Raja Mazumder

List of Publications by Citations

75 papers	7,972 citations	28 h-index	89 g-index
91 ext. papers	9,303 ext. citations	6.6 avg, IF	5.05 L-index

#	Paper	IF	Citations
75	The COG database: an updated version includes eukaryotes. *BMC Bioinformatics*, **2003**, 4, 41	3.6	3212
74	The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Research*, **2006**, 34, D187-91	20.1	839
73	UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics*, **2007**, 23, 1282-8	7.2	788
72	CDD: a curated Entrez database of conserved domain alignments. *Nucleic Acids Research*, **2003**, 31, 383-7	20.1	620
71	A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biology*, **2004**, 5, R7	18.3	602
70	PIRSF: family classification system at the Protein Information Resource. *Nucleic Acids Research*, **2004**, 32, D112-4	20.1	165
69	Updates to the Symbol Nomenclature for Glycans guidelines. *Glycobiology*, **2019**, 29, 620-624	5.8	148
68	MMDB: Entrez's 3D-structure database. *Nucleic Acids Research*, **2003**, 31, 474-7	20.1	128
67	Detection of novel members, structure-function analysis and evolutionary classification of the 2H phosphoesterase superfamily. *Nucleic Acids Research*, **2002**, 30, 5229-43	20.1	117
66	CoreGenes: a computational tool for identifying and cataloging "core" genes in a set of small genomes. *BMC Bioinformatics*, **2002**, 3, 12	3.6	105
65	Analysis of Genomes and Transcriptomes of Hepatocellular Carcinomas Identifies Mutations and Gene Expression Changes in the Transforming Growth Factor-Pathway. *Gastroenterology*, **2018**, 154, 195-210	13.3	68
64	Representative proteomes: a stable, scalable and unbiased proteome set for sequence analysis and functional annotation. *PLoS ONE*, **2011**, 6, e18910	3.7	67
63	BioXpress: an integrated RNA-seq-derived gene expression database for pan-cancer analysis. *Database: the Journal of Biological Databases and Curation*, **2015**, 2015,	5	60
62	Baseline human gut microbiota profile in healthy people and standard reporting template. *PLoS ONE*, **2019**, 14, e0206484	3.7	59
61	GlyGen: Computational and Informatics Resources for Glycoscience. *Glycobiology*, **2020**, 30, 72-73	5.8	53
60	A framework for organizing cancer-related variations from existing databases, publications and NGS data using a High-performance Integrated Virtual Environment (HIVE). *Database: the Journal of Biological Databases and Curation*, **2014**, 2014, bau022	5	52
59	A comprehensive protein-centric ID mapping service for molecular data integration. *Bioinformatics*, **2011**, 27, 1190-1	7.2	51

58	BioMuta and BioXpress: mutation and expression knowledgebases for cancer biomarker discovery. *Nucleic Acids Research*, **2018**, 46, D1128-D1136	20.1	49
57	Determining chemotactic responses by two subsurface microaerophiles using a simplified capillary assay method. *Journal of Microbiological Methods*, **1999**, 37, 255-63	2.8	49
56	Enhancement of Fe(III), Co(III), and Cr(VI) reduction at elevated temperatures and by a thermophilic bacterium. *Applied Biochemistry and Biotechnology*, **1996**, 57-58, 923-932	3.2	48
55	High-Performance Integrated Virtual Environment (HIVE) Tools and Applications for Big Data Analysis. *Genes*, **2014**, 5, 957-81	4.2	44
54	Structure-based comparative analysis and prediction of N-linked glycosylation sites in evolutionarily distant eukaryotes. *Genomics, Proteomics and Bioinformatics*, **2013**, 11, 96-104	6.5	37
53	DiMeX: A Text Mining System for Mutation-Disease Association Extraction. *PLoS ONE*, **2016**, 11, e0152725	3.7	36
52	High-performance integrated virtual environment (HIVE): a robust infrastructure for next-generation sequence data analysis. *Database: the Journal of Biological Databases and Curation*, **2016**, 2016,	5	36
51	Computational analysis and identification of amino acid sites in dengue E proteins relevant to development of diagnostics and vaccines. *Virus Genes*, **2007**, 35, 175-86	2.3	33
50	Generating a focused view of disease ontology cancer terms for pan-cancer data integration and analysis. *Database: the Journal of Biological Databases and Curation*, **2015**, 2015, bav032	5	32
49	Proteome-wide analysis of single-nucleotide variations in the N-glycosylation sequon of human genes. *PLoS ONE*, **2012**, 7, e36212	3.7	30
48	HIVE-hexagon: high-performance, parallelized sequence alignment for next-generation sequencing data analysis. *PLoS ONE*, **2014**, 9, e99033	3.7	29
47	GeneOrder: comparing the order of genes in small genomes. *Bioinformatics*, **2001**, 17, 162-6	7.2	22
46	Mutated CEACAMs Disrupt Transforming Growth Factor Beta Signaling and Alter the Intestinal Microbiome to Promote Colorectal Carcinogenesis. *Gastroenterology*, **2020**, 158, 238-252	13.3	21
45	Human germline and pan-cancer variomes and their distinct functional profiles. *Nucleic Acids Research*, **2014**, 42, 11570-88	20.1	20
44	Identification of key differentially expressed MicroRNAs in cancer patients through pan-cancer analysis. *Computers in Biology and Medicine*, **2018**, 103, 183-197	7	19
43	Biocompute Objects-A Step towards Evaluation and Validation of Biomedical Scientific Computations. *PDA Journal of Pharmaceutical Science and Technology*, **2017**, 71, 136-146	0.6	17
42	Community annotation in biology. *Biology Direct*, **2010**, 5, 12	7.2	17
41	Systems integration of biodefense omics data for analysis of pathogen-host interactions and identification of potential targets. *PLoS ONE*, **2009**, 4, e7162	3.7	16

40	Enabling precision medicine via standard communication of HTS provenance, analysis, and results. *PLoS Biology*, **2018**, 16, e3000099	9.7	16
39	Census-based rapid and accurate metagenome taxonomic profiling. *BMC Genomics*, **2014**, 15, 918	4.5	15
38	Computational identification of strain-, species- and genus-specific proteins. *BMC Bioinformatics*, **2005**, 6, 279	3.6	15
37	Low-substrate regulated microaerophilic behavior as a stress response of aquatic and soil bacteria. *Current Microbiology*, **2000**, 41, 79-83	2.4	15
36	The GlySpace Alliance: toward a collaborative global glycoinformatics community. *Glycobiology*, **2020**, 30, 70-71	5.8	15
35	Whole genome single-nucleotide variation profile-based phylogenetic tree building methods for analysis of viral, bacterial and human genomes. *Genomics*, **2014**, 104, 1-7	4.3	14
34	OncoMX: A Knowledgebase for Exploring Cancer Biomarkers in the Context of Related Cancer and Healthy Data. *JCO Clinical Cancer Informatics*, **2020**, 4, 210-220	5.2	13
33	GeneOrder3.0: software for comparing the order of genes in pairs of small bacterial genomes. *BMC Bioinformatics*, **2004**, 5, 52	3.6	13
32	Structure-function analysis of hepatitis C virus envelope glycoproteins E1 and E2. *Journal of Biomolecular Structure and Dynamics*, **2015**, 33, 1682-94	3.6	12
31	Loss and gain of N-linked glycosylation sequons due to single-nucleotide variation in cancer. *Scientific Reports*, **2018**, 8, 4322	4.9	12
30	SNVDis: a proteome-wide analysis service for evaluating nsSNVs in protein functional sites and pathways. *Genomics, Proteomics and Bioinformatics*, **2013**, 11, 122-6	6.5	12
29	Sequence signatures in envelope protein may determine whether flaviviruses produce hemorrhagic or encephalitic syndromes. *Virus Genes*, **2009**, 39, 1-9	2.3	12
28	Non-synonymous variations in cancer and their effects on the human proteome: workflow for NGS data biocuration and proteome-wide analysis of TCGA data. *BMC Bioinformatics*, **2014**, 15, 28	3.6	11
27	GlyGen data model and processing workflow. *Bioinformatics*, **2020**, 36, 3941-3943	7.2	10
26	Proteome-wide analysis of nonsynonymous single-nucleotide variations in active sites of human proteins. *FEBS Journal*, **2013**, 280, 1542-62	5.7	10
25	Comparisons of gene colinearity in genomes using GeneOrder2.0. *Trends in Biochemical Sciences*, **2001**, 26, 514-6	10.3	10
24	Single-nucleotide variations in cardiac arrhythmias: prospects for genomics and proteomics based biomarker discovery and diagnostics. *Genes*, **2014**, 5, 254-69	4.2	9
23	Separation and assembly of deep sequencing data into discrete sub-population genomes. *Nucleic Acids Research*, **2017**, 45, 10989-11003	20.1	7

22	DEXTER: Disease-Expression Relation Extraction from Text. *Database: the Journal of Biological Databases and Curation*, **2018**, 2018,	5	7
21	Structure-guided comparative analysis of proteins: principles, tools, and applications for predicting function. *PLoS Computational Biology*, **2008**, 4, e1000151	5	7
20	Germ line variants of human N-methylpurine DNA glycosylase show impaired DNA repair activity and facilitate 1,N6-ethenoadenine-induced mutations. *Journal of Biological Chemistry*, **2015**, 290, 4966-4980	5.4	6
19	Investigation of somatic single nucleotide variations in human endogenous retrovirus elements and their potential association with cancer. *PLoS ONE*, **2019**, 14, e0213770	3.7	4
18	Bioinformatics and machine learning in gastrointestinal microbiome research and clinical application. *Progress in Molecular Biology and Translational Science*, **2020**, 176, 141-178	4	4
17	A Primer for Access to Repositories of Cancer-Related Genomic Big Data. *Methods in Molecular Biology*, **2019**, 1878, 1-37	1.4	3
16	Computational clustering for viral reference proteomes. *Bioinformatics*, **2016**, 32, 2041-3	7.2	2
15	COVID-19 Biomarkers in research: Extension of the OncoMX cancer biomarker data model to capture biomarker data from other diseases **2020**,		2
14	Impact of Nonsynonymous Single-Nucleotide Variations on Post-Translational Modification Sites in Human Proteins. *Methods in Molecular Biology*, **2017**, 1558, 159-190	1.4	2
13	Protein functional annotation by homology. *Methods in Molecular Biology*, **2008**, 484, 465-90	1.4	2
12	COVID-19 biomarkers and their overlap with comorbidities in a disease biomarker data model. *Briefings in Bioinformatics*, **2021**, 22,	13.4	2
11	Mice with dysfunctional TGF-β signaling develop altered intestinal microbiome and colorectal cancer resistant to 5FU. *Biochimica Et Biophysica Acta - Molecular Basis of Disease*, **2021**, 1867, 166179	6.9	2
10	Bioinformatics tools developed to support BioCompute Objects. *Database: the Journal of Biological Databases and Curation*, **2021**, 2021,	5	2
9	Higher levels of Bifidobacteria and tumor necrosis factor in children with drug-resistant epilepsy are associated with anti-seizure response to the ketogenic diet. *EBioMedicine*, **2022**, 80, 104061	8.8	2
8	A framework for application of metabolic modeling in yeast to predict the effects of nsSNV in human orthologs. *Biology Direct*, **2014**, 9, 9	7.2	1
7	Enhancing the interoperability of glycan data flow between ChEBI, PubChem, and GlyGen. *Glycobiology*, **2021**,	5.8	1
6	Distribution bias analysis of germline and somatic single-nucleotide variations that impact protein functional site and neighboring amino acids. *Scientific Reports*, **2017**, 7, 42169	4.9	0
5	miCloud: A Plug-n-Play, Extensible, On-Premises Bioinformatics Cloud for Seamless Execution of Complex Next-Generation Sequencing Data Analysis Pipelines. *Journal of Computational Biology*, **2019**, 26, 280-284	1.7	0

4   Glycoinformatics Resources Integrated Through the GlySpace Alliance **2021**, 507-521   0

3   Application of global computational tools GeneOrder and CoreGenes to the comparative analyses of chordopoxvirus genomes. *Information Sciences*, **2002**, 146, 127-135   7.7

2   Non-synonymous Single-Nucleotide Variations as Cardiovascular System Disease Biomarkers and Their Roles in Bridging Genomic and Proteomic Technologies **2015**, 1-27

1   Nonsynonymous Single-Nucleotide Variations as Cardiovascular System Disease Biomarkers and Their Roles in Bridging Genomic and Proteomic Technologies **2016**, 821-847