Haixu Tang

List of Publications by Citations

Source: https://exaly.com/author-pdf/1899162/haixu-tang-publications-by-citations.pdf
Version: 2024-04-29

121 papers	6,269 citations	40 h-index	77 g-index
130 ext. papers	7,732 ext. citations	7.2 avg, IF	5.85 L-index

#	Paper	IF	Citations
121	The ecoresponsive genome of Daphnia pulex. *Science*, **2011**, 331, 555-61	33.3	924
120	FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Research*, **2010**, 38, e191	20.1	535
119	Rate and molecular spectrum of spontaneous mutations in the bacterium Escherichia coli as determined by whole-genome sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, **2012**, 109, E2774-83	11.5	418
118	RAPSearch2: a fast and memory-efficient protein similarity search tool for next-generation sequencing data. *Bioinformatics*, **2012**, 28, 125-6	7.2	292
117	The transcriptional diversity of 25 Drosophila cell lines. *Genome Research*, **2011**, 21, 301-14	9.7	171
116	Ancestral reconstruction of segmental duplications reveals punctuated cores of human genome evolution. *Nature Genetics*, **2007**, 39, 1361-8	36.3	162
115	De novo repeat classification and fragment assembly. *Genome Research*, **2004**, 14, 1786-96	9.7	157
114	Identification of Pol IV and RDR2-dependent precursors of 24 nt siRNAs guiding de novo DNA methylation in Arabidopsis. *ELife*, **2015**, 4, e09591	8.9	146
113	A computational approach toward label-free protein quantification using predicted peptide detectability. *Bioinformatics*, **2006**, 22, e481-8	7.2	141
112	Spatial and functional relationships among Pol V-associated loci, Pol IV-dependent siRNAs, and cytosine methylation in the Arabidopsis epigenome. *Genes and Development*, **2012**, 26, 1825-36	12.6	115
111	Learning your identity and disease from research papers **2009**,		113
110	Diverse CRISPRs evolving in human microbiomes. *PLoS Genetics*, **2012**, 8, e1002441	6	100
109	Regulation of adjacent yeast genes. *Trends in Genetics*, **2000**, 16, 109-11	8.5	99
108	Leaky Cauldron on the Dark Land: Understanding Memory Side-Channel Hazards in SGX **2017**, 2017, 2421-2434		97
107	Whole-genome sequencing and assembly with high-throughput, short-read technologies. *PLoS ONE*, **2007**, 2, e484	3.7	94
106	Interlaboratory study on differential analysis of protein glycosylation by mass spectrometry: the ABRF glycoprotein research multi-institutional study 2012. *Molecular and Cellular Proteomics*, **2013**, 12, 2935-51	7.6	92
105	RAPSearch: a fast protein similarity search tool for short reads. *BMC Bioinformatics*, **2011**, 12, 159	3.6	92

104	Automated interpretation of MS/MS spectra of oligosaccharides. *Bioinformatics*, **2005**, 21 Suppl 1, i431-9	7.2	89
103	A novel method for multiple alignment of sequences with repeated and shuffled elements. *Genome Research*, **2004**, 14, 2336-46	9.7	83
102	ISEScan: automated identification of insertion sequence elements in prokaryotic genomes. *Bioinformatics*, **2017**, 33, 3340-3347	7.2	82
101	Computational framework for identification of intact glycopeptides in complex samples. *Analytical Chemistry*, **2014**, 86, 453-63	7.8	79
100	Fragment assembly with double-barreled data. *Bioinformatics*, **2001**, 17 Suppl 1, S225-33	7.2	79
99	A two-step process for epigenetic inheritance in Arabidopsis. *Molecular Cell*, **2014**, 54, 30-42	17.6	78
98	Determinants of spontaneous mutation in the bacterium Escherichia coli as revealed by whole-genome sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, **2015**, 112, E5990-9	11.5	76
97	Quantitative glycomics strategies. *Molecular and Cellular Proteomics*, **2013**, 12, 874-84	7.6	76
96	De novo identification of LTR retrotransposons in eukaryotic genomes. *BMC Genomics*, **2007**, 8, 90	4.5	64
95	Improving confidence in detection and characterization of protein N-glycosylation sites and microheterogeneity. *Rapid Communications in Mass Spectrometry*, **2011**, 25, 2007-19	2.2	61
94	Strand-biased cytosine deamination at the replication fork causes cytosine to thymine mutations in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, **2016**, 113, 2176-81	11.5	58
93	Efficient Genome-Wide, Privacy-Preserving Similar Patient Query based on Private Edit Distance **2015**,		55
92	Quantitative measurement of phosphoproteome response to osmotic stress in arabidopsis based on Library-Assisted eXtracted Ion Chromatogram (LAXIC). *Molecular and Cellular Proteomics*, **2013**, 12, 2354-69	7.6	55
91	A bayesian approach to protein inference problem in shotgun proteomics. *Journal of Computational Biology*, **2009**, 16, 1183-93	1.7	55
90	A novel alignment method and multiple filters for exclusion of unqualified peptides to enhance label-free quantification using peptide intensity in LC-MS/MS. *Journal of Proteome Research*, **2011**, 10, 4799-812	5.6	54
89	On the mutational topology of the bacterial genome. *G3: Genes, Genomes, Genetics*, **2013**, 3, 399-407	3.2	53
88	MGEScan-non-LTR: computational identification and classification of autonomous non-LTR retrotransposons in eukaryotic genomes. *Nucleic Acids Research*, **2009**, 37, e143	20.1	52
87	Insertion sequence-caused large-scale rearrangements in the genome of Escherichia coli. *Nucleic Acids Research*, **2016**, 44, 7109-19	20.1	49

86	Background Mutational Features of the Radiation-Resistant Bacterium Deinococcus radiodurans. *Molecular Biology and Evolution*, **2015**, 32, 2383-92	8.3	45
85	Shotgun protein sequencing by tandem mass spectra assembly. *Analytical Chemistry*, **2004**, 76, 7221-33	7.8	44
84	Addressing Beacon re-identification attacks: quantification and mitigation of privacy risks. *Journal of the American Medical Informatics Association: JAMIA*, **2017**, 24, 799-805	8.6	43
83	BuildSummary: using a group-based approach to improve the sensitivity of peptide/protein identification in shotgun proteomics. *Journal of Proteome Research*, **2012**, 11, 1494-502	5.6	42
82	Testosterone affects neural gene expression differently in male and female juncos: a role for hormones in mediating sexual dimorphism and conflict. *PLoS ONE*, **2013**, 8, e61784	3.7	42
81	N-glycan profiling by microchip electrophoresis to differentiate disease states related to esophageal adenocarcinoma. *Analytical Chemistry*, **2012**, 84, 3621-7	7.8	40
80	On the accuracy and limits of peptide fragmentation spectrum prediction. *Analytical Chemistry*, **2011**, 83, 790-6	7.8	40
79	Glycoproteomics: identifying the glycosylation of prostate specific antigen at normal and high isoelectric points by LC-MS/MS. *Journal of Proteome Research*, **2014**, 13, 5570-80	5.6	38
78	Mapping site-specific protein N-glycosylations through liquid chromatography/mass spectrometry and targeted tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, **2010**, 24, 965-72	2.2	36
77	Label-free glycopeptide quantification for biomarker discovery in human sera. *Journal of Proteome Research*, **2014**, 13, 4821-32	5.6	34
76	Automated annotation and quantitation of glycans by liquid chromatography/electrospray ionization mass spectrometric analysis using the MultiGlycan-ESI computational tool. *Rapid Communications in Mass Spectrometry*, **2015**, 29, 135-42	2.2	31
75	The importance of peptide detectability for protein identification, quantification, and experiment design in MS/MS proteomics. *Journal of Proteome Research*, **2010**, 9, 6288-97	5.6	31
74	An ORFome assembly approach to metagenomics sequences analysis. *Journal of Bioinformatics and Computational Biology*, **2009**, 7, 455-71	1	31
73	Adaptation of to Long-Term Serial Passage in Complex Medium: Evidence of Parallel Evolution. *MSystems*, **2017**, 2,	7.6	30
72	Protecting genomic data analytics in the cloud: state of the art and opportunities. *BMC Medical Genomics*, **2016**, 9, 63	3.7	30
71	Utilizing de Bruijn graph of metagenome assembly for metatranscriptome analysis. *Bioinformatics*, **2016**, 32, 1001-8	7.2	29
70	RNAMotifScan: automatic identification of RNA structural motifs using secondary structural alignment. *Nucleic Acids Research*, **2010**, 38, e176	20.1	29
69	A Graph-Centric Approach for Metagenome-Guided Peptide and Protein Identification in Metaproteomics. *PLoS Computational Biology*, **2016**, 12, e1005224	5	29

68	Automated Glycan Sequencing from Tandem Mass Spectra of N-Linked Glycopeptides. *Analytical Chemistry*, **2016**, 88, 5725-32	7.8	27
67	Potential for sexual conflict assessed via testosterone-mediated transcriptional changes in liver and muscle of a songbird. *Journal of Experimental Biology*, **2014**, 217, 507-17	3	24
66	Identifying repeat domains in large genomes. *Genome Biology*, **2006**, 7, R7	18.3	23
65	A community effort to protect genomic data sharing, collaboration and outsourcing. *Npj Genomic Medicine*, **2017**, 2, 33	6.2	22
64	LTR retroelements in the genome of Daphnia pulex. *BMC Genomics*, **2010**, 11, 425	4.5	21
63	Rational design of a more stable penicillin G acylase against organic cosolvent. *Journal of Molecular Catalysis B: Enzymatic*, **2002**, 18, 285-290		21
62	A machine-learning approach to combined evidence validation of genome assemblies. *Bioinformatics*, **2008**, 24, 744-50	7.2	20
61	Characterization of the Glycosylation Site of Human PSA Prompted by Missense Mutation using LC-MS/MS. *Journal of Proteome Research*, **2015**, 14, 2872-83	5.6	19
60	Determinants of Base-Pair Substitution Patterns Revealed by Whole-Genome Sequencing of DNA Mismatch Repair Defective. *Genetics*, **2018**, 209, 1029-1042	4	19
59	Genome assembly, rearrangement, and repeats. *Chemical Reviews*, **2007**, 107, 3391-406	68.1	19
58	Integration of time-series meta-omics data reveals how microbial ecosystems respond to disturbance. *Nature Communications*, **2020**, 11, 5281	17.4	19
57	A machine learning approach to predicting peptide fragmentation spectra. *Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing*, **2006**, 219-30	1.3	19
56	Choosing blindly but wisely: differentially private solicitation of DNA datasets for disease marker discovery. *Journal of the American Medical Informatics Association: JAMIA*, **2015**, 22, 100-8	8.6	18
55	8-plex LC-MS/MS Analysis of Permethylated -Glycans Achieved by Using Stable Isotopic Iodomethane. *Analytical Chemistry*, **2019**, 91, 11794-11802	7.8	18
54	Full-Spectrum Prediction of Peptides Tandem Mass Spectra using Deep Neural Network. *Analytical Chemistry*, **2020**, 92, 4275-4283	7.8	17
53	Proteomic changes in the photoreceptor outer segment upon intense light exposure. *Journal of Proteome Research*, **2010**, 9, 1173-81	5.6	17
52	MOGONET integrates multi-omics data using graph convolutional networks allowing patient classification and biomarker identification. *Nature Communications*, **2021**, 12, 3445	17.4	17
51	Roles of bacteriophages, plasmids and CRISPR immunity in microbial community dynamics revealed using time-series integrated meta-omics. *Nature Microbiology*, **2021**, 6, 123-135	26.6	16

50	Insertion Polymorphisms of Mobile Genetic Elements in Sexual and Asexual Populations of Daphnia pulex. *Genome Biology and Evolution*, **2017**, 9, 362-374	3.9	14
49	Combinatorial libraries of synthetic peptides as a model for shotgun proteomics. *Analytical Chemistry*, **2010**, 82, 6559-68	7.8	14
48	XLSearch: a Probabilistic Database Search Algorithm for Identifying Cross-Linked Peptides. *Journal of Proteome Research*, **2016**, 15, 1830-41	5.6	14
47	Detection of structural variants involving repetitive regions in the reference genome. *Journal of Computational Biology*, **2014**, 21, 219-33	1.7	13
46	Probabilistic inference of biochemical reactions in microbial communities from metagenomic sequences. *PLoS Computational Biology*, **2013**, 9, e1002981	5	12
45	Independent mammalian genome contractions following the KT boundary. *Genome Biology and Evolution*, **2009**, 1, 2-12	3.9	12
44	ADVANCEMENT IN PROTEIN INFERENCE FROM SHOTGUN PROTEOMICS USING PEPTIDE DETECTABILITY **2006**,		12
43	A Meta-proteogenomic Approach to Peptide Identification Incorporating Assembly Uncertainty and Genomic Variation. *Molecular and Cellular Proteomics*, **2019**, 18, S183-S192	7.6	11
42	STRScan: targeted profiling of short tandem repeats in whole-genome sequencing data. *BMC Bioinformatics*, **2017**, 18, 398	3.6	11
41	A de Bruijn graph approach to the quantification of closely-related genomes in a microbial community. *Journal of Computational Biology*, **2012**, 19, 814-25	1.7	11
40	A new estimator of significance of correlation in time series data. *Journal of Computational Biology*, **2001**, 8, 463-70	1.7	11
39	DNA sequence templates adjacent nucleosome and ORC sites at gene amplification origins in Drosophila. *Nucleic Acids Research*, **2015**, 43, 8746-61	20.1	10
38	The Spectrum of Replication Errors in the Absence of Error Correction Assayed Across the Whole Genome of. *Genetics*, **2018**, 209, 1043-1054	4	10
37	msCRUSH: Fast Tandem Mass Spectral Clustering Using Locality Sensitive Hashing. *Journal of Proteome Research*, **2019**, 18, 147-158	5.6	10
36	The Symmetrical Wave Pattern of Base-Pair Substitution Rates across the Escherichia coli Chromosome Has Multiple Causes. *MBio*, **2019**, 10,	7.8	9
35	Bioinformatics protocols in glycomics and glycoproteomics. *Current Protocols in Protein Science*, **2014**, 76, 2.15.1-2.15.7	3.1	9
34	Extending the coverage of spectral libraries: a neighbor-based approach to predicting intensities of peptide fragmentation spectra. *Proteomics*, **2013**, 13, 756-65	4.8	9
33	On the estimation of false positives in peptide identifications using decoy search strategy. *Proteomics*, **2009**, 9, 194-204	4.8	9

32	iDASH secure genome analysis competition 2017. *BMC Medical Genomics*, **2018**, 11, 85	3.7	9
31	Identification of Glycopeptides with Multiple Hydroxylysine O-Glycosylation Sites by Tandem Mass Spectrometry. *Journal of Proteome Research*, **2015**, 14, 5099-108	5.6	8
30	Adaptation of Escherichia coli to long-term batch culture in various rich media. *Research in Microbiology*, **2018**, 169, 145-156	4	8
29	Protein identification problem from a Bayesian point of view. *Statistics and Its Interface*, **2012**, 5, 21-37	0.4	8
28	Impact of Amidination on Peptide Fragmentation and Identification in Shotgun Proteomics. *Journal of Proteome Research*, **2016**, 15, 3656-3665	5.6	7
27	Computational Methods in Mass Spectrometry-Based Proteomics. *Advances in Experimental Medicine and Biology*, **2016**, 939, 63-89	3.6	7
26	Correcting base-assignment errors in repeat regions of shotgun assembly. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **2007**, 4, 54-64	3	6
25	Insertion sequence elements-mediated structural variations in bacterial genomes. *Mobile DNA*, **2018**, 9, 29	4.4	6
24	Clonal reconstruction from time course genomic sequencing data. *BMC Genomics*, **2019**, 20, 1002	4.5	5
23	New complexities of SOS-induced "untargeted" mutagenesis in Escherichia coli as revealed by mutation accumulation and whole-genome sequencing. *DNA Repair*, **2020**, 90, 102852	4.3	4
22	Constrained Sequencing of neo-Epitope Peptides using Tandem Mass Spectrometry. *Lecture Notes in Computer Science*, **2018**, 10812, 138-153	0.9	4
21	Computational identification of micro-structural variations and their proteogenomic consequences in cancer. *Bioinformatics*, **2018**, 34, 1672-1681	7.2	4
20	GlycoHybridSeq: Automated Identification of N-Linked Glycopeptides Using Electron Transfer/High-Energy Collision Dissociation (EThcD). *Journal of Proteome Research*, **2021**, 20, 3345-3352	5.6	4
19	Characterization of microbial associations in human oral microbiome. *Bio-Medical Materials and Engineering*, **2014**, 24, 3737-44	1	3
18	A computational approach to characterizing bond linkages of glycan isomers using matrix-assisted laser desorption/ionization tandem time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, **2008**, 22, 3561-9	2.2	3
17	On the privacy risks of sharing clinical proteomics data. *AMIA Summits on Translational Science Proceedings*, **2016**, 2016, 122-31	1.1	3
16	A Computational Approach for the Identification of Site-Specific Protein Glycosylations Through Ion-Trap Mass Spectrometry. *Lecture Notes in Computer Science*, **2007**, 96-107	0.9	3
15	Identification of N-terminal protein processing sites by chemical labeling mass spectrometry. *Rapid Communications in Mass Spectrometry*, **2019**, 33, 1015-1023	2.2	2

14	A Secure Alignment Algorithm for Mapping Short Reads to Human Genome. *Journal of Computational Biology*, **2018**, 25, 529-540	1.7	2
13	Improving phosphopeptide identification in shotgun proteomics by supervised filtering of peptide-spectrum matches **2013**,		2
12	Secure Genomic Computation through Site-Wise Encryption. *AMIA Summits on Translational Science Proceedings*, **2015**, 2015, 227-31	1.1	2
11	Real-time Protection of Genomic Data Sharing in Beacon Services. *AMIA Summits on Translational Science Proceedings*, **2018**, 2017, 45-54	1.1	2
10	Privacy-preserving construction of generalized linear mixed model for biomedical computation. *Bioinformatics*, **2020**, 36, i128-i135	7.2	2
9	A Maximum-Likelihood Approach to Estimating the Insertion Frequencies of Transposable Elements from Population Sequencing Data. *Molecular Biology and Evolution*, **2018**, 35, 2560-2571	8.3	1
8	Bioinformatic Approaches in Glycomics and Glycoproteomics. *Current Proteomics*, **2011**, 8, 309-324	0.7	1
7	The symmetrical pattern of base-pair substitutions rates across the chromosome in Escherichia coli has multiple causes		1
6	Clonal reconstruction from time course genomic sequencing data		1
5	Haplotype-based membership inference from summary genomic data. *Bioinformatics*, **2021**, 37, i161-i168	7.2	1
4	Algorithmic approaches to clonal reconstruction in heterogeneous cell populations. *Quantitative Biology*, **2019**, 7, 255-265	3.9	1
3	A Fast and Memory-Efficient Spectral Library Search Algorithm Using Locality-Sensitive Hashing. *Proteomics*, **2020**, 20, e2000002	4.8	0
2	GlycanGUI: Automated Glycan Annotation and Quantification Using Glucose Unit Index. *Frontiers in Chemistry*, **2021**, 9, 707382	5	0
1	Overlap detection on long, error-prone sequencing reads via smooth q-gram. *Bioinformatics*, **2020**, 36, 4838-4845	7.2	