

Pavel A Pevzner

List of Publications by Citations

Source: <https://exaly.com/author-pdf/2594793/pavel-a-pevzner-publications-by-citations.pdf>

Version: 2024-04-26

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.

124
papers

26,018
citations

50
h-index

132
g-index

132
ext. papers

37,410
ext. citations

12.2
avg, IF

7.11
L-index

#	Paper	IF	Citations
124	SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. <i>Journal of Computational Biology</i> , 2012 , 19, 455-77	1.7	11958
123	Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. <i>Nature Biotechnology</i> , 2016 , 34, 828-837	44.5	1566
122	metaSPAdes: a new versatile metagenomic assembler. <i>Genome Research</i> , 2017 , 27, 824-834	9.7	1305
121	Assembly of long, error-prone reads using repeat graphs. <i>Nature Biotechnology</i> , 2019 , 37, 540-546	44.5	897
120	Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. <i>Journal of Computational Biology</i> , 2013 , 20, 714-37	1.7	815
119	MS-GF+ makes progress towards a universal database search tool for proteomics. <i>Nature Communications</i> , 2014 , 5, 5277	17.4	633
118	PepNovo: de novo peptide sequencing via probabilistic network modeling. <i>Analytical Chemistry</i> , 2005 , 77, 964-73	7.8	506
117	InsPecT: identification of posttranslationally modified peptides from tandem mass spectra. <i>Analytical Chemistry</i> , 2005 , 77, 4626-39	7.8	502
116	De novo peptide sequencing via tandem mass spectrometry. <i>Journal of Computational Biology</i> , 1999 , 6, 327-42	1.7	434
115	How to apply de Bruijn graphs to genome assembly. <i>Nature Biotechnology</i> , 2011 , 29, 987-91	44.5	341
114	Spectral probabilities and generating functions of tandem mass spectra: a strike against decoy databases. <i>Journal of Proteome Research</i> , 2008 , 7, 3354-63	5.6	341
113	American Gut: an Open Platform for Citizen Science Microbiome Research. <i>MSystems</i> , 2018 , 3,	7.6	336
112	Short read fragment assembly of bacterial genomes. <i>Genome Research</i> , 2008 , 18, 324-30	9.7	332
111	hybridSPAdes: an algorithm for hybrid assembly of short and long reads. <i>Bioinformatics</i> , 2016 , 32, 1009-152	15.2	292
110	Human and mouse genomic sequences reveal extensive breakpoint reuse in mammalian evolution. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2003 , 100, 7672-7	11.5	290
109	Telomere-to-telomere assembly of a complete human X chromosome. <i>Nature</i> , 2020 , 585, 79-84	50.4	269
108	Genome rearrangements in mammalian evolution: lessons from human and mouse genomes. <i>Genome Research</i> , 2003 , 13, 37-45	9.7	244

107	plasmidSPAdes: assembling plasmids from whole genome sequencing data. <i>Bioinformatics</i> , 2016 , 32, 3380-3387	7.2	233
106	Identification of post-translational modifications by blind search of mass spectra. <i>Nature Biotechnology</i> , 2005 , 23, 1562-7	44.5	231
105	De novo fragment assembly with short mate-paired reads: Does the read length matter?. <i>Genome Research</i> , 2009 , 19, 336-46	9.7	199
104	1-Tuple DNA sequencing: computer analysis. <i>Journal of Biomolecular Structure and Dynamics</i> , 1989 , 7, 63-73	3.6	188
103	Clustering millions of tandem mass spectra. <i>Journal of Proteome Research</i> , 2008 , 7, 113-22	5.6	187
102	Ancestral reconstruction of segmental duplications reveals punctuated cores of human genome evolution. <i>Nature Genetics</i> , 2007 , 39, 1361-8	36.3	162
101	De novo repeat classification and fragment assembly. <i>Genome Research</i> , 2004 , 14, 1786-96	9.7	157
100	Imaging mass spectrometry of intraspecies metabolic exchange revealed the cannibalistic factors of <i>Bacillus subtilis</i> . <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2010 , 107, 16286-90	11.5	155
99	Whole-genome analysis of Alu repeat elements reveals complex evolutionary history. <i>Genome Research</i> , 2004 , 14, 2245-52	9.7	146
98	Fragment assembly with short reads. <i>Bioinformatics</i> , 2004 , 20, 2067-74	7.2	134
97	Protein identification by spectral networks analysis. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2007 , 104, 6140-5	11.5	133
96	Dereplication of peptidic natural products through database search of mass spectra. <i>Nature Chemical Biology</i> , 2017 , 13, 30-37	11.7	131
95	Assembly of long error-prone reads using de Bruijn graphs. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2016 , 113, E8396-E8405	11.5	120
94	Automated genome mining of ribosomal peptide natural products. <i>ACS Chemical Biology</i> , 2014 , 9, 1545-51	11.9	114
93	Characterization of cyanobacterial hydrocarbon composition and distribution of biosynthetic pathways. <i>PLoS ONE</i> , 2014 , 9, e85140	3.7	114
92	Dereplication of microbial metabolites through database search of mass spectra. <i>Nature Communications</i> , 2018 , 9, 4035	17.4	109
91	Mutation-tolerant protein identification by mass spectrometry. <i>Journal of Computational Biology</i> , 2000 , 7, 777-87	1.7	107
90	The complete sequence of a human genome.. <i>Science</i> , 2022 , 376, 44-53	33.3	107

89	metaFlye: scalable long-read metagenome assembly using repeat graphs. <i>Nature Methods</i> , 2020 , 17, 1103-1110	21.6	101
88	Automated de novo protein sequencing of monoclonal antibodies. <i>Nature Biotechnology</i> , 2008 , 26, 1336-41	11.5	100
87	ExSPAnDer: a universal repeat resolver for DNA fragment assembly. <i>Bioinformatics</i> , 2014 , 30, i293-301	7.2	80
86	Fragment assembly with double-barreled data. <i>Bioinformatics</i> , 2001 , 17 Suppl 1, S225-33	7.2	79
85	Dereplication and de novo sequencing of nonribosomal peptides. <i>Nature Methods</i> , 2009 , 6, 596-9	21.6	75
84	Shotgun protein sequencing: assembly of peptide tandem mass spectra from mixtures of modified proteins. <i>Molecular and Cellular Proteomics</i> , 2007 , 6, 1123-34	7.6	71
83	Spectral archives: extending spectral libraries to analyze both identified and unidentified spectra. <i>Nature Methods</i> , 2011 , 8, 587-91	21.6	67
82	DRIMM-Synteny: decomposing genomes into evolutionary conserved segments. <i>Bioinformatics</i> , 2010 , 26, 2509-16	7.2	65
81	Interpreting top-down mass spectra using spectral alignment. <i>Analytical Chemistry</i> , 2008 , 80, 2499-505	7.8	65
80	NRPquest: Coupling Mass Spectrometry and Genome Mining for Nonribosomal Peptide Discovery. <i>Journal of Natural Products</i> , 2014 , 77, 1902-9	4.9	64
79	Computing has changed biology--biology education must catch up. <i>Science</i> , 2009 , 325, 541-2	33.3	61
78	The complete sequence of a human genome		58
77	Dereplication, sequencing and identification of peptidic natural products: from genome mining to peptidogenomics to spectral networks. <i>Natural Product Reports</i> , 2016 , 33, 73-86	15.1	54
76	Comment on "Protein sequences from mastodon and Tyrannosaurus rex revealed by mass spectrometry". <i>Science</i> , 2008 , 321, 1040; author reply 1040	33.3	51
75	Paired de bruijn graphs: a novel approach for incorporating mate pair information into genome assemblers. <i>Journal of Computational Biology</i> , 2011 , 18, 1625-34	1.7	50
74	Increased diversity of peptidic natural products revealed by modification-tolerant database search of mass spectra. <i>Nature Microbiology</i> , 2018 , 3, 319-327	26.6	49
73	Plasmid detection and assembly in genomic and metagenomic data sets. <i>Genome Research</i> , 2019 , 29, 961-968	9.7	48
72	SEQuel: improving the accuracy of genome assemblies. <i>Bioinformatics</i> , 2012 , 28, i188-96	7.2	46

71	Telomere-to-telomere assembly of a complete human X chromosome		45
70	Metaviral SPAdes: assembly of viruses from metagenomic data. <i>Bioinformatics</i> , 2020 , 36, 4126-4129	7.2	44
69	Single-molecule protein identification by sub-nanopore sensors. <i>PLoS Computational Biology</i> , 2017 , 13, e1005356	5	40
68	Top-down analysis of immunoglobulin G isotypes 1 and 2 with electron transfer dissociation on a high-field Orbitrap mass spectrometer. <i>Journal of Proteomics</i> , 2017 , 159, 67-76	3.9	36
67	Multiplex de novo sequencing of peptide antibiotics. <i>Journal of Computational Biology</i> , 2011 , 18, 1371-81.7		36
66	Single cells within the Puerto Rico trench suggest hadal adaptation of microbial lineages. <i>Applied and Environmental Microbiology</i> , 2015 , 81, 8265-76	4.8	34
65	Cycloquest: identification of cyclopeptides via database search of their mass spectra against genome databases. <i>Journal of Proteome Research</i> , 2011 , 10, 4505-12	5.6	33
64	TruSPAdes: barcode assembly of TruSeq synthetic long reads. <i>Nature Methods</i> , 2016 , 13, 248-50	21.6	32
63	Sequencing cyclic peptides by multistage mass spectrometry. <i>Proteomics</i> , 2011 , 11, 3642-50	4.8	32
62	De Novo Sequencing of Peptides from Top-Down Tandem Mass Spectra. <i>Journal of Proteome Research</i> , 2015 , 14, 4450-62	5.6	28
61	Assembling short reads from jumping libraries with large insert sizes. <i>Bioinformatics</i> , 2015 , 31, 3262-8	7.2	28
60	Assembly of Long Error-Prone Reads Using Repeat Graphs		28
59	IgRepertoireConstructor: a novel algorithm for antibody repertoire construction and immunoproteogenomics analysis. <i>Bioinformatics</i> , 2015 , 31, i53-61	7.2	27
58	BiosyntheticSPAdes: reconstructing biosynthetic gene clusters from assembly graphs. <i>Genome Research</i> , 2019 , 29, 1352-1362	9.7	26
57	MetaMiner: A Scalable Peptidogenomics Approach for Discovery of Ribosomal Peptide Natural Products with Blind Modifications from Microbial Communities. <i>Cell Systems</i> , 2019 , 9, 600-608.e4	10.6	26
56	Multi-spectra peptide sequencing and its applications to multistage mass spectrometry. <i>Bioinformatics</i> , 2008 , 24, i416-23	7.2	25
55	Optimizing sequencing protocols for leaderboard metagenomics by combining long and short reads. <i>Genome Biology</i> , 2019 , 20, 226	18.3	24
54	Automated assembly of centromeres from ultra-long error-prone reads. <i>Nature Biotechnology</i> , 2020 , 38, 1309-1316	44.5	24

53	Detection and analysis of ancient segmental duplications in mammalian genomes. <i>Genome Research</i> , 2018 , 28, 901-909	9.7	23
52	Advanced Proteogenomic Analysis Reveals Multiple Peptide Mutations and Complex Immunoglobulin Peptides in Colon Cancer. <i>Journal of Proteome Research</i> , 2015 , 14, 3555-67	5.6	22
51	Spatial Molecular Architecture of the Microbial Community of a Lichen. <i>MSystems</i> , 2016 , 1,	7.6	22
50	A family of small, cyclic peptides buried in prealbumin since the Eocene epoch. <i>Plant Direct</i> , 2018 , 2, e00042	3.3	21
49	Reconstructing Antibody Repertoires from Error-Prone Immunosequencing Reads. <i>Journal of Immunology</i> , 2017 , 199, 3369-3380	5.3	21
48	TandemTools: mapping long reads and assessing/improving assembly quality in extra-long tandem repeats. <i>Bioinformatics</i> , 2020 , 36, i75-i83	7.2	21
47	Analysis of metagenome-assembled viral genomes from the human gut reveals diverse putative CrAss-like phages with unique genomic features. <i>Nature Communications</i> , 2021 , 12, 1044	17.4	20
46	Thousands of previously unknown phages discovered in whole-community human gut metagenomes. <i>Microbiome</i> , 2021 , 9, 78	16.6	19
45	Complete genomic and epigenetic maps of human centromeres.. <i>Science</i> , 2022 , 376, eabl4178	33.3	19
44	A new approach to evaluating statistical significance of spectral identifications. <i>Journal of Proteome Research</i> , 2013 , 12, 1560-8	5.6	16
43	De Novo Peptide Sequencing Reveals Many Cyclopeptides in the Human Gut and Other Environments. <i>Cell Systems</i> , 2020 , 10, 99-108.e5	10.6	16
42	Ultralow-input single-tube linked-read library method enables short-read second-generation sequencing systems to routinely generate highly accurate and economical long-range sequencing information. <i>Genome Research</i> , 2020 , 30, 898-909	9.7	15
41	Inference of Diversity Genes and Analysis of Non-canonical V(DD)J Recombination in Immunoglobulins. <i>Frontiers in Immunology</i> , 2019 , 10, 987	8.4	14
40	metaFlye: scalable long-read metagenome assembly using repeat graphs		14
39	Generating lineage-resolved, complete metagenome-assembled genomes from complex microbial communities.. <i>Nature Biotechnology</i> , 2022 ,	44.5	13
38	The string decomposition problem and its applications to centromere analysis and assembly. <i>Bioinformatics</i> , 2020 , 36, i93-i101	7.2	13
37	SpectroGene: A Tool for Proteogenomic Annotations Using Top-Down Spectra. <i>Journal of Proteome Research</i> , 2016 , 15, 144-51	5.6	12
36	Paired de Bruijn Graphs: A Novel Approach for Incorporating Mate Pair Information into Genome Assemblers. <i>Lecture Notes in Computer Science</i> , 2011 , 238-251	0.9	12

35	De novo repeat classification and fragment assembly 2004 ,		12
34	Top-down analysis of protein samples by de novo sequencing techniques. <i>Bioinformatics</i> , 2016 , 32, 2753-92		12
33	How bioinformatics and open data can boost basic science in countries and universities with limited resources. <i>Nature Biotechnology</i> , 2019 , 37, 324-326	44.5	11
32	What is the difference between the breakpoint graph and the de Bruijn graph?. <i>BMC Genomics</i> , 2014 , 15 Suppl 6, S6	4.5	11
31	Pathset graphs: a novel approach for comprehensive utilization of paired reads in genome assembly. <i>Journal of Computational Biology</i> , 2013 , 20, 359-71	1.7	11
30	centroFlye: Assembling Centromeres with Long Error-Prone Reads		9
29	Integrating genomics and metabolomics for scalable non-ribosomal peptide discovery. <i>Nature Communications</i> , 2021 , 12, 3225	17.4	8
28	Complete genomic and epigenetic maps of human centromeres		8
27	From de Bruijn Graphs to Rectangle Graphs for Genome Assembly. <i>Lecture Notes in Computer Science</i> , 2012 , 249-261	0.9	7
26	A Multi-Omics Characterization of the Natural Product Potential of Tropical Filamentous Marine Cyanobacteria. <i>Marine Drugs</i> , 2021 , 19,	6	6
25	CentromereArchitect: inference and analysis of the architecture of centromeres. <i>Bioinformatics</i> , 2021 , 37, i196-i204	7.2	6
24	Joint Analysis of Long and Short Reads Enables Accurate Estimates of Microbiome Complexity. <i>Cell Systems</i> , 2018 , 7, 192-200.e3	10.6	5
23	Manifold de Bruijn Graphs. <i>Lecture Notes in Computer Science</i> , 2014 , 296-310	0.9	5
22	GENOME-WIDE ANALYSIS OF BACTERIAL PROMOTER REGIONS 2002 ,		5
21	mosaicFlye: Resolving long mosaic repeats using long error-prone reads		5
20	Trace Reconstruction Problems in Computational Biology. <i>IEEE Transactions on Information Theory</i> , 2021 , 67, 3295-3314	2.8	5
19	Multiplex de Bruijn graphs enable genome assembly from long, high-fidelity reads.. <i>Nature Biotechnology</i> , 2022 ,	44.5	5
18	Automated analysis of immunosequencing datasets reveals novel immunoglobulin D genes across diverse species. <i>PLoS Computational Biology</i> , 2020 , 16, e1007837	5	4

17	Immunoglobulin Classification Using the Colored Antibody Graph. <i>Journal of Computational Biology</i> , 2016 , 23, 483-94	1.7	4
16	The Antibody Repertoire of Colorectal Cancer. <i>Molecular and Cellular Proteomics</i> , 2017 , 16, 2111-2124	7.6	3
15	Pathset Graphs: A Novel Approach for Comprehensive Utilization of Paired Reads in Genome Assembly. <i>Lecture Notes in Computer Science</i> , 2012 , 200-212	0.9	3
14	TandemMapper and TandemQUAST: mapping long reads and assessing/improving assembly quality in extra-long tandem repeats		3
13	Thousands of previously unknown phages discovered in whole-community human gut metagenomes		3
12	V(DD)J recombination is an important and evolutionarily conserved mechanism for generating antibodies with unusually long CDR3s. <i>Genome Research</i> , 2020 , 30, 1547-1558	9.7	3
11	HORmon: automated annotation of human centromeres		2
10	Generation of lineage-resolved complete metagenome-assembled genomes by precision phasing		2
9	MetaRiPPquest: A Peptidogenomics Approach for the Discovery of Ribosomally Synthesized and Post-translationally Modified Peptides		1
8	MosaicFlye: Resolving Long Mosaic Repeats Using Long Reads. <i>Lecture Notes in Computer Science</i> , 2020 , 226-228	0.9	1
7	ORFograph: search for novel insecticidal protein genes in genomic and metagenomic assembly graphs. <i>Microbiome</i> , 2021 , 9, 149	16.6	1
6	viralFlye: assembling viruses and identifying their hosts from long-read metagenomics data.. <i>Genome Biology</i> , 2022 , 23, 57	18.3	1
5	Automated annotation of human centromeres with HORmon.. <i>Genome Research</i> , 2022 ,	9.7	1
4	Preface: 2nd Satellite Meeting on Bioinformatics Education, Research in Computational Molecular Biology (RECOMB-BE 2010). <i>Journal of Computational Biology</i> , 2011 , 18, 865	1.7	
3	Educating Biologists in the 21st Century: Bioinformatics Scientists versus Bioinformatics Technicians 2007 , 1-5		
2	Are There Rearrangement Hotspots in the Human Genome?. <i>PLoS Computational Biology</i> , 2005 , preprint, e209		5
1	Identifying protease cleavage sites by label-free mass spectrometry. <i>FASEB Journal</i> , 2010 , 24, 905.4	0.9	