

# Robert D Finn

## List of Publications by Citations

**Source:** <https://exaly.com/author-pdf/11310124/robert-d-finn-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.

105  
papers

41,925  
citations

56  
h-index

110  
g-index

110  
ext. papers

56,231  
ext. citations

15.2  
avg, IF

7.5  
L-index

| #   | Paper   | IF   | Citations |
|-----|---|------|-----------|
| 105 | Pfam: the protein families database. <i>Nucleic Acids Research</i> , <b>2014</b> , 42, D222-30  | 20.1 | 3975      |
| 104 | The Pfam protein families database: towards a more sustainable future. <i>Nucleic Acids Research</i> , <b>2016</b> , 44, D279-85  | 20.1 | 3678      |
| 103 | HMMER web server: interactive sequence similarity searching. <i>Nucleic Acids Research</i> , <b>2011</b> , 39, W29-37   | 20.1 | 2917      |
| 102 | The Pfam protein families database. <i>Nucleic Acids Research</i> , <b>2012</b> , 40, D290-301  | 20.1 | 2844      |
| 101 | The Pfam protein families database. <i>Nucleic Acids Research</i> , <b>2004</b> , 32, D138-41   | 20.1 | 2720      |
| 100 | The Pfam protein families database. <i>Nucleic Acids Research</i> , <b>2010</b> , 38, D211-22   | 20.1 | 2467      |
| 99  | The Pfam protein families database in 2019. <i>Nucleic Acids Research</i> , <b>2019</b> , 47, D427-D432   | 20.1 | 2298      |
| 98  | The EMBL-EBI search and sequence analysis tools APIs in 2019. <i>Nucleic Acids Research</i> , <b>2019</b> , 47, W636-W641   | 20.1 | 2132      |
| 97  | The Pfam protein families database. <i>Nucleic Acids Research</i> , <b>2008</b> , 36, D281-8  | 20.1 | 1806      |
| 96  | Pfam: clans, web tools and services. <i>Nucleic Acids Research</i> , <b>2006</b> , 34, D247-51  | 20.1 | 1784      |
| 95  | InterPro: the integrative protein signature database. <i>Nucleic Acids Research</i> , <b>2009</b> , 37, D211-5  | 20.1 | 1379      |
| 94  | InterPro in 2017-beyond protein family and domain annotations. <i>Nucleic Acids Research</i> , <b>2017</b> , 45, D190-D199  | 20.1 | 970       |
| 93  | The InterPro protein families database: the classification resource after 15 years. <i>Nucleic Acids Research</i> , <b>2015</b> , 43, D213-21   | 20.1 | 954       |
| 92  | InterPro in 2011: new developments in the family and domain prediction database. <i>Nucleic Acids Research</i> , <b>2012</b> , 40, D306-12  | 20.1 | 844       |
| 91  | InterPro in 2019: improving coverage, classification and access to protein sequence annotations. <i>Nucleic Acids Research</i> , <b>2019</b> , 47, D351-D360  | 20.1 | 835       |
| 90  | Rfam 12.0: updates to the RNA families database. <i>Nucleic Acids Research</i> , <b>2015</b> , 43, D130-7   | 20.1 | 723       |
| 89  | 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> , <b>2018</b> , 46, D624-D632 | 20.1 | 643       |

|    |  |      |     |
|----|--|------|-----|
| 88 | Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. <i>Nucleic Acids Research</i> , <b>2013</b> , 41, e121                                  | 20.1 | 629 |
| 87 | Pfam: The protein families database in 2021. <i>Nucleic Acids Research</i> , <b>2021</b> , 49, D412-D419   | 20.1 | 620 |
| 86 | HMMER web server: 2018 update. <i>Nucleic Acids Research</i> , <b>2018</b> , 46, W200-W204   | 20.1 | 596 |
| 85 | HMMER web server: 2015 update. <i>Nucleic Acids Research</i> , <b>2015</b> , 43, W30-8   | 20.1 | 590 |
| 84 | Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. <i>Nucleic Acids Research</i> , <b>2018</b> , 46, D335-D342                                      | 20.1 | 533 |
| 83 | A new genomic blueprint of the human gut microbiota. <i>Nature</i> , <b>2019</b> , 568, 499-504  | 50.4 | 484 |
| 82 | The InterPro protein families and domains database: 20 years on. <i>Nucleic Acids Research</i> , <b>2021</b> , 49, D344-D354   | 20.1 | 358 |
| 81 | Ensembl Genomes 2018: an integrated omics infrastructure for non-vertebrate species. <i>Nucleic Acids Research</i> , <b>2018</b> , 46, D802-D808                               | 20.1 | 330 |
| 80 | Rfam: Wikipedia, clans and the "decimal" release. <i>Nucleic Acids Research</i> , <b>2011</b> , 39, D141-5   | 20.1 | 321 |
| 79 | The Dfam database of repetitive DNA families. <i>Nucleic Acids Research</i> , <b>2016</b> , 44, D81-9  | 20.1 | 295 |
| 78 | iPfam: visualization of protein-protein interactions in PDB at domain and amino acid resolutions. <i>Bioinformatics</i> , <b>2005</b> , 21, 410-2                              | 7.2  | 263 |
| 77 | A human gut bacterial genome and culture collection for improved metagenomic analyses. <i>Nature Biotechnology</i> , <b>2019</b> , 37, 186-192                                 | 44.5 | 224 |
| 76 | Skyline: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. <i>BMC Bioinformatics</i> , <b>2014</b> , 15, 7 | 3.6  | 196 |
| 75 | A unified catalog of 204,938 reference genomes from the human gut microbiome. <i>Nature Biotechnology</i> , <b>2021</b> , 39, 105-114  | 44.5 | 185 |
| 74 | Non-Coding RNA Analysis Using the Rfam Database. <i>Current Protocols in Bioinformatics</i> , <b>2018</b> , 62, e51  | 24.2 | 152 |
| 73 | Dfam: a database of repetitive DNA based on profile hidden Markov models. <i>Nucleic Acids Research</i> , <b>2013</b> , 41, D70-82   | 20.1 | 147 |
| 72 | DUFs: families in search of function. <i>Acta Crystallographica Section F: Structural Biology Communications</i> , <b>2010</b> , 66, 1148-52                                   |      | 145 |
| 71 | Microbial abundance, activity and population genomic profiling with mOTUs2. <i>Nature Communications</i> , <b>2019</b> , 10, 1014  | 17.4 | 134 |

|    |  |      |     |
|----|--|------|-----|
| 70 | EBI Metagenomics in 2017: enriching the analysis of microbial communities, from sequence reads to assemblies. <i>Nucleic Acids Research</i> , <b>2018</b> , 46, D726-D735  | 20.1 | 130 |
| 69 | MGNify: the microbiome analysis resource in 2020. <i>Nucleic Acids Research</i> , <b>2020</b> , 48, D570-D578  | 20.1 | 127 |
| 68 | Predicting active site residue annotations in the Pfam database. <i>BMC Bioinformatics</i> , <b>2007</b> , 8, 298  | 3.6  | 124 |
| 67 | iPfam: a database of protein family and domain interactions found in the Protein Data Bank. <i>Nucleic Acids Research</i> , <b>2014</b> , 42, D364-73  | 20.1 | 120 |
| 66 | RNAcentral: a comprehensive database of non-coding RNA sequences. <i>Nucleic Acids Research</i> , <b>2017</b> , 45, D128-D134  | 20.1 | 119 |
| 65 | Rfam 14: expanded coverage of metagenomic, viral and microRNA families. <i>Nucleic Acids Research</i> , <b>2021</b> , 49, D192-D200  | 20.1 | 108 |
| 64 | Pfam 10 years on: 10,000 families and still growing. <i>Briefings in Bioinformatics</i> , <b>2008</b> , 9, 210-9   | 13.4 | 101 |
| 63 | 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> , <b>2016</b> , 12, e1004862      | 5    | 95  |
| 62 | Massive expansion of human gut bacteriophage diversity. <i>Cell</i> , <b>2021</b> , 184, 1098-1109.e9  | 56.2 | 93  |
| 61 | The European Bioinformatics Institute in 2016: Data growth and integration. <i>Nucleic Acids Research</i> , <b>2016</b> , 44, D20-6  | 20.1 | 91  |
| 60 | RNAcentral: a hub of information for non-coding RNA sequences. <i>Nucleic Acids Research</i> , <b>2019</b> , 47, D221-D229   | 20.1 | 90  |
| 59 | RNAcentral: an international database of ncRNA sequences. <i>Nucleic Acids Research</i> , <b>2015</b> , 43, D123-9   | 20.1 | 89  |
| 58 | EBI metagenomics in 2016--an expanding and evolving resource for the analysis and archiving of metagenomic data. <i>Nucleic Acids Research</i> , <b>2016</b> , 44, D595-603  | 20.1 | 81  |
| 57 | Integrating biological data--the Distributed Annotation System. <i>BMC Bioinformatics</i> , <b>2008</b> , 9 Suppl 8, S3  | 3.6  | 77  |
| 56 | Defining the in Vivo Role for cytochrome b5 in cytochrome P450 function through the conditional hepatic deletion of microsomal cytochrome b5. <i>Journal of Biological Chemistry</i> , <b>2008</b> , 283, 31385-93 | 5.4  | 69  |
| 55 | Representative proteomes: a stable, scalable and unbiased proteome set for sequence analysis and functional annotation. <i>PLoS ONE</i> , <b>2011</b> , 6, e18910  | 3.7  | 67  |
| 54 | Unsaturated fatty acid regulation of cytochrome P450 expression via a CAR-dependent pathway. <i>Biochemical Journal</i> , <b>2009</b> , 417, 43-54   | 3.8  | 61  |
| 53 | Benchmarking taxonomic assignments based on 16S rRNA gene profiling of the microbiota from commonly sampled environments. <i>GigaScience</i> , <b>2018</b> , 7,  | 7.6  | 61  |

|    |   |      |    |
|----|---|------|----|
| 52 | Deletion of microsomal cytochrome b5 profoundly affects hepatic and extrahepatic drug metabolism. <i>Molecular Pharmacology</i> , <b>2010</b> , 78, 269-78  | 4.3  | 60 |
| 51 | Integrating sequence and structural biology with DAS. <i>BMC Bioinformatics</i> , <b>2007</b> , 8, 333  | 3.6  | 57 |
| 50 | Computational strategies to combat COVID-19: useful tools to accelerate SARS-CoV-2 and coronavirus research. <i>Briefings in Bioinformatics</i> , <b>2021</b> , 22, 642-663   | 13.4 | 56 |
| 49 | HPMCD: the database of human microbial communities from metagenomic datasets and microbial reference genomes. <i>Nucleic Acids Research</i> , <b>2016</b> , 44, D604-9  | 20.1 | 50 |
| 48 | The HMMER Web Server for Protein Sequence Similarity Search. <i>Current Protocols in Bioinformatics</i> , <b>2017</b> , 60, 3.15.1-3.15.23  | 24.2 | 49 |
| 47 | Identifying protein domains with the Pfam database. <i>Current Protocols in Bioinformatics</i> , <b>2008</b> , Chapter 2, Unit 2.5  | 24.2 | 43 |
| 46 | RNAcentral 2021: secondary structure integration, improved sequence search and new member databases. <i>Nucleic Acids Research</i> , <b>2021</b> , 49, D212-D220  | 20.1 | 42 |
| 45 | SCOOP: a simple method for identification of novel protein superfamily relationships. <i>Bioinformatics</i> , <b>2007</b> , 23, 809-14  | 7.2  | 36 |
| 44 | Conditional deletion of cytochrome P450 oxidoreductase in the liver and gastrointestinal tract: a new model for studying the functions of the P450 system. <i>Journal of Pharmacology and Experimental Therapeutics</i> , <b>2007</b> , 322, 40-7 | 4.7  | 33 |
| 43 | COVID-19 pandemic reveals the peril of ignoring metadata standards. <i>Scientific Data</i> , <b>2020</b> , 7, 188   | 8.2  | 30 |
| 42 | Modifier effects between regulatory and protein-coding variation. <i>PLoS Genetics</i> , <b>2008</b> , 4, e1000244  | 6    | 30 |
| 41 | The metagenomic data life-cycle: standards and best practices. <i>GigaScience</i> , <b>2017</b> , 6, 1-11   | 7.6  | 29 |
| 40 | Microbial composition of Kombucha determined using amplicon sequencing and shotgun metagenomics. <i>Journal of Food Science</i> , <b>2020</b> , 85, 455-464   | 3.4  | 28 |
| 39 | Bacterial pleckstrin homology domains: a prokaryotic origin for the PH domain. <i>Journal of Molecular Biology</i> , <b>2010</b> , 396, 31-46   | 6.5  | 25 |
| 38 | Cytochrome b5 null mouse: a new model for studying inherited skin disorders and the role of unsaturated fatty acids in normal homeostasis. <i>Transgenic Research</i> , <b>2011</b> , 20, 491-502   | 3.3  | 24 |
| 37 | Making your database available through Wikipedia: the pros and cons. <i>Nucleic Acids Research</i> , <b>2012</b> , 40, D9-12  | 20.1 | 23 |
| 36 | Experience using web services for biological sequence analysis. <i>Briefings in Bioinformatics</i> , <b>2008</b> , 9, 493-504   | 5.4  | 22 |
| 35 | Identifying accurate metagenome and amplicon software via a meta-analysis of sequence to taxonomy benchmarking studies. <i>PeerJ</i> , <b>2019</b> , 7, e6160   | 3.1  | 22 |

|    |  |      |    |
|----|--|------|----|
| 34 | Phylogenomics of expanding uncultured environmental Tenericutes provides insights into their pathogenicity and evolutionary relationship with Bacilli. <i>BMC Genomics</i> , <b>2020</b> , 21, 408             | 4.5  | 19 |
| 33 | Reporting guidelines for human microbiome research: the STORMS checklist. <i>Nature Medicine</i> , <b>2021</b> , 27, 1885-1892   | 50.5 | 19 |
| 32 | Estimating the quality of eukaryotic genomes recovered from metagenomic analysis with EukCC. <i>Genome Biology</i> , <b>2020</b> , 21, 244   | 18.3 | 19 |
| 31 | 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> , <b>2016</b> , 2016,                  | 5    | 16 |
| 30 | The challenge of increasing Pfam coverage of the human proteome. <i>Database: the Journal of Biological Databases and Curation</i> , <b>2013</b> , 2013, bat023  | 5    | 16 |
| 29 | R2DT is a framework for predicting and visualising RNA secondary structure using templates. <i>Nature Communications</i> , <b>2021</b> , 12, 3494  | 17.4 | 16 |
| 28 | Microbial community drivers of PK/NRP gene diversity in selected global soils. <i>Microbiome</i> , <b>2019</b> , 7, 78   | 16.6 | 15 |
| 27 | Genome properties in 2019: a new companion database to InterPro for the inference of complete functional attributes. <i>Nucleic Acids Research</i> , <b>2019</b> , 47, D564-D572                               | 20.1 | 15 |
| 26 | A unified sequence catalogue of over 280,000 genomes obtained from the human gut microbiome  |      | 14 |
| 25 | 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> , <b>2010</b> , 19, 2131-40 | 6.3  | 12 |
| 24 | Ensembl Genomes 2022: an expanding genome resource for non-vertebrates. <i>Nucleic Acids Research</i> , <b>2021</b> ,  | 20.1 | 12 |
| 23 | TreeGrafter: phylogenetic tree-based annotation of proteins with Gene Ontology terms and other annotations. <i>Bioinformatics</i> , <b>2019</b> , 35, 518-520  | 7.2  | 12 |
| 22 | ELIXIR-EXCELERATE: establishing Europe's data infrastructure for the life science research of the future. <i>EMBO Journal</i> , <b>2021</b> , 40, e107409  | 13   | 11 |
| 21 | A catalogue of 1,167 genomes from the human gut archaeome.. <i>Nature Microbiology</i> , <b>2021</b> ,   | 26.6 | 9  |
| 20 | Unifying the known and unknown microbial coding sequence space   |      | 9  |
| 19 | A vaccinology Approach to the Identification and Characterization of Candidate Protective Antigens for the Control of Poultry Red Mite Infestations. <i>Vaccines</i> , <b>2019</b> , 7,                        | 5.3  | 9  |
| 18 | Predicted Input of Uncultured Fungal Symbionts to a Lichen Symbiosis from Metagenome-Assembled Genomes. <i>Genome Biology and Evolution</i> , <b>2021</b> , 13,  | 3.9  | 8  |
| 17 | Genome3D: integrating a collaborative data pipeline to expand the depth and breadth of consensus protein structure annotation. <i>Nucleic Acids Research</i> , <b>2020</b> , 48, D314-D319                     | 20.1 | 7  |

|    |   |        |
|----|---|--------|
| 16 | Eleven quick tips to build a usable REST API for life sciences. <i>PLoS Computational Biology</i> , <b>2018</b> , 14, e1006542  | 7      |
| 15 | ELIXIR pilot action: Marine metagenomics - towards a domain specific set of sustainable services. <i>F1000Research</i> , <b>2017</b> , 6,   | 3.6 6  |
| 14 | Integrating cultivation and metagenomics for a multi-kingdom view of skin microbiome diversity and functions.. <i>Nature Microbiology</i> , <b>2021</b> ,   | 26.6 6 |
| 13 | The complexity, challenges and benefits of comparing two transporter classification systems in TCDB and Pfam. <i>Briefings in Bioinformatics</i> , <b>2015</b> , 16, 865-72   | 13.4 5 |
| 12 | Creating a specialist protein resource network: a meeting report for the protein bioinformatics and community resources retreat. <i>Database: the Journal of Biological Databases and Curation</i> , <b>2015</b> , 2015, bav063 | 5 5    |
| 11 | Conformational changes of Escherichia coli sigma54-RNA-polymerase upon closed-promoter complex formation. <i>Journal of Molecular Biology</i> , <b>2005</b> , 354, 201-5  | 6.5 4  |
| 10 | Determination of Escherichia coli RNA polymerase structure by single particle cryoelectron microscopy. <i>Methods in Enzymology</i> , <b>2003</b> , 370, 24-42  | 1.7 4  |
| 9  | Unifying the known and unknown microbial coding sequence space.. <i>ELife</i> , <b>2022</b> , 11,   | 8.9 4  |
| 8  | Estimating the quality of eukaryotic genomes recovered from metagenomic analysis  | 3      |
| 7  | R2DT: computational framework for template-based RNA secondary structure visualisation across non-coding RNA types  | 3      |
| 6  | Exploring Non-Coding RNAs in RNACentral. <i>Current Protocols in Bioinformatics</i> , <b>2020</b> , 71, e104  | 24.2 3 |
| 5  | An inter-laboratory study to investigate the impact of the bioinformatics component on microbiome analysis using mock communities. <i>Scientific Reports</i> , <b>2021</b> , 11, 10590  | 4.9 3  |
| 4  | The Ensembl COVID-19 resource: ongoing integration of public SARS-CoV-2 data. <i>Nucleic Acids Research</i> , <b>2021</b> ,   | 20.1 3 |
| 3  | 3DPatch: fast 3D structure visualization with residue conservation. <i>Bioinformatics</i> , <b>2019</b> , 35, 332-334   | 7.2 2  |
| 2  | Recovering prokaryotic genomes from host-associated, short-read shotgun metagenomic sequencing data. <i>Nature Protocols</i> , <b>2021</b> , 16, 2520-2541  | 18.8 2 |
| 1  | Metagenomics approach for Polymyxa betae genome assembly enables comparative analysis towards deciphering the intracellular parasitic lifestyle of the plasmodiophorids. <i>Genomics</i> , <b>2021</b> , 114, 9-22              | 4.3 0  |