

# Robert D. Finn

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/1737471/publications.pdf>

Version: 2024-02-01

198  
papers

110,975  
citations

5876

81  
h-index

2439

197  
g-index

226  
all docs

226  
docs citations

226  
times ranked

105633  
citing authors

#	ARTICLE	IF	CITATIONS
1	Initial sequencing and analysis of the human genome. <i>Nature</i> , 2001, 409, 860-921.	13.7	21,074
2	The Pfam protein families database. <i>Nucleic Acids Research</i> , 2007, 36, D281-D288.	6.5	6,372
3	Pfam: the protein families database. <i>Nucleic Acids Research</i> , 2014, 42, D222-D230.	6.5	5,425
4	The Pfam protein families database: towards a more sustainable future. <i>Nucleic Acids Research</i> , 2016, 44, D279-D285.	6.5	5,391
5	UniProt: the universal protein knowledgebase in 2021. <i>Nucleic Acids Research</i> , 2021, 49, D480-D489.	6.5	4,709
6	HMMER web server: interactive sequence similarity searching. <i>Nucleic Acids Research</i> , 2011, 39, W29-W37.	6.5	4,492
7	The Pfam protein families database in 2019. <i>Nucleic Acids Research</i> , 2019, 47, D427-D432.	6.5	3,937
8	The EMBL-EBI search and sequence analysis tools APIs in 2019. <i>Nucleic Acids Research</i> , 2019, 47, W636-W641.	6.5	3,820
9	The Pfam protein families database. <i>Nucleic Acids Research</i> , 2012, 40, D290-D301.	6.5	3,306
10	The Pfam protein families database. <i>Nucleic Acids Research</i> , 2004, 32, 138D-141.	6.5	3,084
11	Pfam: The protein families database in 2021. <i>Nucleic Acids Research</i> , 2021, 49, D412-D419.	6.5	3,068
12	The Pfam protein families database. <i>Nucleic Acids Research</i> , 2010, 38, D211-D222.	6.5	2,693
13	The Pfam Protein Families Database. <i>Nucleic Acids Research</i> , 2002, 30, 276-280.	6.5	2,067
14	Pfam: clans, web tools and services. <i>Nucleic Acids Research</i> , 2006, 34, D247-D251.	6.5	2,030
15	Highly accurate protein structure prediction for the human proteome. <i>Nature</i> , 2021, 596, 590-596.	13.7	1,773
16	InterPro: the integrative protein signature database. <i>Nucleic Acids Research</i> , 2009, 37, D211-D215.	6.5	1,712
17	HMMER web server: 2018 update. <i>Nucleic Acids Research</i> , 2018, 46, W200-W204.	6.5	1,432
18	The InterPro protein families and domains database: 20 years on. <i>Nucleic Acids Research</i> , 2021, 49, D344-D354.	6.5	1,385

#	ARTICLE	IF	CITATIONS
19	InterPro in 2017â€”beyond protein family and domain annotations. <i>Nucleic Acids Research</i> , 2017, 45, D190-D199.	6.5	1,358
20	Rfam: an RNA family database. <i>Nucleic Acids Research</i> , 2003, 31, 439-441.	6.5	1,296
21	InterPro in 2019: improving coverage, classification and access to protein sequence annotations. <i>Nucleic Acids Research</i> , 2019, 47, D351-D360.	6.5	1,291
22	Rfam: annotating non-coding RNAs in complete genomes. <i>Nucleic Acids Research</i> , 2004, 33, D121-D124.	6.5	1,255
23	The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. <i>Nucleic Acids Research</i> , 2018, 46, D624-D632.	6.5	1,234
24	Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. <i>Nucleic Acids Research</i> , 2013, 41, e121-e121.	6.5	1,214
25	The InterPro protein families database: the classification resource after 15 years. <i>Nucleic Acids Research</i> , 2015, 43, D213-D221.	6.5	1,205
26	The Pfam Protein Families Database. <i>Nucleic Acids Research</i> , 2000, 28, 263-266.	6.5	1,173
27	Rfam 12.0: updates to the RNA families database. <i>Nucleic Acids Research</i> , 2015, 43, D130-D137.	6.5	1,000
28	InterPro in 2011: new developments in the family and domain prediction database. <i>Nucleic Acids Research</i> , 2012, 40, D306-D312.	6.5	921
29	A new genomic blueprint of the human gut microbiota. <i>Nature</i> , 2019, 568, 499-504.	13.7	901
30	HMMER web server: 2015 update. <i>Nucleic Acids Research</i> , 2015, 43, W30-W38.	6.5	849
31	Rfam: updates to the RNA families database. <i>Nucleic Acids Research</i> , 2009, 37, D136-D140.	6.5	820
32	Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. <i>Nucleic Acids Research</i> , 2018, 46, D335-D342.	6.5	819
33	MEROPS: the peptidase database. <i>Nucleic Acids Research</i> , 2010, 38, D227-D233.	6.5	786
34	<i>MEROPS</i>: the database of proteolytic enzymes, their substrates and inhibitors. <i>Nucleic Acids Research</i> , 2014, 42, D503-D509.	6.5	782
35	Rfam 11.0: 10 years of RNA families. <i>Nucleic Acids Research</i> , 2013, 41, D226-D232.	6.5	745
36	Twenty years of the <i>MEROPS</i> database of proteolytic enzymes, their substrates and inhibitors. <i>Nucleic Acids Research</i> , 2016, 44, D343-D350.	6.5	648

#	ARTICLE	IF	CITATIONS
37	A unified catalog of 204,938 reference genomes from the human gut microbiome. <i>Nature Biotechnology</i> , 2021, 39, 105-114.	9.4	628
38	The Dfam database of repetitive DNA families. <i>Nucleic Acids Research</i> , 2016, 44, D81-D89.	6.5	543
39	Single-particle electron cryo-microscopy: towards atomic resolution. <i>Quarterly Reviews of Biophysics</i> , 2000, 33, 307-369.	2.4	535
40	Ensembl Genomes 2018: an integrated omics infrastructure for non-vertebrate species. <i>Nucleic Acids Research</i> , 2018, 46, D802-D808.	6.5	489
41	InterPro, progress and status in 2005. <i>Nucleic Acids Research</i> , 2004, 33, D201-D205.	6.5	478
42	Rfam 14: expanded coverage of metagenomic, viral and microRNA families. <i>Nucleic Acids Research</i> , 2021, 49, D192-D200.	6.5	475
43	New developments in the InterPro database. <i>Nucleic Acids Research</i> , 2007, 35, D224-D228.	6.5	444
44	A human gut bacterial genome and culture collection for improved metagenomic analyses. <i>Nature Biotechnology</i> , 2019, 37, 186-192.	9.4	420
45	Rfam: Wikipedia, clans and the "decimal" release. <i>Nucleic Acids Research</i> , 2011, 39, D141-D145.	6.5	355
46	Massive expansion of human gut bacteriophage diversity. <i>Cell</i> , 2021, 184, 1098-1109.e9.	13.5	331
47	Non-coding RNA Analysis Using the Rfam Database. <i>Current Protocols in Bioinformatics</i> , 2018, 62, e51.	25.8	309
48	Skyline: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. <i>BMC Bioinformatics</i> , 2014, 15, 7.	1.2	298
49	Microbial abundance, activity and population genomic profiling with mOTUs2. <i>Nature Communications</i> , 2019, 10, 1014.	5.8	298
50	MGnify: the microbiome analysis resource in 2020. <i>Nucleic Acids Research</i> , 2020, 48, D570-D578.	6.5	296
51	iPfam: visualization of protein-protein interactions in PDB at domain and amino acid resolutions. <i>Bioinformatics</i> , 2005, 21, 410-412.	1.8	295
52	The Human RNA-Binding Proteome and Its Dynamics during Translational Arrest. <i>Cell</i> , 2019, 176, 391-403.e19.	13.5	289
53	The Systematic Functional Analysis of Plasmodium Protein Kinases Identifies Essential Regulators of Mosquito Transmission. <i>Cell Host and Microbe</i> , 2010, 8, 377-387.	5.1	267
54	QuickTree: building huge Neighbour-Joining trees of protein sequences. <i>Bioinformatics</i> , 2002, 18, 1546-1547.	1.8	249

#	ARTICLE	IF	CITATIONS
55	Dfam: a database of repetitive DNA based on profile hidden Markov models. <i>Nucleic Acids Research</i> , 2012, 41, D70-D82.	6.5	243
56	Predicting active site residue annotations in the Pfam database. <i>BMC Bioinformatics</i> , 2007, 8, 298.	1.2	239
57	Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases. <i>Nucleic Acids Research</i> , 2019, 47, 10994-11006.	6.5	236
58	The PASTA domain: a $\beta$ -lactam-binding domain. <i>Trends in Biochemical Sciences</i> , 2002, 27, 438-440.	3.7	204
59	DUFs: families in search of function. <i>Acta Crystallographica Section F: Structural Biology Communications</i> , 2010, 66, 1148-1152.	0.7	203
60	EBI Metagenomics in 2017: enriching the analysis of microbial communities, from sequence reads to assemblies. <i>Nucleic Acids Research</i> , 2018, 46, D726-D735.	6.5	175
61	RNAcentral: a comprehensive database of non-coding RNA sequences. <i>Nucleic Acids Research</i> , 2017, 45, D128-D134.	6.5	174
62	Reporting guidelines for human microbiome research: the STORMS checklist. <i>Nature Medicine</i> , 2021, 27, 1885-1892.	15.2	170
63	Alternative splicing of intrinsically disordered regions and rewiring of protein interactions. <i>Current Opinion in Structural Biology</i> , 2013, 23, 443-450.	2.6	166
64	The Hotdog fold: wrapping up a superfamily of thioesterases and dehydratases. <i>BMC Bioinformatics</i> , 2004, 5, 109.	1.2	158
65	iPfam: a database of protein family and domain interactions found in the Protein Data Bank. <i>Nucleic Acids Research</i> , 2014, 42, D364-D373.	6.5	156
66	RNAcentral: a hub of information for non-coding RNA sequences. <i>Nucleic Acids Research</i> , 2019, 47, D221-D229.	6.5	153
67	Cache Domains That are Homologous to, but Different from PAS Domains Comprise the Largest Superfamily of Extracellular Sensors in Prokaryotes. <i>PLoS Computational Biology</i> , 2016, 12, e1004862.	1.5	147
68	Ensembl Genomes 2022: an expanding genome resource for non-vertebrates. <i>Nucleic Acids Research</i> , 2022, 50, D996-D1003.	6.5	141
69	Using deep learning to annotate the protein universe. <i>Nature Biotechnology</i> , 2022, 40, 932-937.	9.4	133
70	TreeFam v9: a new website, more species and orthology-on-the-fly. <i>Nucleic Acids Research</i> , 2014, 42, D922-D925.	6.5	125
71	The HMMER Web Server for Protein Sequence Similarity Search. <i>Current Protocols in Bioinformatics</i> , 2017, 60, 3.15.1-3.15.23.	25.8	119
72	Pfam 10 years on: 10 000 families and still growing. <i>Briefings in Bioinformatics</i> , 2008, 9, 210-219.	3.2	114

#	ARTICLE	IF	CITATIONS
73	Protein interactions in human genetic diseases. <i>Genome Biology</i> , 2008, 9, R9.	13.9	110
74	Computational strategies to combat COVID-19: useful tools to accelerate SARS-CoV-2 and coronavirus research. <i>Briefings in Bioinformatics</i> , 2021, 22, 642-663.	3.2	110
75	A comparison of dense transposon insertion libraries in the <i>Salmonella</i> serovars Typhi and Typhimurium. <i>Nucleic Acids Research</i> , 2013, 41, 4549-4564.	6.5	108
76	The European Bioinformatics Institute in 2016: Data growth and integration. <i>Nucleic Acids Research</i> , 2016, 44, D20-D26.	6.5	108
77	The evolution of protein domain families. <i>Biochemical Society Transactions</i> , 2009, 37, 751-755.	1.6	106
78	RNAcentral: an international database of ncRNA sequences. <i>Nucleic Acids Research</i> , 2015, 43, D123-D129.	6.5	103
79	New knowledge from old: in silico discovery of novel protein domains in <i>Streptomyces coelicolor</i> . <i>BMC Microbiology</i> , 2003, 3, 3.	1.3	98
80	EBI metagenomics in 2016 - an expanding and evolving resource for the analysis and archiving of metagenomic data. <i>Nucleic Acids Research</i> , 2016, 44, D595-D603.	6.5	97
81	Benchmarking taxonomic assignments based on 16S rRNA gene profiling of the microbiota from commonly sampled environments. <i>GigaScience</i> , 2018, 7, .	3.3	95
82	Representative Proteomes: A Stable, Scalable and Unbiased Proteome Set for Sequence Analysis and Functional Annotation. <i>PLoS ONE</i> , 2011, 6, e18910.	1.1	94
83	Quantifying the mechanisms of domain gain in animal proteins. <i>Genome Biology</i> , 2010, 11, R74.	13.9	93
84	On expert curation and scalability: UniProtKB/Swiss-Prot as a case study. <i>Bioinformatics</i> , 2017, 33, 3454-3460.	1.8	91
85	Asparagine Peptide Lyases. <i>Journal of Biological Chemistry</i> , 2011, 286, 38321-38328.	1.6	89
86	Structure and distribution of pentapeptide repeats in bacteria. <i>Protein Science</i> , 1998, 7, 1477-1480.	3.1	87
87	Integrating biological data – the Distributed Annotation System. <i>BMC Bioinformatics</i> , 2008, 9, S3.	1.2	87
88	The G5 domain: a potential N-acetylglucosamine recognition domain involved in biofilm formation. <i>Bioinformatics</i> , 2005, 21, 1301-1303.	1.8	78
89	Filling out the structural map of the NTF2-like superfamily. <i>BMC Bioinformatics</i> , 2013, 14, 327.	1.2	74
90	Microbial composition of Kombucha determined using amplicon sequencing and shotgun metagenomics. <i>Journal of Food Science</i> , 2020, 85, 455-464.	1.5	72

#	ARTICLE	IF	CITATIONS
91	A catalogue of 1,167 genomes from the human gut archaeome. <i>Nature Microbiology</i> , 2022, 7, 48-61.	5.9	72
92	RNIE: genome-wide prediction of bacterial intrinsic terminators. <i>Nucleic Acids Research</i> , 2011, 39, 5845-5852.	6.5	71
93	The PepSY domain: a regulator of peptidase activity in the microbial environment?. <i>Trends in Biochemical Sciences</i> , 2004, 29, 169-172.	3.7	68
94	Integrating sequence and structural biology with DAS. <i>BMC Bioinformatics</i> , 2007, 8, 333.	1.2	68
95	RNAcentral: A vision for an international database of RNA sequences. <i>Rna</i> , 2011, 17, 1941-1946.	1.6	67
96	Members of the immunoglobulin superfamily in bacteria. <i>Protein Science</i> , 1996, 5, 1939-1941.	3.1	66
97	The RNA WikiProject: Community annotation of RNA families. <i>Rna</i> , 2008, 14, 2462-2464.	1.6	66
98	Phylogenomics of expanding uncultured environmental Tenericutes provides insights into their pathogenicity and evolutionary relationship with Bacilli. <i>BMC Genomics</i> , 2020, 21, 408.	1.2	65
99	Identifying Protein Domains with the Pfam Database. <i>Current Protocols in Bioinformatics</i> , 2008, 23, Unit 2.5.	25.8	61
100	HPMCD: the database of human microbial communities from metagenomic datasets and microbial reference genomes. <i>Nucleic Acids Research</i> , 2016, 44, D604-D609.	6.5	60
101	Phospholipid scramblases and Tubby-like proteins belong to a new superfamily of membrane tethered transcription factors. <i>Bioinformatics</i> , 2009, 25, 159-162.	1.8	59
102	Biocurators and Biocuration: surveying the 21st century challenges. <i>Database: the Journal of Biological Databases and Curation</i> , 2012, 2012, bar059-bar059.	1.4	59
103	Cryo-EM structures of human RNA polymerase III in its unbound and transcribing states. <i>Nature Structural and Molecular Biology</i> , 2021, 28, 210-219.	3.6	59
104	Estimating the quality of eukaryotic genomes recovered from metagenomic analysis with EukCC. <i>Genome Biology</i> , 2020, 21, 244.	3.8	58
105	R2DT is a framework for predicting and visualising RNA secondary structure using templates. <i>Nature Communications</i> , 2021, 12, 3494.	5.8	58
106	Integrating cultivation and metagenomics for a multi-kingdom view of skin microbiome diversity and functions. <i>Nature Microbiology</i> , 2022, 7, 169-179.	5.9	58
107	COVID-19 pandemic reveals the peril of ignoring metadata standards. <i>Scientific Data</i> , 2020, 7, 188.	2.4	56
108	Clustered Coding Variants in the Glutamate Receptor Complexes of Individuals with Schizophrenia and Bipolar Disorder. <i>PLoS ONE</i> , 2011, 6, e19011.	1.1	54

#	ARTICLE	IF	CITATIONS
109	Structure of the <i>Escherichia coli</i> ProQ RNA-binding protein. <i>Rna</i> , 2017, 23, 696-711.	1.6	50
110	Eros is a novel transmembrane protein that controls the phagocyte respiratory burst and is essential for innate immunity. <i>Journal of Experimental Medicine</i> , 2017, 214, 1111-1128.	4.2	50
111	SCOOP: a simple method for identification of novel protein superfamily relationships. <i>Bioinformatics</i> , 2007, 23, 809-814.	1.8	49
112	Enhanced protein domain discovery by using language modeling techniques from speech recognition. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2003, 100, 4516-4520.	3.3	47
113	The ELIXIR Core Data Resources: fundamental infrastructure for the life sciences. <i>Bioinformatics</i> , 2020, 36, 2636-2642.	1.8	47
114	Toward unrestricted use of public genomic data. <i>Science</i> , 2019, 363, 350-352.	6.0	45
115	Reuse of structural domain-domain interactions in protein networks. <i>BMC Bioinformatics</i> , 2007, 8, 259.	1.2	44
116	How to use the MEROPS database and website to help understand peptidase specificity. <i>Protein Science</i> , 2021, 30, 83-92.	3.1	44
117	The metagenomic data life-cycle: standards and best practices. <i>GigaScience</i> , 2017, 6, 1-11.	3.3	42
118	Unifying the known and unknown microbial coding sequence space. <i>ELife</i> , 2022, 11, .	2.8	41
119	Distant homology recognition using structural classification of proteins. , 1997, 29, 105-112.		40
120	Using the MEROPS Database for Proteolytic Enzymes and Their Inhibitors and Substrates. <i>Current Protocols in Bioinformatics</i> , 2014, 48, 1.25.1-33.	25.8	39
121	AntiFam: a tool to help identify spurious ORFs in protein annotation. <i>Database: the Journal of Biological Databases and Curation</i> , 2012, 2012, bas003-bas003.	1.4	38
122	Pfam. <i>Methods in Molecular Biology</i> , 2007, 396, 43-58.	0.4	38
123	UniProt-DAAC: domain architecture alignment and classification, a new method for automatic functional annotation in UniProtKB. <i>Bioinformatics</i> , 2016, 32, 2264-2271.	1.8	37
124	Ten Simple Rules for Editing Wikipedia. <i>PLoS Computational Biology</i> , 2010, 6, e1000941.	1.5	36
125	<i>Escherichia coli</i> RNA polymerase core and holoenzyme structures. <i>EMBO Journal</i> , 2000, 19, 6833-6844.	3.5	35
126	Visualizing profile-profile alignment: pairwise HMM logos. <i>Bioinformatics</i> , 2005, 21, 2912-2913.	1.8	35



#	ARTICLE	IF	CITATIONS
127	ProServer: a simple, extensible Perl DAS server. <i>Bioinformatics</i> , 2007, 23, 1568-1570.	1.8	34
128	Identifying accurate metagenome and amplicon software via a meta-analysis of sequence to taxonomy benchmarking studies. <i>PeerJ</i> , 2019, 7, e6160.	0.9	34
129	The European Bioinformatics Institute (EMBL-EBI) in 2021. <i>Nucleic Acids Research</i> , 2022, 50, D11-D19.	6.5	34
130	Modifier Effects between Regulatory and Protein-Coding Variation. <i>PLoS Genetics</i> , 2008, 4, e1000244.	1.5	33
131	Sequence analysis of tyrosine recombinases allows annotation of mobile genetic elements in prokaryotic genomes. <i>Molecular Systems Biology</i> , 2021, 17, e9880.	3.2	33
132	Bacterial Pleckstrin Homology Domains: A Prokaryotic Origin for the PH Domain. <i>Journal of Molecular Biology</i> , 2010, 396, 31-46.	2.0	32
133	Towards BioDBcore: a community-defined information specification for biological databases. <i>Nucleic Acids Research</i> , 2011, 39, D7-D10.	6.5	32
134	Making your database available through Wikipedia: the pros and cons. <i>Nucleic Acids Research</i> , 2012, 40, D9-D12.	6.5	31
135	Towards BioDBcore: a community-defined information specification for biological databases. Database: the Journal of Biological Databases and Curation, 2011, 2011, baq027-baq027.	1.4	30
136	Origins of peptidases. <i>Biochimie</i> , 2019, 166, 4-18.	1.3	30
137	Microbial community drivers of PK/NRP gene diversity in selected global soils. <i>Microbiome</i> , 2019, 7, 78.	4.9	30
138	A member of the immunoglobulin superfamily in bacteriophage T4. <i>Virus Genes</i> , 1997, 14, 163-165.	0.7	28
139	Tandem domain swapping: determinants of multidomain protein misfolding. <i>Current Opinion in Structural Biology</i> , 2019, 58, 97-104.	2.6	28
140	Genome properties in 2019: a new companion database to InterPro for the inference of complete functional attributes. <i>Nucleic Acids Research</i> , 2019, 47, D564-D572.	6.5	27
141	Ten simple rules to make your computing more environmentally sustainable. <i>PLoS Computational Biology</i> , 2021, 17, e1009324.	1.5	27
142	Priorities for ocean microbiome research. <i>Nature Microbiology</i> , 2022, 7, 937-947.	5.9	27
143	Novel protein domains and motifs in the marine planctomycete <i>Rhodopirellula baltica</i> . <i>FEMS Microbiology Letters</i> , 2004, 236, 333-340.	0.7	25
144	The yeast noncoding RNA interaction network. <i>Rna</i> , 2017, 23, 1479-1492.	1.6	25

#	ARTICLE	IF	CITATIONS
145	Recovering prokaryotic genomes from host-associated, short-read shotgun metagenomic sequencing data. <i>Nature Protocols</i> , 2021, 16, 2520-2541.	5.5	25
146	Experience using web services for biological sequence analysis. <i>Briefings in Bioinformatics</i> , 2008, 9, 493-505.	3.2	24
147	Outline structures for the extracellular domains of the fibroblast growth factor receptors. <i>Nature Structural and Molecular Biology</i> , 1995, 2, 1068-1074.	3.6	23
148	Domain atrophy creates rare cases of functional partial protein domains. <i>Genome Biology</i> , 2015, 16, 88.	3.8	23
149	The challenge of increasing Pfam coverage of the human proteome. <i>Database: the Journal of Biological Databases and Curation</i> , 2013, 2013, bat023.	1.4	22
150	TreeGrafter: phylogenetic tree-based annotation of proteins with Gene Ontology terms and other annotations. <i>Bioinformatics</i> , 2019, 35, 518-520.	1.8	22
151	Acetylation of Surface Carbohydrates in Bacterial Pathogens Requires Coordinated Action of a Two-Domain Membrane-Bound Acyltransferase. <i>MBio</i> , 2020, 11, .	1.8	22
152	Microbiota Characterization of Agricultural Green Waste-Based Suppressive Composts Using Omics and Classic Approaches. <i>Agriculture (Switzerland)</i> , 2020, 10, 61.	1.4	21
153	Identifying Protein Domains with the Pfam Database. <i>Current Protocols in Bioinformatics</i> , 2003, 1, 2.5.1-2.5.19.	25.8	20
154	GO annotation in InterPro: why stability does not indicate accuracy in a sea of changing annotations. <i>Database: the Journal of Biological Databases and Curation</i> , 2016, 2016, baw027.	1.4	19
155	Predicted Input of Uncultured Fungal Symbionts to a Lichen Symbiosis from Metagenome-Assembled Genomes. <i>Genome Biology and Evolution</i> , 2021, 13, .	1.1	19
156	Enhanced protein domain discovery using taxonomy. <i>BMC Bioinformatics</i> , 2004, 5, 56.	1.2	18
157	Eleven quick tips to build a usable REST API for life sciences. <i>PLoS Computational Biology</i> , 2018, 14, e1006542.	1.5	18
158	ELIXIRâ€™s EXCELERATE: establishing Europe's data infrastructure for the life science research of the future. <i>EMBO Journal</i> , 2021, 40, e107409.	3.5	18
159	An inter-laboratory study to investigate the impact of the bioinformatics component on microbiome analysis using mock communities. <i>Scientific Reports</i> , 2021, 11, 10590.	1.6	17
160	Patterns of database citation in articles and patents indicate long-term scientific and industry value of biological data resources. <i>F1000Research</i> , 2016, 5, 160.	0.8	16
161	DASMI: exchanging, annotating and assessing molecular interaction data. <i>Bioinformatics</i> , 2009, 25, 1321-1328.	1.8	15
162	Defining the remarkable structural malleability of a bacterial surface protein Rib domain implicated in infection. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2019, 116, 26540-26548.	3.3	15

#	ARTICLE	IF	CITATIONS
163	Solution structure of human myeloid-derived growth factor suggests a conserved function in the endoplasmic reticulum. <i>Nature Communications</i> , 2019, 10, 5612.	5.8	15
164	Periscope Proteins are variable-length regulators of bacterial cell surface interactions. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2021, 118, .	3.3	15
165	Key challenges for the creation and maintenance of specialist protein resources. <i>Proteins: Structure, Function and Bioinformatics</i> , 2015, 83, 1005-1013.	1.5	13
166	Genome3D: integrating a collaborative data pipeline to expand the depth and breadth of consensus protein structure annotation. <i>Nucleic Acids Research</i> , 2020, 48, D314-D319.	6.5	13
167	Pfam: the protein families database. , 2005, , .		12
168	The crystal structure of a bacterial Sufu-like protein defines a novel group of bacterial proteins that are similar to the N-terminal domain of human Sufu. <i>Protein Science</i> , 2010, 19, 2131-2140.	3.1	12
169	Gene Unprediction with Spurio: A tool to identify spurious protein sequences. <i>F1000Research</i> , 2018, 7, 261.	0.8	11
170	The Ensembl COVID-19 resource: ongoing integration of public SARS-CoV-2 data. <i>Nucleic Acids Research</i> , 2022, 50, D765-D770.	6.5	10
171	The structure of Jann_2411 (DUF1470) from <i>Jannaschia</i> sp. at 1.45 Å resolution reveals a new fold (the ABATE domain) and suggests its possible role as a transcription regulator. <i>Acta Crystallographica Section F: Structural Biology Communications</i> , 2010, 66, 1198-1204.	0.7	9
172	The structure of BVU2987 from <i>Bacteroides vulgatus</i> reveals a superfamily of bacterial periplasmic proteins with possible inhibitory function. <i>Acta Crystallographica Section F: Structural Biology Communications</i> , 2010, 66, 1265-1273.	0.7	8
173	LUD, a new protein domain associated with lactate utilization. <i>BMC Bioinformatics</i> , 2013, 14, 341.	1.2	8
174	ISCB Computational Biology Wikipedia Competition. <i>PLoS Computational Biology</i> , 2013, 9, e1003242.	1.5	8
175	Creating a specialist protein resource network: a meeting report for the protein bioinformatics and community resources retreat: Figure 1.. Database: the Journal of Biological Databases and Curation, 2015, 2015, bav063.	1.4	8
176	ELIXIR pilot action: Marine metagenomics towards a domain specific set of sustainable services. <i>F1000Research</i> , 2017, 6, 70.	0.8	8
177	A mouse model of occult intestinal colonization demonstrating antibiotic-induced outgrowth of carbapenem-resistant Enterobacteriaceae. <i>Microbiome</i> , 2022, 10, 43.	4.9	8
178	Discovery of fibrillar adhesins across bacterial species. <i>BMC Genomics</i> , 2021, 22, 550.	1.2	7
179	The structure of pyogenecin immunity protein, a novel bacteriocin-like immunity protein from <i>Streptococcus pyogenes</i> . <i>BMC Structural Biology</i> , 2009, 9, 75.	2.3	6
180	The complexity, challenges and benefits of comparing two transporter classification systems in TCDB and Pfam. <i>Briefings in Bioinformatics</i> , 2015, 16, 865-872.	3.2	6

#	ARTICLE	IF	CITATIONS
181	TADOSS: computational estimation of tandem domain swap stability. <i>Bioinformatics</i> , 2019, 35, 2507-2508.	1.8	6
182	Exploring Non-coding RNAs in RNAcentral. <i>Current Protocols in Bioinformatics</i> , 2020, 71, e104.	25.8	6
183	The SHOCT Domain: A Widespread Domain Under-Represented in Model Organisms. <i>PLoS ONE</i> , 2013, 8, e57848.	1.1	6
184	Large-Scale Discovery of Microbial Fibrillar Adhesins and Identification of Novel Members of Adhesive Domain Families. <i>Journal of Bacteriology</i> , 2022, 204, .	1.0	6
185	<i>Caenorhabditis elegans</i> AF4/FMR2 Family Homolog <i>affl-2</i> Regulates Heat-Shock-Induced Gene Expression. <i>Genetics</i> , 2020, 215, 1039-1054.	1.2	5
186	The thrombospondin module 1 domain of the matricellular protein CCN3 shows an atypical disulfide pattern and incomplete CWR layers. <i>Acta Crystallographica Section D: Structural Biology</i> , 2020, 76, 124-134.	1.1	5
187	Determination of <i>Escherichia coli</i> RNA Polymerase Structure by Single Particle Cryoelectron Microscopy. <i>Methods in Enzymology</i> , 2003, 370, 24-42.	0.4	4
188	Conformational Changes of <i>Escherichia coli</i> $\sigma^{54}$ -RNA-Polymerase upon Closed Promoter Complex Formation. <i>Journal of Molecular Biology</i> , 2005, 354, 201-205.	2.0	4
189	Metagenomics approach for <i>Polymyxa betae</i> genome assembly enables comparative analysis towards deciphering the intracellular parasitic lifestyle of the plasmodiophorids. <i>Genomics</i> , 2022, 114, 9-22.	1.3	4
190	The C-Terminal 12 Amino Acids of $\beta$ -N Are Required for Structure and Function. <i>Archives of Biochemistry and Biophysics</i> , 1999, 371, 234-240.	1.4	3
191	The YARHG Domain: An Extracellular Domain in Search of a Function. <i>PLoS ONE</i> , 2012, 7, e35575.	1.1	3
192	Two Pfam protein families characterized by a crystal structure of protein lpg2210 from <i>Legionella pneumophila</i> . <i>BMC Bioinformatics</i> , 2013, 14, 265.	1.2	3
193	3DPatch: fast 3D structure visualization with residue conservation. <i>Bioinformatics</i> , 2019, 35, 332-334.	1.8	2
194	The first structure in a family of peptidase inhibitors reveals an unusual Ig-like fold. <i>F1000Research</i> , 2013, 2, 154.	0.8	2
195	Structure and computational analysis of a novel protein with metallopeptidase-like and circularly permuted winged-helix-turn-helix domains reveals a possible role in modified polysaccharide biosynthesis. <i>BMC Bioinformatics</i> , 2014, 15, 75.	1.2	1
196	Rapid identification of novel protein families using similarity searches. <i>F1000Research</i> , 2018, 7, 1975.	0.8	1
197	Navigating bacterial taxonomy in a world of uncharted microbial organisms.., 2021, , 179-197.		0
198	A Comprehensive Analysis of the Global Human Gut Archaeome from a Thousand Genome Catalogue. <i>SSRN Electronic Journal</i> , 0, , .	0.4	0