Anton Nekrutenko

List of Publications by Year
in descending order

96 papers	22,425 citations	66343 42 h-index	42399 92 g-index
120 all docs	120 docs citations	120 times ranked	32668 citing authors

#	Article	IF	Citations
1	Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biology, 2010, 11, R86.	9.6	3,082
2	The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Research, 2018, 46, W537-W544.	14.5	3,003
3	Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature, 2004, 432, 695-716.	27.8	2,421
4	Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature, 2004, 428, 493-521.	27.8	1,943
5	Galaxy: A platform for interactive large-scale genome analysis. Genome Research, 2005, 15, 1451-1455.	5.5	1,795
6	The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Research, 2016, 44, W3-W10.	14.5	1,751
7	Galaxy: A Web–Based Genome Analysis Tool for Experimentalists. Current Protocols in Molecular Biology, 2010, 89, Unit 19.10.1-21.	2.9	1,159
8	Manipulation of FASTQ data with Galaxy. Bioinformatics, 2010, 26, 1783-1785.	4.1	587
9	Ten Simple Rules for Reproducible Computational Research. PLoS Computational Biology, 2013, 9, e1003285.	3.2	509
10	Evolutionary analyses of the human genome. Nature, 2001, 409, 847-849.	27.8	442
11	Complete Khoisan and Bantu genomes from southern Africa. Nature, 2010, 463, 943-947.	27.8	400
12	Transposable elements are found in a large number of human protein-coding genes. Trends in Genetics, 2001, 17, 619-621.	6.7	383
13	HyPhy 2.5—A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. Molecular Biology and Evolution, 2020, 37, 295-299.	8.9	342
14	The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. Nucleic Acids Research, 2020, 48, W395-W402.	14.5	322
15	Next-generation sequencing data interpretation: enhancing reproducibility and accessibility. Nature Reviews Genetics, 2012, 13, 667-672.	16.3	269
16	28-Way vertebrate alignment and conservation track in the UCSC Genome Browser. Genome Research, 2007, 17, 1797-1808.	5.5	237
17	The KA/KS Ratio Test for Assessing the Protein-Coding Potential of Genomic Regions: An Empirical and Simulation Study. Genome Research, 2002, 12, 198-202.	5.5	233
18	Dissemination of scientific software with Galaxy ToolShed. Genome Biology, 2014, 15, 403.	9.6	205

#	Article	IF	Citations
19	Maternal age effect and severe germ-line bottleneck in the inheritance of human mitochondrial DNA. Proceedings of the National Academy of Sciences of the United States of America, 2014, 111, 15474-15479.	7.1	201
20	Sequencing error profiles of Illumina sequencing instruments. NAR Genomics and Bioinformatics, 2021, 3, lqab019.	3.2	188
21	COMPARATIVE GENOMICS. Annual Review of Genomics and Human Genetics, 2004, 5, 15-56.	6.2	156
22	Community-Driven Data Analysis Training for Biology. Cell Systems, 2018, 6, 752-758.e1.	6.2	141
23	Galaxy CloudMan: delivering cloud compute clusters. BMC Bioinformatics, 2010, 11, S4.	2.6	140
24	A framework for collaborative analysis of ENCODE data: Making large-scale analyses biologist-friendly. Genome Research, 2007, 17, 960-964.	5.5	122
25	Using Galaxy to Perform Large–Scale Interactive Data Analyses. Current Protocols in Bioinformatics, 2007, 19, Unit 10.5.	25.8	100
26	Harnessing cloud computing with Galaxy Cloud. Nature Biotechnology, 2011, 29, 972-974.	17.5	100
27	Practical Computational Reproducibility in the Life Sciences. Cell Systems, 2018, 6, 631-635.	6.2	100
28	Dynamics of mitochondrial heteroplasmy in three families investigated via a repeatable re-sequencing study. Genome Biology, 2011, 12, R59.	8.8	93
29	Densities, length proportions, and other distributional features of repetitive sequences in the human genome estimated from 430 megabases of genomic sequence. Gene, 2000, 259, 81-88.	2.2	91
30	Signatures of Domain Shuffling in the Human Genome. Genome Research, 2002, 12, 1642-1650.	5.5	91
31	Selection Analysis Identifies Clusters of Unusual Mutational Changes in Omicron Lineage BA.1 That Likely Impact Spike Function. Molecular Biology and Evolution, 2022, 39, .	8.9	84
32	Transcriptome of embryonic and neonatal mouse cortex by high-throughput RNA sequencing. Proceedings of the National Academy of Sciences of the United States of America, 2009, 106, 12741-12746.	7.1	72
33	Bottleneck and selection in the germline and maternal age influence transmission of mitochondrial DNA in human pedigrees. Proceedings of the National Academy of Sciences of the United States of America, 2019, 116, 25172-25178.	7.1	71
34	A First Look at ARFome: Dual-Coding Genes in Mammalian Genomes. PLoS Computational Biology, 2007, 3, e91.	3.2	68
35	Windshield splatter analysis with the Galaxy metagenomic pipeline. Genome Research, 2009, 19, 2144-2153.	5.5	68
36	Inverting the model of genomics data sharing with the NHGRI Genomic Data Science Analysis, Visualization, and Informatics Lab-space. Cell Genomics, 2022, 2, 100085.	6.5	59

#	Article	IF	Citations
37	Evolutionary Dynamics of Oncogenes and Tumor Suppressor Genes: Higher Intensities of Purifying Selection than Other Genes. Molecular Biology and Evolution, 2003, 20, 964-968.	8.9	57
38	Jupyter and Galaxy: Easing entry barriers into complex data analyses for biomedical researchers. PLoS Computational Biology, 2017, 13, e1005425.	3.2	53
39	Evolution of Microsatellite Alleles in Four Species of Mice (Genus Apodemus). Journal of Molecular Evolution, 2000, 51, 166-172.	1.8	51
40	A Massively Parallel Sequencing Approach Uncovers Ancient Origins and High Genetic Variability of Endangered Przewalski's Horses. Genome Biology and Evolution, 2011, 3, 1096-1106.	2.5	51
41	RNA–DNA differences in human mitochondria restore ancestral form of 16S ribosomal RNA. Genome Research, 2013, 23, 1789-1796.	5.5	49
42	Integrating diverse databases into an unified analysis framework: a Galaxy approach. Database: the Journal of Biological Databases and Curation, 2011, 2011, bar011-bar011.	3.0	46
43	Rapid and asymmetric divergence of duplicate genes in the human gene coexpression network. BMC Bioinformatics, 2006, 7, 46.	2.6	45
44	Web-based visual analysis for high-throughput genomics. BMC Genomics, 2013, 14, 397.	2.8	45
45	Integrative genomic analysis by interoperation of bioinformatics tools in GenomeSpace. Nature Methods, 2016, 13, 245-247.	19.0	44
46	Making whole genome multiple alignments usable for biologists. Bioinformatics, 2011, 27, 2426-2428.	4.1	43
47	StructureFold: genome-wide RNA secondary structure mapping and reconstruction *in vivo*. Bioinformatics, 2015, 31, 2668-2675.	4.1	43
48	Oscillating Evolution of a Mammalian Locus with Overlapping Reading Frames: An XLÎ±s/ALEX Relay. PLoS Genetics, 2005, 1, e18.	3.5	39
49	Child Weight Gain Trajectories Linked To Oral Microbiota Composition. Scientific Reports, 2018, 8, 14030.	3.3	39
50	Using Galaxy to Perform Large-Scale Interactive Data Analyses. Current Protocols in Bioinformatics, 2012, 38, Unit10.5.	25.8	36
51	An evolutionary approach reveals a high protein-coding capacity of the human genome. Trends in Genetics, 2003, 19, 306-310.	6.7	33
52	Functionality of unspliced XBP1 is required to explain evolution of overlapping reading frames. Trends in Genetics, 2006, 22, 645-648.	6.7	31
53	Wrangling Galaxy's reference data. Bioinformatics, 2014, 30, 1917-1919.	4.1	31
54	Bushbaby Growth Hormone Is Much More Similar to Nonprimate Growth Hormones than to Rhesus Monkey and Human Growth Hormones. Molecular Biology and Evolution, 2001, 18, 55-60.	8.9	29

#	Article	IF	Citations
55	NGS analyses by visualization with Trackster. Nature Biotechnology, 2012, 30, 1036-1039.	17.5	26
56	The anatomy of successful computational biology software. Nature Biotechnology, 2013, 31, 894-897.	17.5	25
57	Streamlined analysis of duplex sequencing data with Du Novo. Genome Biology, 2016, 17, 180.	8.8	24
58	Rapid asymmetric evolution of a dual-coding tumor suppressor INK4a/ARF locus contradicts its function. Proceedings of the National Academy of Sciences of the United States of America, 2007, 104, 12807-12812.	7.1	22
59	Controlling for contamination in re-sequencing studies with a reproducible web-based phylogenetic approach. BioTechniques, 2014, 56, 134-141.	1.8	22
60	No more business as usual: Agile and effective responses to emerging pathogen threats require open data and open analytics. PLoS Pathogens, 2020, 16, e1008643.	4.7	22
61	Detection of SARS-CoV-2 intra-host recombination during superinfection with Alpha and Epsilon variants in New York City. Nature Communications, 2022, 13, .	12.8	22
62	Detection of gene duplications and block duplications in eukaryotic genomes. Journal of Structural and Functional Genomics, 2003, 3, 27-34.	1.2	21
63	Ready-to-use public infrastructure for global SARS-CoV-2 monitoring. Nature Biotechnology, 2021, 39, 1178-1179.	17.5	21
64	Galaxy: A Gateway to Tools in e-Science. Computer Communications and Networks, 2011, , 145-177.	0.8	20
65	ETOPE: evolutionary test of predicted exons. Nucleic Acids Research, 2003, 31, 3564-3567.	14.5	17
66	Reconciling the Numbers: ESTs Versus Protein-Coding Genes. Molecular Biology and Evolution, 2004, 21, 1278-1282.	8.9	17
67	Representational difference analysis to distinguish cryptic species. Molecular Ecology, 1999, 8, 1235-1237.	3.9	16
68	Fostering accessible online education using Galaxy as an e-learning platform. PLoS Computational Biology, 2021, 17, e1008923.	3.2	15
69	Assessment of Compositional Heterogeneity Within and Between Eukaryotic Genomes. Genome Research, 2000, 10, 1986-1995.	5.5	15
70	A single-cell RNA-sequencing training and analysis suite using the Galaxy framework. GigaScience, 2020, 9, .	6.4	14
71	Identification of Novel Exons from Ratâ€"Mouse Comparisons. Journal of Molecular Evolution, 2004, 59, 703-708.	1.8	12
72	Detection of gene duplications and block duplications in eukaryotic genomes. Journal of Structural and Functional Genomics, 2003, 3, 27-34.	1.2	12

#	ARTICLE	IF	CITATIONS
73	Isolation of binary species-specific PCR-based markers and their value for diagnostic applications. Gene, 2000, 249, 47-51.	2.2	11
74	A High-Resolution View of Adaptive Event Dynamics in a Plasmid. Genome Biology and Evolution, 2019, 11, 3022-3034.	2.5	11
75	Predicting runtimes of bioinformatics tools based on historical data: five years of Galaxy usage. Bioinformatics, 2019, 35, 3453-3460.	4.1	11
76	Subgenome-specific markers in allopolyploid cotton Gossypium hirsutum: implications for evolutionary analysis of polyploids. Gene, 2003, 306, 99-103.	2.2	10
77	High-Resolution Mapping of Evolutionary Trajectories in a Phage. Genome Biology and Evolution, 2009, 1, 294-307.	2.5	10
78	Family reunion via error correction: an efficient analysis of duplex sequencing data. BMC Bioinformatics, 2020, 21, 96.	2.6	10
79	Web-Based Analysis of (Epi-) Genome Data Using EpiGRAPH and Galaxy. Methods in Molecular Biology, 2010, 628, 275-296.	0.9	10
80	Wheels within Wheels: Clues to the Evolution of the Gnas and Gnal Loci. Molecular Biology and Evolution, 2008, 25, 2745-2757.	8.9	9
81	Online Resources for Genomic Analysis Using High-Throughput Sequencing. Cold Spring Harbor Protocols, 2015, 2015, pdb.top083667.	0.3	6
82	Biology Needs Evolutionary Software Tools: Let's Build Them Right. Molecular Biology and Evolution, 2018, 35, 1372-1375.	8.9	6
83	Increased yields of duplex sequencing data by a series of quality control tools. NAR Genomics and Bioinformatics, 2021, 3, lqab002.	3.2	6
84	Stepwise Evolution and Exceptional Conservation of ORF1a/b Overlap in Coronaviruses. Molecular Biology and Evolution, 2021, 38, 5678-5684.	8.9	6
85	Lessons learned from Galaxy, a Web-based platform for high-throughput genomic analyses. , 2012, , .		5
86	A reference model for deploying applications in virtualized environments. Concurrency Computation Practice and Experience, 2012, 24, 1349-1361.	2.2	4
87	Using Galaxy to Perform Large-Scale Interactive Data Analyses—An Update. Current Protocols, 2021, 1, e31.	2.9	4
88	Representational difference analysis to distinguish cryptic species. Molecular Ecology, 1999, 8, 1235.	3.9	3
89	A guide and best practices for R/Bioconductor tool integration in Galaxy. F1000Research, 0, 5, 2757.	1.6	3
90	mNSC1 shows no evidence of protein-coding capacity. Gene, 2006, 370, 83-85.	2.2	2

#	Article	IF	Citations
91	In memory of James Taylor: the birth of Galaxy. Genome Biology, 2020, 21, 105.	8.8	1
92	Reproducible and accessible analysis of transposon insertion sequencing in Galaxy for qualitative essentiality analyses. BMC Microbiology, 2021, 21, 168.	3.3	1
93	GYAN: Accelerating Bioinformatics Tools in Galaxy with GPU-Aware Computation Mapping. , 2021, , .		1
94	Detection of gene duplications and block duplications in eukaryotic genomes. , 2003, , 27-34.		1
95	A sustainable national gateway for biological computation. , 2013, , .		0
96	Toward the commoditization of translational genomic research: Design and implementation features of the Galaxy genomic workbench. Summit on Translational Bioinformatics, 2008, 2008, 56-60.	0.7	0