , <b>2012</b> , 1, 18	7.6	3152
78	Non-clairvoyant weighted flow time scheduling with rejection penalty <b>2012</b> ,		1
77	Online Flow Time Scheduling in the Presence of Preemption Overhead. <i>Lecture Notes in Computer Science</i> , <b>2012</b> , 85-97	0.9	1
76	Non-clairvoyant Weighted Flow Time Scheduling on Different Multi-processor Models. <i>Lecture Notes in Computer Science</i> , <b>2012</b> , 137-149	0.9	1

75	Next generation sequencing has lower sequence coverage and poorer SNP-detection capability in the regulatory regions. <i>Scientific Reports</i> , <b>2011</b> , 1, 55	4.9	61
74	Approximating Frequent Items in Asynchronous Data Stream over a Sliding Window. <i>Algorithms</i> , <b>2011</b> , 4, 200-222	1.8	3
73	SOAPsplice: Genome-Wide ab initio Detection of Splice Junctions from RNA-Seq Data. <i>Frontiers in Genetics</i> , <b>2011</b> , 2, 46	4.5	82
72	Tradeoff between energy and throughput for online deadline scheduling. <i>Sustainable Computing: Informatics and Systems</i> , <b>2011</b> , 1, 189-195	3	2
71	A linear size index for approximate pattern matching. <i>Journal of Discrete Algorithms</i> , <b>2011</b> , 9, 358-364		9
70	Nonclairvoyant Speed Scaling for Flow and Energy. <i>Algorithmica</i> , <b>2011</b> , 61, 507-517	0.9	17
69	Structural alignment of RNA with complex pseudoknot structure. <i>Journal of Computational Biology</i> , <b>2011</b> , 18, 97-108	1.7	11
68	Cache-oblivious index for approximate string matching. <i>Theoretical Computer Science</i> , <b>2011</b> , 412, 3579-3588		3
67	Sleep Management on Multiple Machines for Energy and Flow Time. <i>Lecture Notes in Computer Science</i> , <b>2011</b> , 219-231	0.9	5
66	RNASAlign: RNA structural alignment system. <i>Bioinformatics</i> , <b>2011</b> , 27, 2151-2	7.2	5
65	Edit Distance to Monotonicity in Sliding Windows. <i>Lecture Notes in Computer Science</i> , <b>2011</b> , 564-573	0.9	2
64	Online Tracking of the Dominance Relationship of Distributed Multi-dimensional Data. <i>Lecture Notes in Computer Science</i> , <b>2011</b> , 178-189	0.9	2
63	Energy-Efficient Due Date Scheduling. <i>Lecture Notes in Computer Science</i> , <b>2011</b> , 69-80	0.9	
62	Tradeoff between Energy and Throughput for Online Deadline Scheduling. <i>Lecture Notes in Computer Science</i> , <b>2011</b> , 59-70	0.9	4
61	The sequence and de novo assembly of the giant panda genome. <i>Nature</i> , <b>2010</b> , 463, 311-7	50.4	864
60	Compressed Indexes for Approximate String Matching. <i>Algorithmica</i> , <b>2010</b> , 58, 263-281	0.9	12
59	Deadline scheduling and power management for speed bounded processors. <i>Theoretical Computer Science</i> , <b>2010</b> , 411, 3587-3600	1.1	20
58	Approximating Frequent Items in Asynchronous Data Stream over a Sliding Window. <i>Lecture Notes in Computer Science</i> , <b>2010</b> , 49-61	0.9	2

57	Non-clairvoyant Speed Scaling for Weighted Flow Time. <i>Lecture Notes in Computer Science</i> , <b>2010</b> , 23-35	0.9	9
56	Adjacent nucleotide dependence in ncRNA and order-1 SCFG for ncRNA identification. <i>PLoS ONE</i> , <b>2010</b> , 5, e12848	3.7	1
55	Optimizing throughput and energy in online deadline scheduling. <i>ACM Transactions on Algorithms</i> , <b>2009</b> , 6, 1-22	1.2	21
54	SOAP2: an improved ultrafast tool for short read alignment. <i>Bioinformatics</i> , <b>2009</b> , 25, 1966-7	7.2	2784
53	Sleep with Guilt and Work Faster to Minimize Flow Plus Energy. <i>Lecture Notes in Computer Science</i> , <b>2009</b> , 665-676	0.9	12
52	Succinct Text Indexing with Wildcards. <i>Lecture Notes in Computer Science</i> , <b>2009</b> , 39-50	0.9	18
51	Succinct Index for Dynamic Dictionary Matching. <i>Lecture Notes in Computer Science</i> , <b>2009</b> , 1034-1043	0.9	11
50	Compressed Index for Dictionary Matching <b>2008</b> ,		20
49	Extra Unit-Speed Machines Are Almost as Powerful as Speedy Machines for Flow Time Scheduling. <i>SIAM Journal on Computing</i> , <b>2008</b> , 37, 1595-1612	1.1	1
48	Nonmigratory Multiprocessor Scheduling for Response Time and Energy. <i>IEEE Transactions on Parallel and Distributed Systems</i> , <b>2008</b> , 19, 1527-1539	3.7	6
47	Finding alternative splicing patterns with strong support from expressed sequences on individual exons/introns. <i>Journal of Bioinformatics and Computational Biology</i> , <b>2008</b> , 6, 1021-33	1	3
46	Competitive non-migratory scheduling for flow time and energy <b>2008</b> ,		19
45	Improved Approximate String Matching Using Compressed Suffix Data Structures. <i>Algorithmica</i> , <b>2008</b> , 51, 298-314	0.9	8
44	Scheduling for Speed Bounded Processors. <i>Lecture Notes in Computer Science</i> , <b>2008</b> , 409-420	0.9	55
43	Speed Scaling Functions for Flow Time Scheduling Based on Active Job Count. <i>Lecture Notes in Computer Science</i> , <b>2008</b> , 647-659	0.9	41
42	A Space and Time Efficient Algorithm for Constructing Compressed Suffix Arrays. <i>Algorithmica</i> , <b>2007</b> , 48, 23-36	0.9	45
41	Compressed indexes for dynamic text collections. <i>ACM Transactions on Algorithms</i> , <b>2007</b> , 3, 21	1.2	52
40	A MEMORY EFFICIENT ALGORITHM FOR STRUCTURAL ALIGNMENT OF RNAs WITH EMBEDDED SIMPLE PSEUDOKNOTS <b>2007</b> ,		2

39	Cache-Oblivious Index for Approximate String Matching. <i>Lecture Notes in Computer Science</i> , <b>2007</b> , 40-51	0.9	
38	Space Efficient Indexes for String Matching with Don't Cares <b>2007</b> , 846-857		15
37	Reconstructing an ultrametric galled phylogenetic network from a distance matrix. <i>Journal of Bioinformatics and Computational Biology</i> , <b>2006</b> , 4, 807-32	1	17
36	A Linear Size Index for Approximate Pattern Matching. <i>Lecture Notes in Computer Science</i> , <b>2006</b> , 49-59	0.9	14
35	Compressed Indexes for Approximate String Matching. <i>Lecture Notes in Computer Science</i> , <b>2006</b> , 208-219	0.9	1
34	On-line Stream Merging with Max Span and Min Coverage. <i>Theory of Computing Systems</i> , <b>2005</b> , 38, 461-476		1
33	A Tighter Extra-Resource Analysis of Online Deadline Scheduling. <i>Journal of Combinatorial Optimization</i> , <b>2005</b> , 9, 157-165	0.9	1
32	Nonmigratory Online Deadline Scheduling on Multiprocessors. <i>SIAM Journal on Computing</i> , <b>2005</b> , 34, 669-682	1.1	11
31	Improved Approximate String Matching Using Compressed Suffix Data Structures. <i>Lecture Notes in Computer Science</i> , <b>2005</b> , 339-348	0.9	9
30	Approximate String Matching Using Compressed Suffix Arrays. <i>Lecture Notes in Computer Science</i> , <b>2004</b> , 434-444	0.9	13
29	SUBTREE TRANSFER DISTANCE FOR DEGREE-D PHYLOGENIES. <i>International Journal of Foundations of Computer Science</i> , <b>2004</b> , 15, 893-909	0.6	1
28	Aggressive Online Deadline Scheduling. <i>Electronic Notes in Theoretical Computer Science</i> , <b>2004</b> , 91, 148-157		
27	Extra Processors versus Future Information in Optimal Deadline Scheduling. <i>Theory of Computing Systems</i> , <b>2004</b> , 37, 323-341	0.6	4
26	Performance guarantee for edf under overload. <i>Journal of Algorithms</i> , <b>2004</b> , 52, 193-206		7
25	Non-shared edges and nearest neighbor interchanges revisited. <i>Information Processing Letters</i> , <b>2004</b> , 91, 129-134	0.8	2
24	Competitive Deadline Scheduling via Additional or Faster Processors. <i>Journal of Scheduling</i> , <b>2003</b> , 6, 213-223		1
23	Improving the efficiency of parallel minimum spanning tree algorithms. <i>Discrete Applied Mathematics</i> , <b>2003</b> , 126, 33-54	1	4
22	On-line stream merging in a general setting. <i>Theoretical Computer Science</i> , <b>2003</b> , 296, 27-46	1.1	5

21	Predicting RNA secondary structures with arbitrary pseudoknots by maximizing the number of stacking pairs. <i>Journal of Computational Biology</i> , <b>2003</b> , 10, 981-95	1.7	31
20	Extra processors versus future information in optimal deadline scheduling <b>2002</b> ,		7
19	Automatic construction of online catalog topologies. <i>IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews</i> , <b>2002</b> , 32, 382-391		4
18	An Even Faster and More Unifying Algorithm for Comparing Trees via Unbalanced Bipartite Matchings. <i>Journal of Algorithms</i> , <b>2001</b> , 40, 212-233		39
17	APPROXIMATING THE NEAREST NEIGHBOR INTERCHARGE DISTANCE FOR NON-UNIFORM-DEGREE EVOLUTIONARY TREES. <i>International Journal of Foundations of Computer Science</i> , <b>2001</b> , 12, 533-550	0.6	1
16	Optimal Edge Ranking of Trees in Linear Time. <i>Algorithmica</i> , <b>2001</b> , 30, 12-33	0.9	14
15	A Decomposition Theorem for Maximum Weight Bipartite Matchings. <i>SIAM Journal on Computing</i> , <b>2001</b> , 31, 18-26	1.1	32
14	Predicting RNA secondary structures with arbitrary pseudoknots by maximizing the number of stacking pairs <b>2001</b> ,		4
13	Cavity Matchings, Label Compressions, and Unrooted Evolutionary Trees. <i>SIAM Journal on Computing</i> , <b>2000</b> , 30, 602-624	1.1	8
12	Unbalanced and Hierarchical Bipartite Matchings with Applications to Labeled Tree Comparison. <i>Lecture Notes in Computer Science</i> , <b>2000</b> , 479-490	0.9	
11	Improved Phylogeny Comparisons: Non-shared Edges, Nearest Neighbor Interchanges, and Subtree Transfers. <i>Lecture Notes in Computer Science</i> , <b>2000</b> , 527-538	0.9	
10	Approximating the Nearest Neighbor Interchange Distance for Evolutionary Trees with Non-uniform Degrees. <i>Lecture Notes in Computer Science</i> , <b>1999</b> , 61-70	0.9	2
9	A Decomposition Theorem for MaximumWeight Bipartite Matchings with Applications to Evolutionary Trees. <i>Lecture Notes in Computer Science</i> , <b>1999</b> , 438-449	0.9	13
8	General techniques for comparing unrooted evolutionary trees <b>1997</b> ,		9
7	Finding Connected Components in $O(\log n \log \log n)$ Time on the EREW PRAM. <i>Journal of Algorithms</i> , <b>1995</b> , 18, 378-402		19
6	Results on communication complexity classes. <i>Journal of Computer and System Sciences</i> , <b>1992</b> , 44, 324-342		11
5	Trade-offs between communication and space. <i>Journal of Computer and System Sciences</i> , <b>1992</b> , 45, 296-315		6
4	Finding least-weight subsequences with fewer processors. <i>Lecture Notes in Computer Science</i> , <b>1990</b> , 318-327		2



- 3 Clairvoyante: a multi-task convolutional deep neural network for variant calling in Single Molecule Sequencing 11
- 2 Skyhawk: An Artificial Neural Network-based discriminator for reviewing clinically significant genomic variants 3
- 1 Clair: Exploring the limit of using a deep neural network on pileup data for germline variant calling 5