

# MMseqs2 enables sensitive protein sequence searching sets

Nature Biotechnology

35, 1026-1028

DOI: [10.1038/nbt.3988](https://doi.org/10.1038/nbt.3988)

Citation Report

| #  | ARTICLE                                                                                                                                                                               | IF  | CITATIONS |
|----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 2  | Neutral Theory, Microbial Practice: Challenges in Bacterial Population Genetics. <i>Molecular Biology and Evolution</i> , 2018, 35, 1338-1347.                                        | 3.5 | 93        |
| 3  | fMLC: fast multi-level clustering and visualization of large molecular datasets. <i>Bioinformatics</i> , 2018, 34, 1577-1579.                                                         | 1.8 | 11        |
| 4  | Cas13d Is a Compact RNA-Targeting Type VI CRISPR Effector Positively Modulated by a WYL-Domain-Containing Accessory Protein. <i>Molecular Cell</i> , 2018, 70, 327-339.e5.            | 4.5 | 356       |
| 5  | A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. <i>Journal of Molecular Biology</i> , 2018, 430, 2237-2243.                               | 2.0 | 1,956     |
| 6  | Reflections on a journey: a retrospective of the ISCB Student Council symposium series. <i>BMC Bioinformatics</i> , 2018, 19, 347.                                                    | 1.2 | 9         |
| 7  | Genetic exchanges are more frequent in bacteria encoding capsules. <i>PLoS Genetics</i> , 2018, 14, e1007862.                                                                         | 1.5 | 31        |
| 8  | DynBench3D, a Web-Resource to Dynamically Generate Benchmark Sets of Large Heteromeric Protein Complexes. <i>Journal of Molecular Biology</i> , 2018, 430, 4431-4438.                 | 2.0 | 3         |
| 9  | The Origin of Mitochondria-Specific Outer Membrane $\beta$ -Barrels from an Ancestral Bacterial Fragment. <i>Genome Biology and Evolution</i> , 2018, 10, 2759-2765.                  | 1.1 | 16        |
| 10 | Molecular recognition of the beta-glucans laminarin and pustulan by a SusD-like glycan-binding protein of a marine <i>Bacteroidetes</i> . <i>FEBS Journal</i> , 2018, 285, 4465-4481. | 2.2 | 13        |
| 11 | Hidden in plain sight—highly abundant and diverse planktonic freshwater Chloroflexi. <i>Microbiome</i> , 2018, 6, 176.                                                                | 4.9 | 130       |
| 12 | The Ocean Gene Atlas: exploring the biogeography of plankton genes online. <i>Nucleic Acids Research</i> , 2018, 46, W289-W295.                                                       | 6.5 | 126       |
| 13 | AC-DIAMOND v1: accelerating large-scale DNA-protein alignment. <i>Bioinformatics</i> , 2018, 34, 3744-3746.                                                                           | 1.8 | 8         |
| 14 | Clustering huge protein sequence sets in linear time. <i>Nature Communications</i> , 2018, 9, 2542.                                                                                   | 5.8 | 497       |
| 15 | Mantis: A Fast, Small, and Exact Large-Scale Sequence-Search Index. <i>Cell Systems</i> , 2018, 7, 201-207.e4.                                                                        | 2.9 | 79        |
| 16 | HFSP: high speed homology-driven function annotation of proteins. <i>Bioinformatics</i> , 2018, 34, i304-i312.                                                                        | 1.8 | 22        |
| 17 | RaptorX-Angle: real-value prediction of protein backbone dihedral angles through a hybrid method of clustering and deep learning. <i>BMC Bioinformatics</i> , 2018, 19, 100.          | 1.2 | 44        |
| 18 | SonicParanoid: fast, accurate and easy orthology inference. <i>Bioinformatics</i> , 2019, 35, 149-151.                                                                                | 1.8 | 103       |
| 19 | Phylo-PFP: improved automated protein function prediction using phylogenetic distance of distantly related sequences. <i>Bioinformatics</i> , 2019, 35, 753-759.                      | 1.8 | 29        |

| #  | ARTICLE                                                                                                                                                                                                           | IF   | CITATIONS |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 20 | Sequence Similarity Searching. <i>Current Protocols in Protein Science</i> , 2019, 95, e71.                                                                                                                       | 2.8  | 50        |
| 21 | A Reverse Ecology Approach Based on a Biological Definition of Microbial Populations. <i>Cell</i> , 2019, 178, 820-834.e14.                                                                                       | 13.5 | 118       |
| 22 | Benchmarking Metagenomics Tools for Taxonomic Classification. <i>Cell</i> , 2019, 178, 779-794.                                                                                                                   | 13.5 | 364       |
| 23 | RAFTS3G: an efficient and versatile clustering software to analyses in large protein datasets. <i>BMC Bioinformatics</i> , 2019, 20, 392.                                                                         | 1.2  | 3         |
| 24 | Effects of microbiota-directed foods in gnotobiotic animals and undernourished children. <i>Science</i> , 2019, 365, .                                                                                            | 6.0  | 305       |
| 25 | Multiple C2 domains and transmembrane region proteins ( <scp>MCTP</scp> s) tether membranes at plasmodesmata. <i>EMBO Reports</i> , 2019, 20, e47182.                                                             | 2.0  | 92        |
| 26 | Non-catalytic Binding Sites Induce Weaker Long-Range Evolutionary Rate Gradients than Catalytic Sites in Enzymes. <i>Journal of Molecular Biology</i> , 2019, 431, 3860-3870.                                     | 2.0  | 9         |
| 28 | OMA standalone: orthology inference among public and custom genomes and transcriptomes. <i>Genome Research</i> , 2019, 29, 1152-1163.                                                                             | 2.4  | 111       |
| 29 | Fecal Short-Chain Fatty Acids Are Not Predictive of Colonic Tumor Status and Cannot Be Predicted Based on Bacterial Community Structure. <i>MBio</i> , 2019, 10, .                                                | 1.8  | 32        |
| 30 | Genes functioned in kleptoplastids of <i>Dinophysis</i> are derived from haptophytes rather than from cryptophytes. <i>Scientific Reports</i> , 2019, 9, 9009.                                                    | 1.6  | 15        |
| 31 | The draft genome of <i>Actinia tenebrosa</i> reveals insights into toxin evolution. <i>Ecology and Evolution</i> , 2019, 9, 11314-11328.                                                                          | 0.8  | 28        |
| 32 | Cyclic GMP&#x2013;AMP signalling protects bacteria against viral infection. <i>Nature</i> , 2019, 574, 691-695.                                                                                                   | 13.7 | 370       |
| 33 | The <i>Rhododendron</i> Genome and Chromosomal Organization Provide Insight into Shared Whole-Genome Duplications across the Heath Family (Ericaceae). <i>Genome Biology and Evolution</i> , 2019, 11, 3353-3371. | 1.1  | 47        |
| 34 | Gene Expression Changes and Community Turnover Differentially Shape the Global Ocean Metatranscriptome. <i>Cell</i> , 2019, 179, 1068-1083.e21.                                                                   | 13.5 | 268       |
| 35 | A Conserved Streptococcal Virulence Regulator Controls the Expression of a Distinct Class of M-Like Proteins. <i>MBio</i> , 2019, 10, .                                                                           | 1.8  | 8         |
| 36 | Phage-centric ecological interactions in aquatic ecosystems revealed through ultra-deep metagenomics. <i>Microbiome</i> , 2019, 7, 135.                                                                           | 4.9  | 68        |
| 37 | TreeCluster: Clustering biological sequences using phylogenetic trees. <i>PLoS ONE</i> , 2019, 14, e0221068.                                                                                                      | 1.1  | 99        |
| 38 | Risk Factors Associated With Malignant Transformation in Patients With Oral Leukoplakia in a Chinese Population: A Retrospective Study. <i>Journal of Oral and Maxillofacial Surgery</i> , 2019, 77, 2483-2493.   | 0.5  | 15        |

| #  | ARTICLE                                                                                                                                                                                                                                    | IF  | CITATIONS |
|----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 39 | Graph Splitting: A Graph-Based Approach for Superfamily-Scale Phylogenetic Tree Reconstruction. <i>Systematic Biology</i> , 2020, 69, 265-279.                                                                                             | 2.7 | 12        |
| 40 | Gene Tags Assessment by Comparative Genomics (GTACG): A User-Friendly Framework for Bacterial Comparative Genomics. <i>Frontiers in Genetics</i> , 2019, 10, 725.                                                                          | 1.1 | 4         |
| 41 | Systematic prediction of functionally linked genes in bacterial and archaeal genomes. <i>Nature Protocols</i> , 2019, 14, 3013-3031.                                                                                                       | 5.5 | 21        |
| 42 | HH-suite3 for fast remote homology detection and deep protein annotation. <i>BMC Bioinformatics</i> , 2019, 20, 473.                                                                                                                       | 1.2 | 706       |
| 43 | Identification of a distinct lineage of aviadenovirus from crane feces. <i>Virus Genes</i> , 2019, 55, 815-824.                                                                                                                            | 0.7 | 7         |
| 44 | The distinction of CPR bacteria from other bacteria based on protein family content. <i>Nature Communications</i> , 2019, 10, 4173.                                                                                                        | 5.8 | 112       |
| 45 | MBGD update 2018: microbial genome database based on hierarchical orthology relations covering closely related and distantly related comparisons. <i>Nucleic Acids Research</i> , 2019, 47, D382-D389.                                     | 6.5 | 34        |
| 46 | ProteinNet: a standardized data set for machine learning of protein structure. <i>BMC Bioinformatics</i> , 2019, 20, 311.                                                                                                                  | 1.2 | 84        |
| 47 | Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. <i>Nature Methods</i> , 2019, 16, 603-606.                                                                                                   | 9.0 | 262       |
| 48 | M1CROB1A1Z3Râ€”a user-friendly web server for the analysis of large-scale microbial genomics data. <i>Nucleic Acids Research</i> , 2019, 47, W88-W92.                                                                                      | 6.5 | 86        |
| 49 | Human contamination in bacterial genomes has created thousands of spurious proteins. <i>Genome Research</i> , 2019, 29, 954-960.                                                                                                           | 2.4 | 111       |
| 50 | TASmania: A bacterial Toxin-Antitoxin Systems database. <i>PLoS Computational Biology</i> , 2019, 15, e1006946.                                                                                                                            | 1.5 | 95        |
| 51 | De Novo Clustering of Long-Read Transcriptome Data Using a Greedy, Quality-Value Based Algorithm. <i>Lecture Notes in Computer Science</i> , 2019, , 227-242.                                                                              | 1.0 | 11        |
| 52 | Reiterative Synthesis by the Ribosome and Recognition of the N-Terminal Formyl Group by Biosynthetic Machinery Contribute to Evolutionary Conservation of the Length of Antibiotic Microcin C Peptide Precursor. <i>MBio</i> , 2019, 10, . | 1.8 | 6         |
| 53 | Genomic Analyses of <i>Bifidobacterium moukalabense</i> Reveal Adaptations to Frugivore/Folivore Feeding Behavior. <i>Microorganisms</i> , 2019, 7, 99.                                                                                    | 1.6 | 6         |
| 54 | Microbial abundance, activity and population genomic profiling with mOTUs2. <i>Nature Communications</i> , 2019, 10, 1014.                                                                                                                 | 5.8 | 298       |
| 55 | iTRAQâ€”Based Quantitative Proteomic Analysis of Duck Eggshell During Biomineralization. <i>Proteomics</i> , 2019, 19, 1900011.                                                                                                            | 1.3 | 14        |
| 56 | Research in Computational Molecular Biology. <i>Lecture Notes in Computer Science</i> , 2019, , .                                                                                                                                          | 1.0 | 0         |

| #  | ARTICLE                                                                                                                                                                                                                  | IF   | CITATIONS |
|----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 57 | Casting light on Asgardarchaeota metabolism in a sunlit microoxic niche. <i>Nature Microbiology</i> , 2019, 4, 1129-1137.                                                                                                | 5.9  | 96        |
| 58 | WGDdetector: a pipeline for detecting whole genome duplication events using the genome or transcriptome annotations. <i>BMC Bioinformatics</i> , 2019, 20, 75.                                                           | 1.2  | 26        |
| 59 | NetSurfPâ€2.0: Improved prediction of protein structural features by integrated deep learning. <i>Proteins: Structure, Function and Bioinformatics</i> , 2019, 87, 520-527.                                              | 1.5  | 439       |
| 60 | Wide Distribution of Phage That Infect Freshwater SAR11 Bacteria. <i>MSystems</i> , 2019, 4, .                                                                                                                           | 1.7  | 11        |
| 61 | SPOT-Disorder2: Improved Protein Intrinsic Disorder Prediction by Ensembled Deep Learning. <i>Genomics, Proteomics and Bioinformatics</i> , 2019, 17, 645-656.                                                           | 3.0  | 106       |
| 62 | Parallelized Pipeline for Whole Genome Shotgun Metagenomics with GHOSTZ-GPU and MEGAN. , 2019, , .                                                                                                                       |      | 0         |
| 63 | Extracellular electron transfer powers flavinylated extracellular reductases in Gram-positive bacteria. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2019, 116, 26892-26899. | 3.3  | 68        |
| 64 | DGCF: A Distributed Greedy Clustering Framework for Large-scale Genomic Sequences. , 2019, , .                                                                                                                           |      | 1         |
| 65 | OrthoFinder: phylogenetic orthology inference for comparative genomics. <i>Genome Biology</i> , 2019, 20, 238.                                                                                                           | 3.8  | 3,367     |
| 66 | The MG-RAST API explorer: an on-ramp for RESTful query composition. <i>BMC Bioinformatics</i> , 2019, 20, 561.                                                                                                           | 1.2  | 10        |
| 67 | Hadrurid Scorpion Toxins: Evolutionary Conservation and Selective Pressures. <i>Toxins</i> , 2019, 11, 637.                                                                                                              | 1.5  | 9         |
| 68 | Modeling aspects of the language of life through transfer-learning protein sequences. <i>BMC Bioinformatics</i> , 2019, 20, 723.                                                                                         | 1.2  | 319       |
| 69 | Necrotizing enterocolitis is preceded by increased gut bacterial replication, <i>Klebsiella</i> , and fimbriae-encoding bacteria. <i>Science Advances</i> , 2019, 5, eaax5727.                                           | 4.7  | 120       |
| 70 | Charting the Complexity of the Marine Microbiome through Single-Cell Genomics. <i>Cell</i> , 2019, 179, 1623-1635.e11.                                                                                                   | 13.5 | 158       |
| 71 | DeepCoilâ€”a fast and accurate prediction of coiled-coil domains in protein sequences. <i>Bioinformatics</i> , 2019, 35, 2790-2795.                                                                                      | 1.8  | 131       |
| 72 | OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. <i>Nucleic Acids Research</i> , 2019, 47, D807-D811.        | 6.5  | 715       |
| 73 | MMseqs2 desktop and local web server app for fast, interactive sequence searches. <i>Bioinformatics</i> , 2019, 35, 2856-2858.                                                                                           | 1.8  | 321       |
| 74 | Analysis of several key factors influencing deep learning-based inter-residue contact prediction. <i>Bioinformatics</i> , 2020, 36, 1091-1098.                                                                           | 1.8  | 29        |

| #  | ARTICLE                                                                                                                                                                                                                                           | IF  | CITATIONS |
|----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 75 | Insilico Alpha-Helical Structural Recognition of Temporin Antimicrobial Peptides and Its Interactions with Middle East Respiratory Syndrome-Coronavirus. <i>International Journal of Peptide Research and Therapeutics</i> , 2020, 26, 1473-1483. | 0.9 | 20        |
| 76 | Graph Theory-Based Sequence Descriptors as Remote Homology Predictors. <i>Biomolecules</i> , 2020, 10, 26.                                                                                                                                        | 1.8 | 13        |
| 77 | Expression of Heat Shock Proteins in Thermally Challenged Pacific Abalone <i>Haliotis discus hannai</i> . <i>Genes</i> , 2020, 11, 22.                                                                                                            | 1.0 | 17        |
| 78 | Evolution of a major virion protein of the giant pandoraviruses from an inactivated bacterial glycoside hydrolase. <i>Virus Evolution</i> , 2020, 6, veaa059.                                                                                     | 2.2 | 22        |
| 79 | ADEPT: a domain independent sequence alignment strategy for gpu architectures. <i>BMC Bioinformatics</i> , 2020, 21, 406.                                                                                                                         | 1.2 | 19        |
| 80 | Coral-Associated Viral Assemblages From the Central Red Sea Align With Host Species and Contribute to Holobiont Genetic Diversity. <i>Frontiers in Microbiology</i> , 2020, 11, 572534.                                                           | 1.5 | 16        |
| 81 | Redox stratification within cryoconite granules influences the nitrogen cycle on glaciers. <i>FEMS Microbiology Ecology</i> , 2020, 96, .                                                                                                         | 1.3 | 19        |
| 82 | Conserved and Opposite Transcriptome Patterns during Germination in <i>Hordeum vulgare</i> and <i>Arabidopsis thaliana</i> . <i>International Journal of Molecular Sciences</i> , 2020, 21, 7404.                                                 | 1.8 | 6         |
| 83 | Seeker: alignment-free identification of bacteriophage genomes by deep learning. <i>Nucleic Acids Research</i> , 2020, 48, e121-e121.                                                                                                             | 6.5 | 78        |
| 84 | Phylogeny and Structure of Fatty Acid Photodecarboxylases and Glucose-Methanol-Choline Oxidoreductases. <i>Catalysts</i> , 2020, 10, 1072.                                                                                                        | 1.6 | 16        |
| 85 | Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome. <i>Nature Microbiology</i> , 2020, 5, 1262-1270.                                                                                                           | 5.9 | 156       |
| 86 | Genomic Analysis of Diverse Members of the Fungal Genus <i>Monosporascus</i> Reveals Novel Lineages, Unique Genome Content and a Potential Bacterial Associate. <i>G3: Genes, Genomes, Genetics</i> , 2020, 10, 2573-2583.                        | 0.8 | 5         |
| 87 | Groundwater <i>Elusimicrobia</i> are metabolically diverse compared to gut microbiome <i>Elusimicrobia</i> and some have a novel nitrogenase paralog. <i>ISME Journal</i> , 2020, 14, 2907-2922.                                                  | 4.4 | 51        |
| 88 | Identifying Genes Involved in Alkaloid Biosynthesis in <i>Vinca minor</i> through Transcriptomics and Gene Co-Expression Analysis. <i>Biomolecules</i> , 2020, 10, 1595.                                                                          | 1.8 | 12        |
| 89 | Metagenomic Association Analysis of Gut Symbiont <i>Limosilactobacillus reuteri</i> Without Host-Specific Genome Isolation. <i>Frontiers in Microbiology</i> , 2020, 11, 585622.                                                                  | 1.5 | 4         |
| 90 | Structural annotation of the conserved carbohydrate esterase vb_24B_21 from Shiga toxin-encoding bacteriophage Î¶24B. <i>Journal of Structural Biology</i> , 2020, 212, 107596.                                                                   | 1.3 | 2         |
| 91 | DRAM for distilling microbial metabolism to automate the curation of microbiome function. <i>Nucleic Acids Research</i> , 2020, 48, 8883-8900.                                                                                                    | 6.5 | 410       |
| 92 | Modular prophage interactions driven by capsule serotype select for capsule loss under phage predation. <i>ISME Journal</i> , 2020, 14, 2980-2996.                                                                                                | 4.4 | 43        |

| #   | ARTICLE                                                                                                                                                                                             | IF   | CITATIONS |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 93  | Machine-learning approach expands the repertoire of anti-CRISPR protein families. <i>Nature Communications</i> , 2020, 11, 3784.                                                                    | 5.8  | 64        |
| 94  | Progress in quickly finding orthologs as reciprocal best hits: comparing blast, last, diamond and MMseqs2. <i>BMC Genomics</i> , 2020, 21, 741.                                                     | 1.2  | 43        |
| 95  | Accurate reconstruction of bacterial pan- and core genomes with PEPPAN. <i>Genome Research</i> , 2020, 30, 1667-1679.                                                                               | 2.4  | 56        |
| 96  | Evolutionary Biology – A Transdisciplinary Approach. , 2020, , .                                                                                                                                    |      | 5         |
| 97  | Uncovering the genomic potential of the Amazon River microbiome to degrade rainforest organic matter. <i>Microbiome</i> , 2020, 8, 151.                                                             | 4.9  | 18        |
| 99  | RRE-Finder: a Genome-Mining Tool for Class-Independent RiPP Discovery. <i>MSystems</i> , 2020, 5, .                                                                                                 | 1.7  | 60        |
| 100 | Sequencing of five poultry strains elucidates phylogenetic relationships and divergence in virulence genes in <i>Morganella morganii</i> . <i>BMC Genomics</i> , 2020, 21, 579.                     | 1.2  | 8         |
| 101 | Diversity and classification of cyclic-oligonucleotide-based anti-phage signalling systems. <i>Nature Microbiology</i> , 2020, 5, 1608-1615.                                                        | 5.9  | 160       |
| 102 | A freshwater radiation of diplomonads. <i>Environmental Microbiology</i> , 2020, 22, 4658-4668.                                                                                                     | 1.8  | 17        |
| 103 | Diverse enzymatic activities mediate antiviral immunity in prokaryotes. <i>Science</i> , 2020, 369, 1077-1084.                                                                                      | 6.0  | 302       |
| 104 | Analytical Performance Validation of Next-Generation Sequencing Based Clinical Microbiology Assays Using a K-mer Analysis Workflow. <i>Frontiers in Microbiology</i> , 2020, 11, 1883.              | 1.5  | 12        |
| 105 | Potential causes and consequences of rapid mitochondrial genome evolution in thermoacidophilic <i>Galdieria</i> (Rhodophyta). <i>BMC Evolutionary Biology</i> , 2020, 20, 112.                      | 3.2  | 13        |
| 106 | Large freshwater phages with the potential to augment aerobic methane oxidation. <i>Nature Microbiology</i> , 2020, 5, 1504-1515.                                                                   | 5.9  | 66        |
| 107 | Discovery of Novel Biosynthetic Gene Cluster Diversity From a Soil Metagenomic Library. <i>Frontiers in Microbiology</i> , 2020, 11, 585398.                                                        | 1.5  | 16        |
| 112 | Analysis of Spatio-Temporal Transcriptome Profiles of Soybean ( <i>Glycine max</i> ) Tissues during Early Seed Development. <i>International Journal of Molecular Sciences</i> , 2020, 21, 7603.    | 1.8  | 9         |
| 113 | Bacterial Retrons Function In Anti-Phage Defense. <i>Cell</i> , 2020, 183, 1551-1561.e12.                                                                                                           | 13.5 | 208       |
| 114 | Comparative Genome Analysis of 33 <i>Chlamydia</i> Strains Reveals Characteristic Features of <i>Chlamydia Psittaci</i> and Closely Related Species. <i>Pathogens</i> , 2020, 9, 899.               | 1.2  | 24        |
| 115 | Amino Acid k-mer Feature Extraction for Quantitative Antimicrobial Resistance (AMR) Prediction by Machine Learning and Model Interpretation for Biological Insights. <i>Biology</i> , 2020, 9, 365. | 1.3  | 21        |

| #   | ARTICLE                                                                                                                                                                                                                | IF  | CITATIONS |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 116 | Comparative (Within Species) Genomics of the <i>Vitis vinifera</i> L. Terpene Synthase Family to Explore the Impact of Genotypic Variation Using Phased Diploid Genomes. <i>Frontiers in Genetics</i> , 2020, 11, 421. | 1.1 | 8         |
| 117 | Terminating contamination: large-scale search identifies more than 2,000,000 contaminated entries in GenBank. <i>Genome Biology</i> , 2020, 21, 115.                                                                   | 3.8 | 130       |
| 118 | Similar yet different: phylogenomic analysis to delineate <i>Salmonella</i> and <i>Citrobacter</i> species boundaries. <i>BMC Genomics</i> , 2020, 21, 377.                                                            | 1.2 | 15        |
| 119 | Catalase and Ascorbate Peroxidase in Euglenozoan Protists. <i>Pathogens</i> , 2020, 9, 317.                                                                                                                            | 1.2 | 12        |
| 120 | EnzymeMiner: automated mining of soluble enzymes with diverse structures, catalytic properties and stabilities. <i>Nucleic Acids Research</i> , 2020, 48, W104-W109.                                                   | 6.5 | 51        |
| 121 | Phylogenetic background and habitat drive the genetic diversification of <i>Escherichia coli</i> . <i>PLoS Genetics</i> , 2020, 16, e1008866.                                                                          | 1.5 | 131       |
| 122 | The parallelism motifs of genomic data analysis. <i>Philosophical Transactions Series A, Mathematical, Physical, and Engineering Sciences</i> , 2020, 378, 20190394.                                                   | 1.6 | 11        |
| 123 | ATLAS: a Snakemake workflow for assembly, annotation, and genomic binning of metagenome sequence data. <i>BMC Bioinformatics</i> , 2020, 21, 257.                                                                      | 1.2 | 91        |
| 124 | Unprecedented Diversity of Unique CRISPR-Cas-Related Systems and Cas1 Homologs in Asgard Archaea. <i>CRISPR Journal</i> , 2020, 3, 156-163.                                                                            | 1.4 | 17        |
| 125 | ProtPCV: A Fixed Dimensional Numerical Representation of Protein Sequence to Significantly Reduce Sequence Search Time. <i>Interdisciplinary Sciences, Computational Life Sciences</i> , 2020, 12, 276-287.            | 2.2 | 3         |
| 126 | De Novo Clustering of Long-Read Transcriptome Data Using a Greedy, Quality Value-Based Algorithm. <i>Journal of Computational Biology</i> , 2020, 27, 472-484.                                                         | 0.8 | 55        |
| 127 | Synthesis and future directions. , 2020, , 222-226.                                                                                                                                                                    |     | 0         |
| 128 | ClusPro in rounds 38 to 45 of CAPRI: Toward combining template-based methods with free docking. <i>Proteins: Structure, Function and Bioinformatics</i> , 2020, 88, 1082-1090.                                         | 1.5 | 5         |
| 129 | PPanGGOLIN: Depicting microbial diversity via a partitioned pangenome graph. <i>PLoS Computational Biology</i> , 2020, 16, e1007732.                                                                                   | 1.5 | 102       |
| 130 | Learning supervised embeddings for large scale sequence comparisons. <i>PLoS ONE</i> , 2020, 15, e0216636.                                                                                                             | 1.1 | 2         |
| 131 | A Family of T6SS Antibacterial Effectors Related to L,d-Transpeptidases Targets the Peptidoglycan. <i>Cell Reports</i> , 2020, 31, 107813.                                                                             | 2.9 | 39        |
| 132 | Broccoli: Combining Phylogenetic and Network Analyses for Orthology Assignment. <i>Molecular Biology and Evolution</i> , 2020, 37, 3389-3396.                                                                          | 3.5 | 65        |
| 133 | Prickly waterlily and rigid hornwort genomes shed light on early angiosperm evolution. <i>Nature Plants</i> , 2020, 6, 215-222.                                                                                        | 4.7 | 88        |



| #   | ARTICLE                                                                                                                                                                                                                        | IF   | CITATIONS |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 134 | On the evolution of protein-adenine binding. Proceedings of the National Academy of Sciences of the United States of America, 2020, 117, 4701-4709.                                                                            | 3.3  | 28        |
| 135 | Ab Initio Construction and Evolutionary Analysis of Protein-Coding Gene Families with Partially Homologous Relationships: Closely Related Drosophila Genomes as a Case Study. Genome Biology and Evolution, 2020, 12, 185-202. | 1.1  | 2         |
| 136 | RNAse: neural network-based assessment of the protein-coding potential of RNA sequences. NAR Genomics and Bioinformatics, 2020, 2, lqz024.                                                                                     | 1.5  | 65        |
| 137 | FeGenie: A Comprehensive Tool for the Identification of Iron Genes and Iron Gene Neighborhoods in Genome and Metagenome Assemblies. Frontiers in Microbiology, 2020, 11, 37.                                                   | 1.5  | 195       |
| 138 | Transcription factor AP2 controls cnidarian germ cell induction. Science, 2020, 367, 757-762.                                                                                                                                  | 6.0  | 58        |
| 139 | Novel phylogenetic methods are needed for understanding gene function in the era of mega-scale genome sequencing. Nucleic Acids Research, 2020, 48, 2209-2219.                                                                 | 6.5  | 32        |
| 140 | Deep learning methods in protein structure prediction. Computational and Structural Biotechnology Journal, 2020, 18, 1301-1310.                                                                                                | 1.9  | 169       |
| 141 | Transcriptome reconstruction and functional analysis of eukaryotic marine plankton communities via high-throughput metagenomics and metatranscriptomics. Genome Research, 2020, 30, 647-659.                                   | 2.4  | 50        |
| 142 | The Pangenome. , 2020, , .                                                                                                                                                                                                     |      | 32        |
| 143 | Diversity and evolution of surface polysaccharide synthesis loci in Enterobacteriales. ISME Journal, 2020, 14, 1713-1730.                                                                                                      | 4.4  | 40        |
| 144 | The Remarkable Dual-Level Diversity of Prokaryotic Flagellins. MSystems, 2020, 5, .                                                                                                                                            | 1.7  | 11        |
| 145 | Essential dynamic interdependence of FtsZ and SepF for Z-ring and septum formation in <i>Corynebacterium glutamicum</i> . Nature Communications, 2020, 11, 1641.                                                               | 5.8  | 29        |
| 146 | Canfam_GSD: De novo chromosome-length genome assembly of the German Shepherd Dog ( <i>Canis lupus</i> ) Tj ETQq0,0,0 rgBT /Overlock I                                                                                          | 3.3  | 47        |
| 147 | MetaEuk-sensitive, high-throughput gene discovery, and annotation for large-scale eukaryotic metagenomics. Microbiome, 2020, 8, 48.                                                                                            | 4.9  | 119       |
| 148 | Prokaryotic viperins produce diverse antiviral molecules. Nature, 2021, 589, 120-124.                                                                                                                                          | 13.7 | 172       |
| 149 | Amoeba Genome Reveals Dominant Host Contribution to Plastid Endosymbiosis. Molecular Biology and Evolution, 2021, 38, 344-357.                                                                                                 | 3.5  | 23        |
| 150 | Life-history strategies of soil microbial communities in an arid ecosystem. ISME Journal, 2021, 15, 649-657.                                                                                                                   | 4.4  | 84        |
| 151 | The genome of <i>Draba nivalis</i> shows signatures of adaptation to the extreme environmental stresses of the Arctic. Molecular Ecology Resources, 2021, 21, 661-676.                                                         | 2.2  | 14        |

| #   | ARTICLE                                                                                                                                                                                                          | IF  | CITATIONS |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 152 | Phylogeny and biogeography of <i>Fagus</i> (Fagaceae) based on 28 nuclear single-copy loci. <i>Journal of Systematics and Evolution</i> , 2022, 60, 759-772.                                                     | 1.6 | 15        |
| 153 | RCSB Protein Data Bank: Architectural Advances Towards Integrated Searching and Efficient Access to Macromolecular Structure Data from the PDB Archive. <i>Journal of Molecular Biology</i> , 2021, 433, 166704. | 2.0 | 106       |
| 154 | OrthoDB in 2020: evolutionary and functional annotations of orthologs. <i>Nucleic Acids Research</i> , 2021, 49, D389-D393.                                                                                      | 6.5 | 103       |
| 155 | Probing periodontal microbial dark matter using metataxonomics and metagenomics. <i>Periodontology</i> 2000, 2021, 85, 12-27.                                                                                    | 6.3 | 16        |
| 156 | Time-calibrated genomic evolution of a monomorphic bacterium during its establishment as an endemic crop pathogen. <i>Molecular Ecology</i> , 2021, 30, 1823-1835.                                               | 2.0 | 9         |
| 157 | DescribePROT: database of amino acid-level protein structure and function predictions. <i>Nucleic Acids Research</i> , 2021, 49, D298-D308.                                                                      | 6.5 | 46        |
| 158 | Embryo-Like Features in Developing <i>Bacillus subtilis</i> Biofilms. <i>Molecular Biology and Evolution</i> , 2021, 38, 31-47.                                                                                  | 3.5 | 25        |
| 159 | Evolution of DNA packaging in gene transfer agents. <i>Virus Evolution</i> , 2021, 7, veab015.                                                                                                                   | 2.2 | 36        |
| 160 | Expansion and re-classification of the extracytoplasmic function (ECF) $\sigma$ factor family. <i>Nucleic Acids Research</i> , 2021, 49, 986-1005.                                                               | 6.5 | 32        |
| 161 | Identification and classification of antiviral defence systems in bacteria and archaea with PADLOC reveals new system types. <i>Nucleic Acids Research</i> , 2021, 49, 10868-10878.                              | 6.5 | 92        |
| 162 | Rapid discovery of novel prophages using biological feature engineering and machine learning. <i>NAR Genomics and Bioinformatics</i> , 2021, 3, lqaa109.                                                         | 1.5 | 39        |
| 164 | Harnessing CRISPR-Cas system diversity for gene editing technologies. <i>Journal of Biomedical Research</i> , 2021, 35, 91.                                                                                      | 0.7 | 1         |
| 168 | Improved DNA-versus-Protein Homology Search for Protein Fossils. <i>Lecture Notes in Computer Science</i> , 2021, , 146-158.                                                                                     | 1.0 | 4         |
| 169 | The complexity of protein interactions unravelled from structural disorder. <i>PLoS Computational Biology</i> , 2021, 17, e1008546.                                                                              | 1.5 | 15        |
| 172 | ProtTrans: Toward Understanding the Language of Life Through Self-Supervised Learning. <i>IEEE Transactions on Pattern Analysis and Machine Intelligence</i> , 2022, 44, 7112-7127.                              | 9.7 | 496       |
| 173 | Distinct clonal lineages and within-host diversification shape invasive <i>Staphylococcus epidermidis</i> populations. <i>PLoS Pathogens</i> , 2021, 17, e1009304.                                               | 2.1 | 41        |
| 174 | Balrog: A universal protein model for prokaryotic gene prediction. <i>PLoS Computational Biology</i> , 2021, 17, e1008727.                                                                                       | 1.5 | 21        |
| 175 | Detecting high-scoring local alignments in pangenome graphs. <i>Bioinformatics</i> , 2021, 37, 2266-2274.                                                                                                        | 1.8 | 6         |

| #   | ARTICLE                                                                                                                                                                                                         | IF  | CITATIONS |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 176 | Generating functional protein variants with variational autoencoders. <i>PLoS Computational Biology</i> , 2021, 17, e1008736.                                                                                   | 1.5 | 106       |
| 177 | Causes and Consequences of Bacteriophage Diversification via Genetic Exchanges across Lifestyles and Bacterial Taxa. <i>Molecular Biology and Evolution</i> , 2021, 38, 2497-2512.                              | 3.5 | 48        |
| 182 | Genome-scale phylogenetic analyses confirm <i>Olpidium</i> as the closest living zoosporic fungus to the non-flagellated, terrestrial fungi. <i>Scientific Reports</i> , 2021, 11, 3217.                        | 1.6 | 24        |
| 183 | VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. <i>Microbiome</i> , 2021, 9, 37.                                                                                  | 4.9 | 441       |
| 185 | Bacteria have numerous distinctive groups of phage-plasmids with conserved phage and variable plasmid gene repertoires. <i>Nucleic Acids Research</i> , 2021, 49, 2655-2673.                                    | 6.5 | 91        |
| 188 | The Genome of Banana Leaf Blight Pathogen <i>Fusarium sacchari</i> str. FS66 Harbors Widespread Gene Transfer From <i>Fusarium oxysporum</i> . <i>Frontiers in Plant Science</i> , 2021, 12, 629859.            | 1.7 | 5         |
| 189 | A large genome with chromosome-scale assembly sheds light on the evolutionary success of a true toad ( <i>Bufo gargarizans</i> ). <i>Molecular Ecology Resources</i> , 2021, 21, 1256-1273.                     | 2.2 | 32        |
| 190 | Overview of bioinformatic methods for analysis of antibiotic resistome from genome and metagenome data. <i>Journal of Microbiology</i> , 2021, 59, 270-280.                                                     | 1.3 | 16        |
| 191 | Highly multiplexed oligonucleotide probe-ligation testing enables efficient extraction-free SARS-CoV-2 detection and viral genotyping. <i>Modern Pathology</i> , 2021, 34, 1093-1103.                           | 2.9 | 12        |
| 193 | A chromosome-level assembly of the black tiger shrimp ( <i>Penaeus monodon</i> ) genome facilitates the identification of growth-associated genes. <i>Molecular Ecology Resources</i> , 2021, 21, 1620-1640.    | 2.2 | 43        |
| 194 | An international virtual hackathon to build tools for the analysis of structural variants within species ranging from coronaviruses to vertebrates. <i>F1000Research</i> , 2021, 10, 246.                       | 0.8 | 3         |
| 196 | A chromosome-level genome of the spider <i>Trichonephila antipodiana</i> reveals the genetic basis of its polyphagy and evidence of an ancient whole-genome duplication event. <i>GigaScience</i> , 2021, 10, . | 3.3 | 187       |
| 197 | Chromosome-length genome assembly and structural variations of the primal Basenji dog ( <i>Canis lupus</i> ) Tj ETQq0 0.0rgBT /Overlock 10                                                                      | 1.2 | 22        |
| 198 | Expanding functional protein sequence spaces using generative adversarial networks. <i>Nature Machine Intelligence</i> , 2021, 3, 324-333.                                                                      | 8.3 | 165       |
| 200 | Ultrafast functional profiling of RNA-seq data for nonmodel organisms. <i>Genome Research</i> , 2021, 31, 713-720.                                                                                              | 2.4 | 15        |
| 201 | Accurate prediction of inter-protein residue-residue contacts for homo-oligomeric protein complexes. <i>Briefings in Bioinformatics</i> , 2021, 22, .                                                           | 3.2 | 46        |
| 202 | Current State and Future Directions of Genetics and Genomics of Endophytic Fungi for Bioprospecting Efforts. <i>Frontiers in Bioengineering and Biotechnology</i> , 2021, 9, 649906.                            | 2.0 | 23        |
| 204 | Introducing EzAAI: a pipeline for high throughput calculations of prokaryotic average amino acid identity. <i>Journal of Microbiology</i> , 2021, 59, 476-480.                                                  | 1.3 | 155       |

| #   | ARTICLE                                                                                                                                                                                                                      | IF   | CITATIONS |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 207 | Quantifying evolutionary importance of protein sites: A Tale of two measures. PLoS Genetics, 2021, 17, e1009476.                                                                                                             | 1.5  | 8         |
| 209 | De novo Genome Assembly of the Raccoon Dog ( <i>Nyctereutes procyonoides</i> ). Frontiers in Genetics, 2021, 12, 658256.                                                                                                     | 1.1  | 11        |
| 210 | On the use of direct-coupling analysis with a reduced alphabet of amino acids combined with super-secondary structure motifs for protein fold prediction. NAR Genomics and Bioinformatics, 2021, 3, lqab027.                 | 1.5  | 0         |
| 211 | Expanded diversity of Asgard archaea and their relationships with eukaryotes. Nature, 2021, 593, 553-557.                                                                                                                    | 13.7 | 161       |
| 213 | SpacePHARER: sensitive identification of phages from CRISPR spacers in prokaryotic hosts. Bioinformatics, 2021, 37, 3364-3366.                                                                                               | 1.8  | 36        |
| 214 | Putative plasmid prophages of <i>Bacillus cereus sensu lato</i> may hold the key to undiscovered phage diversity. Scientific Reports, 2021, 11, 7611.                                                                        | 1.6  | 15        |
| 215 | Dual modes of CRISPR-associated transposon homing. Cell, 2021, 184, 2441-2453.e18.                                                                                                                                           | 13.5 | 86        |
| 216 | Sensitive protein alignments at tree-of-life scale using DIAMOND. Nature Methods, 2021, 18, 366-368.                                                                                                                         | 9.0  | 1,195     |
| 218 | Characterization of Three Pleiotropic Drug Resistance Transporter Genes and Their Participation in the Azole Resistance of <i>Mucor circinelloides</i> . Frontiers in Cellular and Infection Microbiology, 2021, 11, 660347. | 1.8  | 15        |
| 219 | Common Presence of Phototrophic <i>Gemmatimonadota</i> in Temperate Freshwater Lakes. MSystems, 2021, 6, .                                                                                                                   | 1.7  | 20        |
| 221 | Complementary Contribution of Fungi and Bacteria to Lignocellulose Digestion in the Food Stored by a Neotropical Higher Termite. Frontiers in Ecology and Evolution, 2021, 9, .                                              | 1.1  | 9         |
| 225 | Identification and characterization of andalusicin: N-terminally dimethylated class III lantibiotic from <i>Bacillus thuringiensis</i> sv. <i>andalousiensis</i> . IScience, 2021, 24, 102480.                               | 1.9  | 18        |
| 226 | PredictProtein - Predicting Protein Structure and Function for 29 Years. Nucleic Acids Research, 2021, 49, W535-W540.                                                                                                        | 6.5  | 135       |
| 228 | Unification of functional annotation descriptions using text mining. Biological Chemistry, 2021, 402, 983-990.                                                                                                               | 1.2  | 6         |
| 230 | Gut microbiome contributions to altered metabolism in a pig model of undernutrition. Proceedings of the National Academy of Sciences of the United States of America, 2021, 118, .                                           | 3.3  | 18        |
| 232 | <i>De novo</i> genome assembly of the land snail <i>Candidula unifasciata</i> (Mollusca: Gastropoda). G3: Genes, Genomes, Genetics, 2021, 11, .                                                                              | 0.8  | 8         |
| 233 | Positive Selection in Gene Regulatory Factors Suggests Adaptive Pleiotropic Changes During Human Evolution. Frontiers in Genetics, 2021, 12, 662239.                                                                         | 1.1  | 8         |
| 234 | ELASPIC2 (EL2): Combining Contextualized Language Models and Graph Neural Networks to Predict Effects of Mutations. Journal of Molecular Biology, 2021, 433, 166810.                                                         | 2.0  | 24        |

| #   | ARTICLE                                                                                                                                                                                                                                                              | IF   | CITATIONS |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 237 | SPOT-1D-Single: improving the single-sequence-based prediction of protein secondary structure, backbone angles, solvent accessibility and half-sphere exposures using a large training set and ensembled deep learning. <i>Bioinformatics</i> , 2021, 37, 3464-3472. | 1.8  | 22        |
| 238 | PHROG: families of prokaryotic virus proteins clustered using remote homology. <i>NAR Genomics and Bioinformatics</i> , 2021, 3, lqab067.                                                                                                                            | 1.5  | 134       |
| 239 | Wheat Rhizosphere Metagenome Reveals Newfound Potential Soil Zn-Mobilizing Bacteria Contributing to Cultivarsâ€™ Variation in Grain Zn Concentration. <i>Frontiers in Microbiology</i> , 2021, 12, 689855.                                                           | 1.5  | 4         |
| 242 | No Evidence Known Viruses Play a Role in the Pathogenesis of Onchocerciasis-Associated Epilepsy. An Explorative Metagenomic Case-Control Study. <i>Pathogens</i> , 2021, 10, 787.                                                                                    | 1.2  | 7         |
| 243 | <i>RAPID</i> omics v2.0: going deeper and wider in the prediction of proteinâ€“RNA interactions. <i>Nucleic Acids Research</i> , 2021, 49, W72-W79.                                                                                                                  | 6.5  | 79        |
| 244 | Bioinformatics Accelerates the Major Tetrad: A Real Boost for the Pharmaceutical Industry. <i>International Journal of Molecular Sciences</i> , 2021, 22, 6184.                                                                                                      | 1.8  | 16        |
| 246 | Mosaic Ends Tagmentation (METa) Assembly for Highly Efficient Construction of Functional Metagenomic Libraries. <i>MSystems</i> , 2021, 6, e0052421.                                                                                                                 | 1.7  | 2         |
| 247 | A Comparative Analysis on the Structure and Function of the <i>Panax notoginseng</i> Rhizosphere Microbiome. <i>Frontiers in Microbiology</i> , 2021, 12, 673512.                                                                                                    | 1.5  | 16        |
| 248 | Computational generation of proteins with predetermined three-dimensional shapes using ProteinSolver. <i>STAR Protocols</i> , 2021, 2, 100505.                                                                                                                       | 0.5  | 5         |
| 251 | Spinal Cord Injury Changes the Structure and Functional Potential of Gut Bacterial and Viral Communities. <i>MSystems</i> , 2021, 6, .                                                                                                                               | 1.7  | 28        |
| 252 | The Evolution of Interdependence in a Four-Way Mealybug Symbiosis. <i>Genome Biology and Evolution</i> , 2021, 13, .                                                                                                                                                 | 1.1  | 9         |
| 253 | Ecology of inorganic sulfur auxiliary metabolism in widespread bacteriophages. <i>Nature Communications</i> , 2021, 12, 3503.                                                                                                                                        | 5.8  | 97        |
| 255 | A catalog of the diversity and ubiquity of bacterial microcompartments. <i>Nature Communications</i> , 2021, 12, 3809.                                                                                                                                               | 5.8  | 55        |
| 256 | Characterization of Bacterial Community Dynamics of the Human Mouth Throughout Decomposition via Metagenomic, Metatranscriptomic, and Culturing Techniques. <i>Frontiers in Microbiology</i> , 2021, 12, 689493.                                                     | 1.5  | 7         |
| 258 | Going to extremes â€“ a metagenomic journey into the dark matter of life. <i>FEMS Microbiology Letters</i> , 2021, 368, .                                                                                                                                            | 0.7  | 16        |
| 260 | pKPDB: a protein data bank extension database of p <i>Ka</i> and p <i>I</i> theoretical values. <i>Bioinformatics</i> , 2021, 38, 297-298.                                                                                                                           | 1.8  | 6         |
| 262 | RaFAH: Host prediction for viruses of Bacteria and Archaea based on protein content. <i>Patterns</i> , 2021, 2, 100274.                                                                                                                                              | 3.1  | 53        |
| 263 | Highly accurate protein structure prediction with AlphaFold. <i>Nature</i> , 2021, 596, 583-589.                                                                                                                                                                     | 13.7 | 17,754    |

| #   | ARTICLE                                                                                                                                                                                                                    | IF  | CITATIONS |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 264 | An astonishing wealth of new proteasome homologs. <i>Bioinformatics</i> , 2021, 37, 4694-4703.                                                                                                                             | 1.8 | 3         |
| 265 | PZLAST: an ultra-fast amino acid sequence similarity search server against public metagenomes. <i>Bioinformatics</i> , 2021, 37, 3944-3946.                                                                                | 1.8 | 5         |
| 266 | Spectrum of Protein Location in Proteomes Captures Evolutionary Relationship Between Species. <i>Journal of Molecular Evolution</i> , 2021, 89, 544-553.                                                                   | 0.8 | 4         |
| 267 | Metatranscriptomics for the Human Microbiome and Microbial Community Functional Profiling. <i>Annual Review of Biomedical Data Science</i> , 2021, 4, 279-311.                                                             | 2.8 | 36        |
| 268 | Nationwide genomic atlas of soil-dwelling <i>Listeria</i> reveals effects of selection and population ecology on pangenome evolution. <i>Nature Microbiology</i> , 2021, 6, 1021-1030.                                     | 5.9 | 54        |
| 269 | Interplay between the cell envelope and mobile genetic elements shapes gene flow in populations of the nosocomial pathogen <i>Klebsiella pneumoniae</i> . <i>PLoS Biology</i> , 2021, 19, e3001276.                        | 2.6 | 44        |
| 271 | Increasing the accuracy of single sequence prediction methods using a deep semi-supervised learning framework. <i>Bioinformatics</i> , 2021, 37, 3744-3751.                                                                | 1.8 | 24        |
| 274 | Dispersal strategies shape persistence and evolution of human gut bacteria. <i>Cell Host and Microbe</i> , 2021, 29, 1167-1176.e9.                                                                                         | 5.1 | 66        |
| 275 | tsRNAsearch: a pipeline for the identification of tRNA and ncRNA fragments from small RNA-sequencing data. <i>Bioinformatics</i> , 2021, 37, 4424-4430.                                                                    | 1.8 | 6         |
| 277 | Public good exploitation in natural bacterioplankton communities. <i>Science Advances</i> , 2021, 7, .                                                                                                                     | 4.7 | 22        |
| 278 | Comparative genome and transcriptome analyses reveal innate differences in response to host plants by two color forms of the two-spotted spider mite <i>Tetranychus urticae</i> . <i>BMC Genomics</i> , 2021, 22, 569.     | 1.2 | 4         |
| 279 | Genetic and phylogenetic analysis of dissimilatory iodate-reducing bacteria identifies potential niches across the world's oceans. <i>ISME Journal</i> , 2022, 16, 38-49.                                                  | 4.4 | 21        |
| 281 | Genomic diversity and ecology of human-associated <i>Akkermansia</i> species in the gut microbiome revealed by extensive metagenomic assembly. <i>Genome Biology</i> , 2021, 22, 209.                                      | 3.8 | 65        |
| 282 | BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. <i>Molecular Biology and Evolution</i> , 2021, 38, 4647-4654. | 3.5 | 1,968     |
| 283 | Technology dictates algorithms: recent developments in read alignment. <i>Genome Biology</i> , 2021, 22, 249.                                                                                                              | 3.8 | 51        |
| 284 | GNNfam. , 2021, , .                                                                                                                                                                                                        |     | 0         |
| 288 | Protein tertiary structure prediction and refinement using deep learning and Rosetta in <sc>CASP14</sc>. <i>Proteins: Structure, Function and Bioinformatics</i> , 2021, 89, 1722-1733.                                    | 1.5 | 40        |
| 291 | Transposable elements and introgression introduce genetic variation in the invasive ant <i>Cardiocondyla obscurior</i>. <i>Molecular Ecology</i> , 2021, 30, 6211-6228.                                                    | 2.0 | 20        |

| #   | ARTICLE                                                                                                                                                                                    | IF  | CITATIONS |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 295 | <scp>TMEM106B</scp> in humans and Vac7 and Tag1 in yeast are predicted to be lipid transfer proteins. <i>Proteins: Structure, Function and Bioinformatics</i> , 2022, 90, 164-175.         | 1.5 | 13        |
| 297 | A Genomic Catalog of Stress Response Genes in Anaerobic Fungi for Applications in Bioproduction. <i>Frontiers in Fungal Biology</i> , 2021, 2, .                                           | 0.9 | 1         |
| 298 | Functional meta-omics provide critical insights into long- and short-read assemblies. <i>Briefings in Bioinformatics</i> , 2021, 22, .                                                     | 3.2 | 6         |
| 301 | Compact RNA editors with small Cas13 proteins. <i>Nature Biotechnology</i> , 2022, 40, 194-197.                                                                                            | 9.4 | 86        |
| 302 | Closely related Lak megaphages replicate in the microbiomes of diverse animals. <i>IScience</i> , 2021, 24, 102875.                                                                        | 1.9 | 20        |
| 304 | MOSGA 2: Comparative genomics and validation tools. <i>Computational and Structural Biotechnology Journal</i> , 2021, 19, 5504-5509.                                                       | 1.9 | 4         |
| 306 | Remote homology clustering identifies lowly conserved families of effector proteins in plant-pathogenic fungi. <i>Microbial Genomics</i> , 2021, 7, .                                      | 1.0 | 7         |
| 307 | Prevalence and Specificity of Chemoreceptor Profiles in Plant-Associated Bacteria. <i>MSystems</i> , 2021, 6, e0095121.                                                                    | 1.7 | 20        |
| 309 | OrtSuite: from genomes to prediction of microbial interactions within targeted ecosystem processes. <i>Life Science Alliance</i> , 2021, 4, e202101167.                                    | 1.3 | 4         |
| 310 | Organelle Genome Variation in the Red Algal Genus <i>Ahnfeltia</i> (Florideophyceae). <i>Frontiers in Genetics</i> , 2021, 12, 724734.                                                     | 1.1 | 0         |
| 311 | Rossmann-toolbox: a deep learning-based protocol for the prediction and design of cofactor specificity in Rossmann fold proteins. <i>Briefings in Bioinformatics</i> , 2022, 23, .         | 3.2 | 14        |
| 313 | Altered microbial CAZyme families indicated dead biomass decomposition following afforestation. <i>Soil Biology and Biochemistry</i> , 2021, 160, 108362.                                  | 4.2 | 17        |
| 314 | Prokaryotic viruses impact functional microorganisms in nutrient removal and carbon cycle in wastewater treatment plants. <i>Nature Communications</i> , 2021, 12, 5398.                   | 5.8 | 49        |
| 315 | Struo2: efficient metagenome profiling database construction for ever-expanding microbial genome datasets. <i>PeerJ</i> , 2021, 9, e12198.                                                 | 0.9 | 24        |
| 316 | The gRAMP CRISPR-Cas effector is an RNA endonuclease complexed with a caspase-like peptidase. <i>Science</i> , 2021, 373, 1349-1353.                                                       | 6.0 | 76        |
| 318 | Structural characterization of NrnC identifies unifying features of dinucleases. <i>ELife</i> , 2021, 10, .                                                                                | 2.8 | 6         |
| 320 | Evolution of Type IV CRISPR-Cas Systems: Insights from CRISPR Loci in Integrative Conjugative Elements of <i>Acidithiobacillus</i> . <i>CRISPR Journal</i> , 2021, 4, 656-672.             | 1.4 | 21        |
| 322 | The first chromosome-level genome assembly of a green lacewing <i>Chrysopa pallens</i> and its implication for biological control. <i>Molecular Ecology Resources</i> , 2022, 22, 755-767. | 2.2 | 7         |

| #   | ARTICLE                                                                                                                                                                                                              | IF   | CITATIONS |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 323 | An international virtual hackathon to build tools for the analysis of structural variants within a species ranging from coronaviruses to vertebrates. <i>F1000Research</i> , 2021, 10, 246.                          | 0.8  | 2         |
| 324 | A guide to machine learning for biologists. <i>Nature Reviews Molecular Cell Biology</i> , 2022, 23, 40-55.                                                                                                          | 16.1 | 626       |
| 326 | Multi-Omics Analyses Show Disease, Diet, and Transcriptome Interactions With the Virome. <i>Gastroenterology</i> , 2021, 161, 1194-1207.e8.                                                                          | 0.6  | 28        |
| 327 | The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided endonucleases. <i>Science</i> , 2021, 374, 57-65.                                                                               | 6.0  | 152       |
| 329 | Phylogenomics of the <i>Maverick</i> Virus-Like Mobile Genetic Elements of Vertebrates. <i>Molecular Biology and Evolution</i> , 2021, 38, 1731-1743.                                                                | 3.5  | 22        |
| 330 | SoluProt: prediction of soluble protein expression in <i>Escherichia coli</i> . <i>Bioinformatics</i> , 2021, 37, 23-28.                                                                                             | 1.8  | 66        |
| 332 | Accurate Annotation of Microbial Metagenomic Genes and Identification of Core Sets. <i>Methods in Molecular Biology</i> , 2021, 2242, 115-138.                                                                       | 0.4  | 0         |
| 333 | The language of proteins: NLP, machine learning & protein sequences. <i>Computational and Structural Biotechnology Journal</i> , 2021, 19, 1750-1758.                                                                | 1.9  | 175       |
| 334 | ProtRe-CN: Protein Remote Homology Detection by Combining Classification Methods and Network Methods via Learning to Rank. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2021, PP, 1-1. | 1.9  | 2         |
| 335 | The impact of genetic diversity on gene essentiality within the <i>Escherichia coli</i> species. <i>Nature Microbiology</i> , 2021, 6, 301-312.                                                                      | 5.9  | 76        |
| 336 | Prediction of GPI-anchored proteins with pointer neural networks. <i>Current Research in Biotechnology</i> , 2021, 3, 6-13.                                                                                          | 1.9  | 59        |
| 337 | Embeddings from deep learning transfer GO annotations beyond homology. <i>Scientific Reports</i> , 2021, 11, 1160.                                                                                                   | 1.6  | 95        |
| 338 | Inferring Orthology and Paralogy. <i>Methods in Molecular Biology</i> , 2019, 1910, 149-175.                                                                                                                         | 0.4  | 84        |
| 339 | 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.                            | 5.8  | 80        |
| 340 | Verrucomicrobia use hundreds of enzymes to digest the algal polysaccharide fucoidan. <i>Nature Microbiology</i> , 2020, 5, 1026-1039.                                                                                | 5.9  | 182       |
| 341 | CRISPR Arrays Away from <i>cas</i> Genes. <i>CRISPR Journal</i> , 2020, 3, 535-549.                                                                                                                                  | 1.4  | 18        |
| 342 | LAMPA, LARge Multidomain Protein Annotator, and its application to RNA virus polyproteins. <i>Bioinformatics</i> , 2020, 36, 2731-2739.                                                                              | 1.8  | 7         |
| 343 | panRGP: a pangenome-based method to predict genomic islands and explore their diversity. <i>Bioinformatics</i> , 2020, 36, i651-i658.                                                                                | 1.8  | 26        |



| #   | ARTICLE                                                                                                                                                                                                                                                                                       | IF  | CITATIONS |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 344 | RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. <i>Nucleic Acids Research</i> , 2021, 49, D437-D451. | 6.5 | 918       |
| 345 | Systematic prediction of genes functionally associated with bacterial retrons and classification of the encoded tripartite systems. <i>Nucleic Acids Research</i> , 2020, 48, 12632-12647.                                                                                                    | 6.5 | 41        |
| 346 | PanACoTA: a modular tool for massive microbial comparative genomics. <i>NAR Genomics and Bioinformatics</i> , 2021, 3, lqaa106.                                                                                                                                                               | 1.5 | 26        |
| 347 | Computational genomic discovery of diverse gene clusters harbouring Fe-S flavoenzymes in anaerobic gut microbiota. <i>Microbial Genomics</i> , 2020, 6, .                                                                                                                                     | 1.0 | 5         |
| 348 | Phylogenetic distribution and evolutionary dynamics of nod and T3SS genes in the genus <i>Bradyrhizobium</i> . <i>Microbial Genomics</i> , 2020, 6, .                                                                                                                                         | 1.0 | 9         |
| 407 | Distributed Many-to-Many Protein Sequence Alignment using Sparse Matrices. , 2020, , .                                                                                                                                                                                                        |     | 9         |
| 408 | Horizontal gene transfer rate is not the primary determinant of observed antibiotic resistance frequencies in <i>Streptococcus pneumoniae</i> . <i>Science Advances</i> , 2020, 6, eaaz6137.                                                                                                  | 4.7 | 19        |
| 409 | Template-based prediction of protein structure with deep learning. <i>BMC Genomics</i> , 2020, 21, 878.                                                                                                                                                                                       | 1.2 | 15        |
| 410 | PremPS: Predicting the impact of missense mutations on protein stability. <i>PLoS Computational Biology</i> , 2020, 16, e1008543.                                                                                                                                                             | 1.5 | 122       |
| 411 | Peptide-based quorum sensing systems in <i>Paenibacillus polymyxa</i> . <i>Life Science Alliance</i> , 2020, 3, e202000847.                                                                                                                                                                   | 1.3 | 11        |
| 413 | Metage2Metabo, microbiota-scale metabolic complementarity for the identification of key species. <i>ELife</i> , 2020, 9, .                                                                                                                                                                    | 2.8 | 44        |
| 414 | A putative origin of the insect chemosensory receptor superfamily in the last common eukaryotic ancestor. <i>ELife</i> , 2020, 9, .                                                                                                                                                           | 2.8 | 16        |
| 415 | ViroidDB: a database of viroids and viroid-like circular RNAs. <i>Nucleic Acids Research</i> , 2022, 50, D432-D438.                                                                                                                                                                           | 6.5 | 14        |
| 416 | EXAGRAPH: Graph and combinatorial methods for enabling exascale applications. <i>International Journal of High Performance Computing Applications</i> , 2021, 35, 553-571.                                                                                                                    | 2.4 | 9         |
| 417 | SPOT-1D2: Improving Protein Secondary Structure Prediction using High Sequence Identity Training Set and an Ensemble of Recurrent and Residual-convolutional Neural Networks. , 2021, , .                                                                                                     |     | 1         |
| 419 | An automated and combinative method for the predictive ranking of candidate effector proteins of fungal plant pathogens. <i>Scientific Reports</i> , 2021, 11, 19731.                                                                                                                         | 1.6 | 27        |
| 421 | Genome sequencing of the NIES Cyanobacteria collection with a focus on the heterocyst-forming clade. <i>DNA Research</i> , 2021, 28, .                                                                                                                                                        | 1.5 | 12        |
| 422 | DeepREX-WS: A web server for characterising protein-solvent interaction starting from sequence. <i>Computational and Structural Biotechnology Journal</i> , 2021, 19, 5791-5799.                                                                                                              | 1.9 | 4         |

| #   | ARTICLE                                                                                                                                                                                                     | IF   | CITATIONS |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 424 | Comparative Genomics of Mycobacterium avium Complex Reveals Signatures of Environment-Specific Adaptation and Community Acquisition. <i>MSystems</i> , 2021, 6, e0119421.                                   | 1.7  | 7         |
| 425 | Scaling up DNA digital data storage by efficiently predicting DNA hybridisation using deep learning. <i>Scientific Reports</i> , 2021, 11, 20517.                                                           | 1.6  | 6         |
| 426 | Microbial colonization and resistome dynamics in food processing environments of a newly opened pork cutting industry during 1.5 years of activity. <i>Microbiome</i> , 2021, 9, 204.                       | 4.9  | 20        |
| 427 | Cyclic CMP and cyclic UMP mediate bacterial immunity against phages. <i>Cell</i> , 2021, 184, 5728-5739.e16.                                                                                                | 13.5 | 156       |
| 428 | Graph-Based Approaches Significantly Improve the Recovery of Antibiotic Resistance Genes From Complex Metagenomic Datasets. <i>Frontiers in Microbiology</i> , 2021, 12, 714836.                            | 1.5  | 7         |
| 429 | Rapid evolutionary turnover of mobile genetic elements drives bacterial resistance to phages. <i>Science</i> , 2021, 374, 488-492.                                                                          | 6.0  | 96        |
| 431 | Chromosome-Level Genome Assembly of <i>Nephotettix cincticeps</i> (Uhler, 1896) (Hemiptera: Tj ETQq0 0 0 rgBT /Overlock 10 Tf 50                                                                            | 1.1  | 7         |
| 432 | Comprehensive discovery of CRISPR-targeted terminally redundant sequences in the human gut metagenome: Viruses, plasmids, and more. <i>PLoS Computational Biology</i> , 2021, 17, e1009428.                 | 1.5  | 7         |
| 434 | FoldHSphere: deep hyperspherical embeddings for protein fold recognition. <i>BMC Bioinformatics</i> , 2021, 22, 490.                                                                                        | 1.2  | 8         |
| 435 | Vertebrate host phylogeny influences gut archaeal diversity. <i>Nature Microbiology</i> , 2021, 6, 1443-1454.                                                                                               | 5.9  | 34        |
| 456 | Orthologous Revelation between <i>Elaeis guineensis</i> , <i>Arabidopsis thaliana</i> and <i>Solanum lycopersicum</i> . <i>International Journal of Life Sciences and Biotechnology</i> , 2020, 3, 164-179. | 0.2  | 0         |
| 458 | Opfi: A Python package for identifying gene clusters in large genomics and metagenomics data sets. <i>Journal of Open Source Software</i> , 2021, 6, 3678.                                                  | 2.0  | 2         |
| 460 | A High Rate Algal Pond Hosting a Dynamic Community of RNA Viruses. <i>Viruses</i> , 2021, 13, 2163.                                                                                                         | 1.5  | 6         |
| 463 | RNANet: an automatically built dual-source dataset integrating homologous sequences and RNA structures. <i>Bioinformatics</i> , 2021, 37, 1218-1224.                                                        | 1.8  | 7         |
| 464 | De novo transcriptome assembly, functional annotation and characterization of the Atlantic bluefin tuna ( <i>Thunnus thynnus</i> ) larval stage. <i>Marine Genomics</i> , 2021, 58, 100834.                 | 0.4  | 2         |
| 467 | Bacterial Microevolution and the Pangenome. , 2020, , 129-149.                                                                                                                                              |      | 1         |
| 468 | Orthology: Promises and Challenges. , 2020, , 203-228.                                                                                                                                                      |      | 10        |
| 469 | A Practical Guide to Orthology Resources. , 2020, , 41-77.                                                                                                                                                  |      | 2         |

| #   | ARTICLE                                                                                                                                                                                                      | IF  | CITATIONS |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 471 | A Comprehensive Analysis of the Global Human Gut Archaeome from a Thousand Genome Catalogue. SSRN Electronic Journal, 0, , .                                                                                 | 0.4 | 0         |
| 474 | Accurate protein function prediction via graph attention networks with predicted structure information. Briefings in Bioinformatics, 2022, 23, .                                                             | 3.2 | 39        |
| 475 | Bacterial adaptation by a transposition burst of an invading IS element. Genome Biology and Evolution, 2021, 13, .                                                                                           | 1.1 | 6         |
| 478 | KH domain proteins: Another family of bacterial RNA matchmakers?. Molecular Microbiology, 2022, 117, 10-19.                                                                                                  | 1.2 | 25        |
| 481 | Citrullination Was Introduced into Animals by Horizontal Gene Transfer from Cyanobacteria. Molecular Biology and Evolution, 2022, 39, .                                                                      | 3.5 | 16        |
| 482 | Density-based binning of gene clusters to infer function or evolutionary history using GeneGrouper. Bioinformatics, 2022, 38, 612-620.                                                                       | 1.8 | 4         |
| 483 | Discovering the Ultimate Limits of Protein Secondary Structure Prediction. Biomolecules, 2021, 11, 1627.                                                                                                     | 1.8 | 4         |
| 484 | PhyloCSF++: a fast and user-friendly implementation of PhyloCSF with annotation tools. Bioinformatics, 2022, 38, 1440-1442.                                                                                  | 1.8 | 8         |
| 492 | Massive rhizobial genomic variation associated with partner quality in <i>Lotus</i> – <i>Mesorhizobium</i> symbiosis. FEMS Microbiology Ecology, 2020, 96, .                                                 | 1.3 | 2         |
| 496 | MinIsoClust. , 2020, , .                                                                                                                                                                                     |     | 1         |
| 497 | Excavating the functionally crucial active-site residues of the DXS protein of <i>Bacillus subtilis</i> by exploring its closest homologues. Journal of Genetic Engineering and Biotechnology, 2020, 18, 76. | 1.5 | 3         |
| 498 | Light attention predicts protein location from the language of life. Bioinformatics Advances, 2021, 1, .                                                                                                     | 0.9 | 57        |
| 499 | Characterization of gut microbial structural variations as determinants of human bile acid metabolism. Cell Host and Microbe, 2021, 29, 1802-1814.e5.                                                        | 5.1 | 43        |
| 500 | Gene sharing among plasmids and chromosomes reveals barriers for antibiotic resistance gene transfer. Philosophical Transactions of the Royal Society B: Biological Sciences, 2022, 377, 20200467.           | 1.8 | 23        |
| 501 | Novel genome characteristics contribute to the invasiveness of <i>Phragmites australis</i> (common) Tj ETQq0 0 0 rgBT /Overlock 10 T                                                                         | 2.8 | 10        |
| 502 | The impact of reductants on the catalytic efficiency of a lytic polysaccharide monooxygenase and the special role of dehydroascorbic acid. FEBS Letters, 2022, 596, 53-70.                                   | 1.3 | 29        |
| 504 | A primer and discussion on DNA-based microbiome data and related bioinformatics analyses. , 0, 1, .                                                                                                          |     | 6         |
| 505 | NetSolP: predicting protein solubility in <i>Escherichia coli</i> using language models. Bioinformatics, 2022, 38, 941-946.                                                                                  | 1.8 | 16        |

| #   | ARTICLE                                                                                                                                                                                                                                                  | IF   | CITATIONS |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 506 | Persistence and plasticity in bacterial gene regulation. <i>Nature Methods</i> , 2021, 18, 1499-1505.                                                                                                                                                    | 9.0  | 23        |
| 507 | Functional and Taxonomic Effects of Organic Amendments on the Restoration of Semiarid Quarry Soils. <i>MSystems</i> , 2021, 6, e0075221.                                                                                                                 | 1.7  | 4         |
| 508 | Metagenomic discovery of CRISPR-associated transposons. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2021, 118, .                                                                                            | 3.3  | 38        |
| 509 | Metagenome-assembled genomes and gene catalog from the chicken gut microbiome aid in deciphering antibiotic resistomes. <i>Communications Biology</i> , 2021, 4, 1305.                                                                                   | 2.0  | 49        |
| 512 | To catch a hijacker: abundance, evolution and genetic diversity of P4-like bacteriophage satellites. <i>Philosophical Transactions of the Royal Society B: Biological Sciences</i> , 2022, 377, 20200475.                                                | 1.8  | 21        |
| 514 | Autism-related dietary preferences mediate autism-gut microbiome associations. <i>Cell</i> , 2021, 184, 5916-5931.e17.                                                                                                                                   | 13.5 | 172       |
| 515 | Unexpected cryptic species among streptophyte algae most distant to land plants. <i>Proceedings of the Royal Society B: Biological Sciences</i> , 2021, 288, 20212168.                                                                                   | 1.2  | 22        |
| 518 | Effector-mediated membrane disruption controls cell death in CBASS antiphage defense. <i>Molecular Cell</i> , 2021, 81, 5039-5051.e5.                                                                                                                    | 4.5  | 59        |
| 519 | The coiled-coil domain of <i>Escherichia coli</i> FtsLB is a structurally detuned element critical for modulating its activation in bacterial cell division. <i>Journal of Biological Chemistry</i> , 2022, 298, 101460.                                 | 1.6  | 8         |
| 521 | Animal gut microbiome mediates the effects of antibiotic pollution on an artificial freshwater system. <i>Journal of Hazardous Materials</i> , 2022, 425, 127968.                                                                                        | 6.5  | 23        |
| 523 | A simple guide to <i>de novo</i> transcriptome assembly and annotation. <i>Briefings in Bioinformatics</i> , 2022, 23, .                                                                                                                                 | 3.2  | 42        |
| 524 | A comprehensive study on the exposure of nanoplastics to constructed wetland ecological systems: Macrophyte physiology and microbial enzymology, community composition and metabolic functions. <i>Chemical Engineering Journal</i> , 2022, 434, 134592. | 6.6  | 28        |
| 525 | Harnessing protein folding neural networks for peptide-protein docking. <i>Nature Communications</i> , 2022, 13, 176.                                                                                                                                    | 5.8  | 682       |
| 526 | The Mouse Gastrointestinal Bacteria Catalogue enables translation between the mouse and human gut microbiotas via functional mapping. <i>Cell Host and Microbe</i> , 2022, 30, 124-138.e8.                                                               | 5.1  | 59        |
| 527 | Soils and sediments host Thermoplasmata archaea encoding novel copper membrane monooxygenases (CuMMOs). <i>ISME Journal</i> , 2022, 16, 1348-1362.                                                                                                       | 4.4  | 17        |
| 528 | Simplified quality assessment for small-molecule ligands in the Protein Data Bank. <i>Structure</i> , 2022, 30, 252-262.e4.                                                                                                                              | 1.6  | 12        |
| 529 | Resolving the structure of phage-bacteria interactions in the context of natural diversity. <i>Nature Communications</i> , 2022, 13, 372.                                                                                                                | 5.8  | 48        |
| 530 | Investigation of the Importance of Protein 3D Structure for Assessing Conservation of Lysine Acetylation Sites in Protein Homologs. <i>Frontiers in Microbiology</i> , 2021, 12, 805181.                                                                 | 1.5  | 1         |

| #   | ARTICLE                                                                                                                                                                                                           | IF  | CITATIONS |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 531 | Bacteriophages Roam the Wheat Phyllosphere. <i>Viruses</i> , 2022, 14, 244.                                                                                                                                       | 1.5 | 13        |
| 532 | Enteric virome negatively affects seroconversion following oral rotavirus vaccination in a longitudinally sampled cohort of Ghanaian infants. <i>Cell Host and Microbe</i> , 2022, 30, 110-123.e5.                | 5.1 | 23        |
| 533 | Incorporating genome-based phylogeny and functional similarity into diversity assessments helps to resolve a global collection of human gut metagenomes. <i>Environmental Microbiology</i> , 2022, 24, 3966-3984. | 1.8 | 2         |
| 537 | RdRp-based sensitive taxonomic classification of RNA viruses for metagenomic data. <i>Briefings in Bioinformatics</i> , 2022, , .                                                                                 | 3.2 | 1         |
| 539 | Bacterial gasdermins reveal an ancient mechanism of cell death. <i>Science</i> , 2022, 375, 221-225.                                                                                                              | 6.0 | 132       |
| 542 | From complete cross-docking to partners identification and binding sites predictions. <i>PLoS Computational Biology</i> , 2022, 18, e1009825.                                                                     | 1.5 | 9         |
| 543 | Secondary Structures of Proteins Follow Menzerath-Altmann Law. <i>International Journal of Molecular Sciences</i> , 2022, 23, 1569.                                                                               | 1.8 | 1         |
| 544 | SPOT-Contact-LM: improving single-sequence-based prediction of protein contact map using a transformer language model. <i>Bioinformatics</i> , 2022, 38, 1888-1894.                                               | 1.8 | 25        |
| 545 | The stem rust effector protein AvrSr50 escapes Sr50 recognition by a substitution in a single surface-exposed residue. <i>New Phytologist</i> , 2022, 234, 592-606.                                               | 3.5 | 32        |
| 546 | Rapid and accurate identification of ribosomal RNA sequences via deep learning. <i>Nucleic Acids Research</i> , 2022, 50, e60-e60.                                                                                | 6.5 | 25        |
| 547 | Improved Red Fluorescent Redox Indicators for Monitoring Cytosolic and Mitochondrial Thioredoxin Redox Dynamics. <i>Biochemistry</i> , 2022, 61, 377-384.                                                         | 1.2 | 5         |
| 548 | Cross-species transcriptomic analyses reveals common and opposite responses in Arabidopsis, rice and barley following oxidative stress and hormone treatment. <i>BMC Plant Biology</i> , 2022, 22, 62.            | 1.6 | 11        |
| 549 | Worldwide Protein Data Bank (wwPDB): A virtual treasure for research in biotechnology. <i>European Journal of Microbiology and Immunology</i> , 2022, 11, 77-86.                                                  | 1.5 | 13        |
| 551 | Chromosome-Level Genome Assembly of <i>Anthidium xuezhongi</i> Niu & Zhu, 2020 (Hymenoptera: Apoidea): Tj ETQq1_1_0.784314 rgBT                                                                                   | 1.1 | 1         |
| 552 | A high-quality functional genome assembly of <i>Delia radicum</i> L. (Diptera: Anthomyiidae) annotated from egg to adult. <i>Molecular Ecology Resources</i> , 2022, 22, 1954-1971.                               | 2.2 | 6         |
| 555 | Single-cell genomics for resolution of conserved bacterial genes and mobile genetic elements of the human intestinal microbiota using flow cytometry. <i>Gut Microbes</i> , 2022, 14, 2029673.                    | 4.3 | 5         |
| 557 | Context-aware genomic surveillance reveals hidden transmission of a carbapenemase-producing <i>Klebsiella pneumoniae</i> . <i>Microbial Genomics</i> , 2021, 7, .                                                 | 1.0 | 9         |
| 558 | A catalogue of 1,167 genomes from the human gut archaeome. <i>Nature Microbiology</i> , 2022, 7, 48-61.                                                                                                           | 5.9 | 72        |

| #   | ARTICLE                                                                                                                                                                                                                     | IF   | CITATIONS |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 559 | Towards the biogeography of prokaryotic genes. <i>Nature</i> , 2022, 601, 252-256.                                                                                                                                          | 13.7 | 85        |
| 561 | Comparative Genomics of Novel <i>Agrobacterium</i> G3 Strains Isolated From the International Space Station and Description of <i>Agrobacterium tomkonis</i> sp. nov.. <i>Frontiers in Microbiology</i> , 2021, 12, 765943. | 1.5  | 22        |
| 562 | A Five-Fold Expansion of the Global RNA Virome Reveals Multiple New Clades of RNA Bacteriophages. <i>SSRN Electronic Journal</i> , 0, , .                                                                                   | 0.4  | 4         |
| 563 | Profile and Actual Transmissibility of Carbapenem Resistance Genes: Intracellular and Extracellular DNA in Hospital Wastewater. <i>SSRN Electronic Journal</i> , 0, , .                                                     | 0.4  | 0         |
| 564 | Functional Genes Highlight Contrasting Elevational Pattern of Bacteria- and Fungi-Derived Compound Decompositions in Forest Soils. <i>SSRN Electronic Journal</i> , 0, , .                                                  | 0.4  | 0         |
| 567 | Microbial Community Structure and Functional Potential of Deep-Sea Sediments on Low Activity Hydrothermal Area in the Central Indian Ridge. <i>Frontiers in Marine Science</i> , 2022, 9, .                                 | 1.2  | 6         |
| 568 | Using deep learning to annotate the protein universe. <i>Nature Biotechnology</i> , 2022, 40, 932-937.                                                                                                                      | 9.4  | 133       |
| 570 | Gene age shapes the transcriptional landscape of sexual morphogenesis in mushroom-forming fungi ( <i>Agaricomycetes</i> ). <i>ELife</i> , 2022, 11, .                                                                       | 2.8  | 18        |
| 571 | CRISPR-Cas Systems in Gut Microbiome of Children with Autism Spectrum Disorders. <i>Life</i> , 2022, 12, 367.                                                                                                               | 1.1  | 0         |
| 574 | N-Terminal Modifications of Ubiquitin via Methionine Excision, Deamination, and Arginylation Expand the Ubiquitin Code. <i>Molecules and Cells</i> , 2022, 45, 158-167.                                                     | 1.0  | 7         |
| 576 | Correlational networking guides the discovery of unclustered lanthipeptide protease-encoding genes. <i>Nature Communications</i> , 2022, 13, 1647.                                                                          | 5.8  | 12        |
| 577 | Phages and their satellites encode hotspots of antiviral systems. <i>Cell Host and Microbe</i> , 2022, 30, 740-753.e5.                                                                                                      | 5.1  | 129       |
| 578 | Construct a variable-length fragment library for de novo protein structure prediction. <i>Briefings in Bioinformatics</i> , 2022, 23, .                                                                                     | 3.2  | 4         |
| 580 | A-Prot: protein structure modeling using MSA transformer. <i>BMC Bioinformatics</i> , 2022, 23, 93.                                                                                                                         | 1.2  | 6         |
| 582 | IntegronFinder 2.0: Identification and Analysis of Integrons across Bacteria, with a Focus on Antibiotic Resistance in <i>Klebsiella</i> . <i>Microorganisms</i> , 2022, 10, 700.                                           | 1.6  | 61        |
| 583 | Sampling alternative conformational states of transporters and receptors with AlphaFold2. <i>ELife</i> , 2022, 11, .                                                                                                        | 2.8  | 180       |
| 585 | Quantifying and Cataloguing Unknown Sequences within Human Microbiomes. <i>MSystems</i> , 2022, 7, e0146821.                                                                                                                | 1.7  | 2         |
| 586 | B-Cell Epitope Mapping of TprC and TprD Variants of <i>Treponema pallidum</i> Subspecies Informs Vaccine Development for Human Treponematoses. <i>Frontiers in Immunology</i> , 2022, 13, 862491.                           | 2.2  | 7         |

| #   | ARTICLE                                                                                                                                                                                                            | IF  | CITATIONS |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 587 | SHOOT: phylogenetic gene search and ortholog inference. <i>Genome Biology</i> , 2022, 23, 85.                                                                                                                      | 3.8 | 28        |
| 592 | InParanoid-DIAMOND: faster orthology analysis with the InParanoid algorithm. <i>Bioinformatics</i> , 2022, 38, 2918-2919.                                                                                          | 1.8 | 12        |
| 593 | Worldwide Population Structure of the Coffee Rust Fungus <i>Hemileia vastatrix</i> Is Strongly Shaped by Local Adaptation and Breeding History. <i>Phytopathology</i> , 2022, 112, 1998-2011.                      | 1.1 | 5         |
| 594 | <i>Streptococcus pneumoniae</i> Strains Isolated From a Single Pediatric Patient Display Distinct Phenotypes. <i>Frontiers in Cellular and Infection Microbiology</i> , 2022, 12, 866259.                          | 1.8 | 3         |
| 595 | <i>Leclercia pneumoniae</i> sp. nov., a bacterium isolated from clinical specimen in Leipzig, Germany. <i>International Journal of Systematic and Evolutionary Microbiology</i> , 2022, 72, .                      | 0.8 | 7         |
| 596 | Conserved exchange of paralog proteins during neuronal differentiation. <i>Life Science Alliance</i> , 2022, 5, e202201397.                                                                                        | 1.3 | 0         |
| 597 | Clustering biological sequences with dynamic sequence similarity threshold. <i>BMC Bioinformatics</i> , 2022, 23, 108.                                                                                             | 1.2 | 5         |
| 599 | Three new inhibitors of class A $\beta$ -lactamases evaluated by molecular docking and dynamics simulations methods: relebactam, enmetazobactam, and QPX7728. <i>Journal of Molecular Modeling</i> , 2022, 28, 76. | 0.8 | 1         |
| 600 | A Web Server for GPCR-GPCR Interaction Pair Prediction. <i>Frontiers in Endocrinology</i> , 2022, 13, 825195.                                                                                                      | 1.5 | 3         |
| 601 | Diverse Protein Architectures and $\beta$ -N-Methylation Patterns Define Split Borosin RiPP Biosynthetic Gene Clusters. <i>ACS Chemical Biology</i> , 2022, 17, 908-917.                                           | 1.6 | 11        |
| 603 | Metagenomics reveals global-scale contrasts in nitrogen cycling and cyanobacterial light-harvesting mechanisms in glacier cryoconite. <i>Microbiome</i> , 2022, 10, 50.                                            | 4.9 | 10        |
| 604 | Unifying the known and unknown microbial coding sequence space. <i>ELife</i> , 2022, 11, .                                                                                                                         | 2.8 | 41        |
| 605 | A mouse model of occult intestinal colonization demonstrating antibiotic-induced outgrowth of carbapenem-resistant Enterobacteriaceae. <i>Microbiome</i> , 2022, 10, 43.                                           | 4.9 | 8         |
| 606 | Prediction of RNA-protein interactions using a nucleotide language model. <i>Bioinformatics Advances</i> , 2022, 2, .                                                                                              | 0.9 | 15        |
| 607 | Different Genes are Recruited During Convergent Evolution of Pregnancy and the Placenta. <i>Molecular Biology and Evolution</i> , 2022, 39, .                                                                      | 3.5 | 9         |
| 609 | Analyzing structural features of proteins from deep-sea organisms. <i>Proteins: Structure, Function and Bioinformatics</i> , 2022, 90, 1521-1537.                                                                  | 1.5 | 2         |
| 613 | Human enteric viruses autonomously shape inflammatory bowel disease phenotype through divergent innate immunomodulation. <i>Science Immunology</i> , 2022, 7, eabn6660.                                            | 5.6 | 38        |
| 614 | HOMELETTE: a unified interface to homology modelling software. <i>Bioinformatics</i> , 2022, 38, 1749-1751.                                                                                                        | 1.8 | 4         |

| #   | ARTICLE                                                                                                                                                                                                   | IF  | CITATIONS |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 616 | A unified evolutionary origin for the ubiquitous protein transporters SecY and YidC. <i>BMC Biology</i> , 2021, 19, 266.                                                                                  | 1.7 | 24        |
| 620 | Protein embeddings and deep learning predict binding residues for various ligand classes. <i>Scientific Reports</i> , 2021, 11, 23916.                                                                    | 1.6 | 63        |
| 622 | An Interplay between Viruses and Bacteria Associated with the White Sea Sponges Revealed by Metagenomics. <i>Life</i> , 2022, 12, 25.                                                                     | 1.1 | 3         |
| 623 | Transposable Prophages in <i>Leptospira</i> : An Ancient, Now Diverse, Group Predominant in Causative Agents of Weil's Disease. <i>International Journal of Molecular Sciences</i> , 2021, 22, 13434.     | 1.8 | 4         |
| 624 | An optimized FM-index library for nucleotide and amino acid search. <i>Algorithms for Molecular Biology</i> , 2021, 16, 25.                                                                               | 0.3 | 1         |
| 626 | DARTS: An Algorithm for Domain-Associated Retrotransposon Search in Genome Assemblies. <i>Genes</i> , 2022, 13, 9.                                                                                        | 1.0 | 6         |
| 627 | Embeddings from protein language models predict conservation and variant effects. <i>Human Genetics</i> , 2022, 141, 1629-1647.                                                                           | 1.8 | 60        |
| 629 | Phage-bacterial contig association prediction with a convolutional neural network. <i>Bioinformatics</i> , 2022, 38, i45-i52.                                                                             | 1.8 | 3         |
| 630 | DEPhT: a novel approach for efficient prophage discovery and precise extraction. <i>Nucleic Acids Research</i> , 2022, 50, e75-e75.                                                                       | 6.5 | 13        |
| 633 | A combined bioinformatics and LC-MS-based approach for the development and benchmarking of a comprehensive database of <i>Lymnaea</i> CNS proteins. <i>Journal of Experimental Biology</i> , 2022, 225, . | 0.8 | 3         |
| 635 | 3DLigandSite: structure-based prediction of protein-ligand binding sites. <i>Nucleic Acids Research</i> , 2022, 50, W13-W20.                                                                              | 6.5 | 25        |
| 636 | How many single-copy orthologous genes from whole genomes reveal deep gastropod relationships?. <i>PeerJ</i> , 2022, 10, e13285.                                                                          | 0.9 | 1         |
| 638 | Novel Viral DNA Polymerases From Metagenomes Suggest Genomic Sources of Strand-Displacing Biochemical Phenotypes. <i>Frontiers in Microbiology</i> , 2022, 13, 858366.                                    | 1.5 | 2         |
| 687 | S51 Family Peptidases Provide Resistance to Peptidyl-Nucleotide Antibiotic McC. <i>MBio</i> , 2022, 13, e0080522.                                                                                         | 1.8 | 0         |
| 688 | CHERRY: a Computational method for accurate prediction of virus-prokaryotic interactions using a graph encoder-decoder model. <i>Briefings in Bioinformatics</i> , 2022, 23, .                            | 3.2 | 21        |
| 689 | GNAT toxins evolve toward narrow tRNA target specificities. <i>Nucleic Acids Research</i> , 2022, 50, 5807-5817.                                                                                          | 6.5 | 2         |
| 690 | Analysis of environmental driving factors on Core Functional Community during Daqu fermentation. <i>Food Research International</i> , 2022, 157, 111286.                                                  | 2.9 | 30        |
| 691 | Metagenome-assembled genomes of phytoplankton microbiomes from the Arctic and Atlantic Oceans. <i>Microbiome</i> , 2022, 10, 67.                                                                          | 4.9 | 17        |



| #   | ARTICLE                                                                                                                                                                                              | IF   | CITATIONS |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 692 | A deep siamese neural network improves metagenome-assembled genomes in microbiome datasets across different environments. <i>Nature Communications</i> , 2022, 13, 2326.                             | 5.8  | 47        |
| 694 | Mechanism of integrin activation by talin and its cooperation with kindlin. <i>Nature Communications</i> , 2022, 13, 2362.                                                                           | 5.8  | 30        |
| 695 | Reaching alignment-profile-based accuracy in predicting protein secondary and tertiary structural properties without alignment. <i>Scientific Reports</i> , 2022, 12, 7607.                          | 1.6  | 14        |
| 697 | Transitions of foliar mycobiota community and transcriptome in response to pathogenic conifer needle interactions. <i>Scientific Reports</i> , 2022, 12, 7832.                                       | 1.6  | 3         |
| 700 | Divergence in the ABA gene regulatory network underlies differential growth control. <i>Nature Plants</i> , 2022, 8, 549-560.                                                                        | 4.7  | 19        |
| 701 | vRhyne enables binning of viral genomes from metagenomes. <i>Nucleic Acids Research</i> , 2022, 50, e83-e83.                                                                                         | 6.5  | 30        |
| 702 | Impact of salivary and pancreatic amylase gene copy numbers on diabetes, obesity, and functional profiles of microbiome in Northern Japanese population. <i>Scientific Reports</i> , 2022, 12, 7628. | 1.6  | 3         |
| 703 | Systematic and quantitative view of the antiviral arsenal of prokaryotes. <i>Nature Communications</i> , 2022, 13, 2561.                                                                             | 5.8  | 208       |
| 704 | The potential for bacteria from carbon-limited deep terrestrial environments to participate in chlorine cycling. <i>FEMS Microbiology Ecology</i> , 2022, 98, .                                      | 1.3  | 3         |
| 705 | A catalogue of resistance gene homologs and a chromosome-scale reference sequence support resistance gene mapping in winter wheat. <i>Plant Biotechnology Journal</i> , 2022, 20, 1730-1742.         | 4.1  | 21        |
| 706 | Distinct gene clusters drive formation of ferrosome organelles in bacteria. <i>Nature</i> , 2022, 606, 160-164.                                                                                      | 13.7 | 15        |
| 707 | A network approach to decipher the dynamics of <i>Lysobacteraceae</i> plasmid gene sharing. <i>Molecular Ecology</i> , 2023, 32, 2660-2673.                                                          | 2.0  | 6         |
| 708 | Improved DNA-Versus-Protein Homology Search for Protein Fossils. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2023, 20, 1691-1699.                                     | 1.9  | 3         |
| 712 | Protein language-model embeddings for fast, accurate, and alignment-free protein structure prediction. <i>Structure</i> , 2022, 30, 1169-1177.e4.                                                    | 1.6  | 52        |
| 714 | Disentangling the Protein Order/Disorder Continuum Using a Sequence-Based Foldability Score. <i>SSRN Electronic Journal</i> , 0, , .                                                                 | 0.4  | 0         |
| 715 | Computationally Decoding NudF Residues To Enhance the Yield of the DXP Pathway. <i>ACS Omega</i> , 0, , .                                                                                            | 1.6  | 0         |
| 716 | Substantiation of propitious "Enzybiotic" from two novel bacteriophages isolated from a wastewater treatment plant in Qatar. <i>Scientific Reports</i> , 2022, 12, .                                 | 1.6  | 2         |
| 717 | Bacterial colonisation dynamics of household plastics in a coastal environment. <i>Science of the Total Environment</i> , 2022, 838, 156199.                                                         | 3.9  | 12        |

| #   | ARTICLE                                                                                                                                                                                                                      | IF  | CITATIONS |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 720 | PepNN: a deep attention model for the identification of peptide binding sites. <i>Communications Biology</i> , 2022, 5, .                                                                                                    | 2.0 | 16        |
| 721 | ColabFold: making protein folding accessible to all. <i>Nature Methods</i> , 2022, 19, 679-682.                                                                                                                              | 9.0 | 3,242     |
| 727 | ContactPFP: Protein Function Prediction Using Predicted Contact Information. <i>Frontiers in Bioinformatics</i> , 2022, 2, .                                                                                                 | 1.0 | 6         |
| 728 | Flexible protein database based on amino acid k-mers. <i>Scientific Reports</i> , 2022, 12, .                                                                                                                                | 1.6 | 4         |
| 729 | Prior knowledge facilitates low homologous protein secondary structure prediction with DSM distillation. <i>Bioinformatics</i> , 2022, 38, 3574-3581.                                                                        | 1.8 | 6         |
| 731 | A scaling law in CRISPR repertoire sizes arises from the avoidance of autoimmunity. <i>Current Biology</i> , 2022, , .                                                                                                       | 1.8 | 5         |
| 732 | Co-component signal transduction systems: Fast-evolving virulence regulation cassettes discovered in enteric bacteria. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2022, 119, . | 3.3 | 14        |
| 733 | Ecogenomics sheds light on diverse lifestyle strategies in freshwater CPR. <i>Microbiome</i> , 2022, 10, .                                                                                                                   | 4.9 | 22        |
| 734 | Transcription factor lineages in plant-pathogenic fungi, connecting diversity with fungal virulence. <i>Fungal Genetics and Biology</i> , 2022, 161, 103712.                                                                 | 0.9 | 4         |
| 735 | The microbiome of cryospheric ecosystems. <i>Nature Communications</i> , 2022, 13, .                                                                                                                                         | 5.8 | 20        |
| 736 | Noncanonical prokaryotic X family DNA polymerases lack polymerase activity and act as exonucleases. <i>Nucleic Acids Research</i> , 0, , .                                                                                   | 6.5 | 0         |
| 737 | Multi-level analysis of intrinsically disordered protein docking methods. <i>Methods</i> , 2022, 204, 55-63.                                                                                                                 | 1.9 | 3         |
| 739 | Protein Sequence Classification Using Convolutional Neural Network and Natural Language Processing. <i>Studies in Big Data</i> , 2022, , 133-144.                                                                            | 0.8 | 2         |
| 740 | Contrastive learning on protein embeddings enlightens midnight zone. <i>NAR Genomics and Bioinformatics</i> , 2022, 4, .                                                                                                     | 1.5 | 38        |
| 743 | Deep embeddings to comprehend and visualize microbiome protein space. <i>Scientific Reports</i> , 2022, 12, .                                                                                                                | 1.6 | 3         |
| 746 | Nanopore sequencing of a monkeypox virus strain isolated from a pustular lesion in the Central African Republic. <i>Scientific Reports</i> , 2022, 12, .                                                                     | 1.6 | 42        |
| 747 | Novel Unspecific Peroxygenase from <i>Truncatella angustata</i> Catalyzes the Synthesis of Bioactive Lipid Mediators. <i>Microorganisms</i> , 2022, 10, 1267.                                                                | 1.6 | 2         |
| 748 | Robust variation in infant gut microbiome assembly across a spectrum of lifestyles. <i>Science</i> , 2022, 376, 1220-1223.                                                                                                   | 6.0 | 63        |

| #   | ARTICLE                                                                                                                                                                                                         | IF   | CITATIONS |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 750 | MAGNETO: An Automated Workflow for Genome-Resolved Metagenomics. <i>MSystems</i> , 2022, 7, .                                                                                                                   | 1.7  | 7         |
| 751 | Disentangling the genetic basis of rhizosphere microbiome assembly in tomato. <i>Nature Communications</i> , 2022, 13, .                                                                                        | 5.8  | 53        |
| 752 | Multi-color RNA imaging with CRISPR-Cas13b systems in living cells. , 2022, 1, 100044.                                                                                                                          |      | 13        |
| 753 | Odd one out? Functional tuning of <i>Zymomonas mobilis</i> pyruvate kinase is narrower than its allosteric, human counterpart. <i>Protein Science</i> , 2022, 31, .                                             | 3.1  | 5         |
| 754 | The DarTG toxin-antitoxin system provides phage defence by ADP-ribosylating viral DNA. <i>Nature Microbiology</i> , 2022, 7, 1028-1040.                                                                         | 5.9  | 78        |
| 755 | Growth rate determines prokaryote-provirus network modulated by temperature and host genetic traits. <i>Microbiome</i> , 2022, 10, .                                                                            | 4.9  | 0         |
| 757 | Biosynthetic potential of the global ocean microbiome. <i>Nature</i> , 2022, 607, 111-118.                                                                                                                      | 13.7 | 128       |
| 760 | Research progress of reduced amino acid alphabets in protein analysis and prediction. <i>Computational and Structural Biotechnology Journal</i> , 2022, 20, 3503-3510.                                          | 1.9  | 7         |
| 761 | A chromosome-level genome assembly and intestinal transcriptome of <i>Trypoxylus dichotomus</i> (Coleoptera: Scarabaeidae) to understand its lignocellulose digestion ability. <i>GigaScience</i> , 2022, 11, . | 3.3  | 5         |
| 762 | A genome and gene catalog of glacier microbiomes. <i>Nature Biotechnology</i> , 2022, 40, 1341-1348.                                                                                                            | 9.4  | 50        |
| 763 | <i>Variovorax terrae</i> sp. nov. Isolated from Soil with Potential Antioxidant Activity. <i>Journal of Microbiology and Biotechnology</i> , 2022, , .                                                          | 0.9  | 1         |
| 764 | Marine Microeukaryote Metatranscriptomics: Sample Processing and Bioinformatic Workflow Recommendations for Ecological Applications. <i>Frontiers in Marine Science</i> , 0, 9, .                               | 1.2  | 8         |
| 770 | Structure of cytoplasmic ring of nuclear pore complex by integrative cryo-EM and AlphaFold. <i>Science</i> , 2022, 376, .                                                                                       | 6.0  | 89        |
| 774 | Scalable multiplex co-fractionation/mass spectrometry platform for accelerated protein interactome discovery. <i>Nature Communications</i> , 2022, 13, .                                                        | 5.8  | 20        |
| 778 | Transcriptome and annotation-guided genome assembly of the European starling. <i>Molecular Ecology Resources</i> , 2022, 22, 3141-3160.                                                                         | 2.2  | 9         |
| 779 | A Chromosome-Scale Genome Assembly of the Okapi ( <i>Okapia johnstoni</i> ). <i>Journal of Heredity</i> , 2022, 113, 568-576.                                                                                   | 1.0  | 2         |
| 780 | Expanded Dataset Reveals the Emergence and Evolution of DNA Gyrase in Archaea. <i>Molecular Biology and Evolution</i> , 2022, 39, .                                                                             | 3.5  | 4         |
| 782 | Core packing of well-defined X-ray and NMR structures is the same. <i>Protein Science</i> , 2022, 31, .                                                                                                         | 3.1  | 1         |

| #   | ARTICLE                                                                                                                                                                                                            | IF  | CITATIONS |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 783 | Theoretical studies on RNA recognition by Musashi 1 RNA-binding protein. <i>Scientific Reports</i> , 2022, 12, .                                                                                                   | 1.6 | 5         |
| 784 | Comprehensive Evaluation of RNA and DNA Viromic Methods Based on Species Richness and Abundance Analyses Using Marmot Rectal Samples. <i>MSystems</i> , 2022, 7, .                                                 | 1.7 | 8         |
| 785 | A chromosome-level genome assembly of <i>Neotoxoptera formosana</i> (Takahashi, 1921) (Hemiptera: Tj ETQq0.0.0 rgBT <sub>0</sub> Overlock                                                                          | 0.8 | 0         |
| 787 | Structurally derived universal mechanism for the catalytic cycle of the tail-anchored targeting factor Get3. <i>Nature Structural and Molecular Biology</i> , 2022, 29, 820-830.                                   | 3.6 | 7         |
| 788 | Molecular Evolution of the Activating Transcription Factors Shapes the Adaptive Cellular Responses to Oxidative Stress. <i>Oxidative Medicine and Cellular Longevity</i> , 2022, 2022, 1-13.                       | 1.9 | 4         |
| 789 | Recent Acquisition of Functional m6A RNA Demethylase Domain in Orchid Ty3/Gypsy Elements. <i>Frontiers in Plant Science</i> , 0, 13, .                                                                             | 1.7 | 0         |
| 790 | Structural Insights into the Interaction between <i>Bacillus subtilis</i> SepF Assembly and FtsZ by Solid-State NMR Spectroscopy. <i>Journal of Physical Chemistry B</i> , 2022, 126, 5219-5230.                   | 1.2 | 2         |
| 791 | Helical remodeling augments 5-lipoxygenase activity in the synthesis of proinflammatory mediators. <i>Journal of Biological Chemistry</i> , 2022, 298, 102282.                                                     | 1.6 | 6         |
| 792 | Benchmarking <i>AlphaFold</i> for protein complex modeling reveals accuracy determinants. <i>Protein Science</i> , 2022, 31, .                                                                                     | 3.1 | 133       |
| 793 | Structural and functional characterization of DdrC, a novel DNA damage-induced nucleoid associated protein involved in DNA compaction. <i>Nucleic Acids Research</i> , 2022, 50, 7680-7696.                        | 6.5 | 8         |
| 794 | Conserved and lineage-specific hypothetical proteins may have played a central role in the rise and diversification of major archaeal groups. <i>BMC Biology</i> , 2022, 20, .                                     | 1.7 | 4         |
| 795 | Stimulation of Distinct Rhizosphere Bacteria Drives Phosphorus and Nitrogen Mineralization in Oilseed Rape under Field Conditions. <i>MSystems</i> , 2022, 7, .                                                    | 1.7 | 7         |
| 797 | Bacteria deplete deoxynucleotides to defend against bacteriophage infection. <i>Nature Microbiology</i> , 2022, 7, 1200-1209.                                                                                      | 5.9 | 58        |
| 799 | A preorganization oriented computational method for de novo design of Kemp elimination enzymes. <i>Enzyme and Microbial Technology</i> , 2022, 160, 110093.                                                        | 1.6 | 3         |
| 801 | Integrated viral elements suggest the dual lifestyle of <i>Tetraselmis</i> spp. Polinton-like viruses. <i>Virus Evolution</i> , 2022, 8, .                                                                         | 2.2 | 6         |
| 802 | Soil pH, Calcium Content and Bacteria as Major Factors Responsible for the Distribution of the Known Fraction of the DNA Bacteriophage Populations in Soils of Luxembourg. <i>Microorganisms</i> , 2022, 10, 1458. | 1.6 | 4         |
| 805 | FastANI, Mash and Dashing equally differentiate between <i>Klebsiella</i> species. <i>PeerJ</i> , 0, 10, e13784.                                                                                                   | 0.9 | 8         |
| 808 | Scaffolding protein functional sites using deep learning. <i>Science</i> , 2022, 377, 387-394.                                                                                                                     | 6.0 | 191       |

| #   | ARTICLE                                                                                                                                                                                                                     | IF  | CITATIONS |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 809 | Cargo Recognition Mechanisms of Yeast Myo2 Revealed by AlphaFold2-Powered Protein Complex Prediction. <i>Biomolecules</i> , 2022, 12, 1032.                                                                                 | 1.8 | 2         |
| 811 | A Vast World of Viroid-Like Circular RNAs Revealed by Mining Metatranscriptomes. <i>SSRN Electronic Journal</i> , 0, , .                                                                                                    | 0.4 | 0         |
| 812 | mOTUp: a robust Bayesian approach to leverage metagenome-assembled genomes for core-genome estimation. <i>NAR Genomics and Bioinformatics</i> , 2022, 4, .                                                                  | 1.5 | 6         |
| 813 | New <i>Microviridae</i> isolated from <i>Sulfitobacter</i> reveals two cosmopolitan subfamilies of single-stranded DNA phages infecting marine and terrestrial Alphaproteobacteria. <i>Virus Evolution</i> , 2022, 8, .     | 2.2 | 5         |
| 814 | De novo determination of mosquitocidal Cry11Aa and Cry11Ba structures from naturally-occurring nanocrystals. <i>Nature Communications</i> , 2022, 13, .                                                                     | 5.8 | 9         |
| 819 | Omics profiles of fecal and oral microbiota change in irritable bowel syndrome patients with diarrhea and symptom exacerbation. <i>Journal of Gastroenterology</i> , 2022, 57, 748-760.                                     | 2.3 | 18        |
| 820 | A balanced gut microbiota is essential to maintain health in captive sika deer. <i>Applied Microbiology and Biotechnology</i> , 2022, 106, 5659-5674.                                                                       | 1.7 | 4         |
| 821 | Insights from analyses of low complexity regions with canonical methods for protein sequence comparison. <i>Briefings in Bioinformatics</i> , 2022, 23, .                                                                   | 3.2 | 7         |
| 824 | Adaptation to novel spatially-structured environments is driven by the capsule and alters virulence-associated traits. <i>Nature Communications</i> , 2022, 13, .                                                           | 5.8 | 17        |
| 827 | Taxonomic classification of DNA sequences beyond sequence similarity using deep neural networks. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2022, 119, .                      | 3.3 | 22        |
| 829 | TALE-cmap: Protein function prediction based on a TALE-based architecture and the structure information from contact map. <i>Computers in Biology and Medicine</i> , 2022, 149, 105938.                                     | 3.9 | 2         |
| 831 | Intraintestinal Analysis of the Functional Activity of Microbiomes and Its Application to the Common Marmoset Intestine. <i>MSystems</i> , 0, , .                                                                           | 1.7 | 0         |
| 835 | High-quality genomes reveal significant genetic divergence and cryptic speciation in the model organism <i>Folsomia candida</i> (collembola). <i>Molecular Ecology Resources</i> , 2023, 23, 273-293.                       | 2.2 | 6         |
| 836 | Prokaryotic innate immunity through pattern recognition of conserved viral proteins. <i>Science</i> , 2022, 377, .                                                                                                          | 6.0 | 90        |
| 838 | iPiDA-LTR: Identifying piwi-interacting RNA-disease associations based on Learning to Rank. <i>PLoS Computational Biology</i> , 2022, 18, e1010404.                                                                         | 1.5 | 10        |
| 839 | RelE3 toxin-antitoxin system of <i>Salmonella Typhimurium</i> with a ribosome-independent toxin and a mutated non-neutralising antitoxin. <i>International Journal of Biological Macromolecules</i> , 2022, 219, 1080-1086. | 3.6 | 0         |
| 841 | Microalgae and Bacteria Interaction—Evidence for Division of Diligence in the Alga Microbiota. <i>Microbiology Spectrum</i> , 2022, 10, .                                                                                   | 1.2 | 17        |
| 842 | Integrated AlphaFold2 and DEER investigation of the conformational dynamics of a pH-dependent APC antiporter. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2022, 119, .         | 3.3 | 21        |

| #   | ARTICLE                                                                                                                                                                                                                                                | IF   | CITATIONS |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 843 | Varidnaviruses in the Human Gut: A Major Expansion of the Order Vinavirales. <i>Viruses</i> , 2022, 14, 1842.                                                                                                                                          | 1.5  | 3         |
| 846 | Mixing plant-based proteins: Gel properties of hemp, pea, lentil proteins and their binary mixtures. <i>Food Research International</i> , 2022, 161, 111752.                                                                                           | 2.9  | 15        |
| 848 | NLRexpressâ€”A bundle of machine learning motif predictorsâ€”Reveals motif stability underlying plant Nod-like receptors diversity. <i>Frontiers in Plant Science</i> , 0, 13, .                                                                       | 1.7  | 5         |
| 850 | Metagenomic Insights into the Effects of Benzyl Dodecyl Dimethyl Ammonium Bromide (Bdab) Shock on Bacterial-Driven Nitrogen Removal in Moving Bed Biofilm Reactor (Mbbr). <i>SSRN Electronic Journal</i> , 0, , .                                      | 0.4  | 0         |
| 851 | Metagenomic Approaches for the Discovery of Pollutant-Remediating Enzymes: Recent Trends and Challenges. , 2022, , 571-604.                                                                                                                            |      | 0         |
| 852 | d-StructMAN: Containerized structural annotation on the scale from genetic variants to whole proteomes. <i>GigaScience</i> , 2022, 11, .                                                                                                               | 3.3  | 0         |
| 853 | Analysis of lineage-specific protein family variability in prokaryotes combined with evolutionary reconstructions. <i>Biology Direct</i> , 2022, 17, .                                                                                                 | 1.9  | 5         |
| 855 | Metabolomic and metagenomic analyses of the Chinese mitten crab <i>Eriocheir sinensis</i> after challenge with <i>Metschnikowia bicuspidata</i> . <i>Frontiers in Microbiology</i> , 0, 13, .                                                          | 1.5  | 5         |
| 856 | Homologous recombinationâ€”deficient mutation cluster in tumor suppressor <i>RAD51C</i> identified by comprehensive analysis of cancer variants. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2022, 119, . | 3.3  | 12        |
| 857 | Chromosomeâ€”level assembly of <i>Culex pipiens molestus</i> and improved reference genome of <i>Culex pipiens pallens</i> (Culicidae, Diptera). <i>Molecular Ecology Resources</i> , 2023, 23, 486-498.                                               | 2.2  | 2         |
| 858 | Phage-Plasmids Spread Antibiotic Resistance Genes through Infection and Lysogenic Conversion. <i>MBio</i> , 2022, 13, .                                                                                                                                | 1.8  | 40        |
| 859 | Expansion of the global RNA virome reveals diverse clades of bacteriophages. <i>Cell</i> , 2022, 185, 4023-4037.e18.                                                                                                                                   | 13.5 | 96        |
| 860 | A functional selection reveals previously undetected anti-phage defence systems in the <i>E. coli</i> pangenome. <i>Nature Microbiology</i> , 2022, 7, 1568-1579.                                                                                      | 5.9  | 108       |
| 861 | Genome Report: Genomic novelty within a â€”great speciatorâ€” revealed by a high-quality reference genome of the collared kingfisher ( <i>Todiramphus chloris collaris</i> ). <i>G3: Genes, Genomes, Genetics</i> , 0, , .                             | 0.8  | 2         |
| 862 | Crystallographic and X-ray scattering study of RdfS, a recombination directionality factor from an integrative and conjugative element. <i>Acta Crystallographica Section D: Structural Biology</i> , 2022, 78, 1210-1220.                             | 1.1  | 1         |
| 864 | Variability of strain engraftment and predictability of microbiome composition after fecal microbiota transplantation across different diseases. <i>Nature Medicine</i> , 2022, 28, 1913-1923.                                                         | 15.2 | 81        |
| 865 | High-quality metagenome-assembled genomes from proximal colonic microbiomes of synbiotic-treated korean native black pigs reveal changes in functional capacity. <i>Scientific Reports</i> , 2022, 12, .                                               | 1.6  | 2         |
| 866 | Microseek: A Protein-Based Metagenomic Pipeline for Virus Diagnostic and Discovery. <i>Viruses</i> , 2022, 14, 1990.                                                                                                                                   | 1.5  | 10        |

| #   | ARTICLE                                                                                                                                                                                                                                                                                                                                                                                                                   | IF   | CITATIONS |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 867 | PhaMMseqs: a new pipeline for constructing phage gene families using MMseqs2. <i>G3: Genes, Genomes, Genetics</i> , 2022, 12, .                                                                                                                                                                                                                                                                                           | 0.8  | 14        |
| 868 | Viruses inhibit TIR gcADPR signalling to overcome bacterial defence. <i>Nature</i> , 2022, 611, 326-331.                                                                                                                                                                                                                                                                                                                  | 13.7 | 66        |
| 869 | Long-term excessive application of K2SO4 fertilizer alters bacterial community and functional pathway of tobacco-planting soil. <i>Frontiers in Plant Science</i> , 0, 13, .                                                                                                                                                                                                                                              | 1.7  | 7         |
| 870 | A geometric deep learning framework for drug repositioning over heterogeneous information networks. <i>Briefings in Bioinformatics</i> , 2022, 23, .                                                                                                                                                                                                                                                                      | 3.2  | 20        |
| 871 | The Evolutionary Kaleidoscope of Rhodopsins. <i>MSystems</i> , 2022, 7, .                                                                                                                                                                                                                                                                                                                                                 | 1.7  | 1         |
| 872 | InÂvitro characterization of the full-length human dynein-1 cargo adaptor BicD2. <i>Structure</i> , 2022, 30, 1470-1478.e3.                                                                                                                                                                                                                                                                                               | 1.6  | 3         |
| 873 | Novel species identification and deep functional annotation of electrogenic biofilms, selectively enriched in a microbial fuel cell array. <i>Frontiers in Microbiology</i> , 0, 13, .                                                                                                                                                                                                                                    | 1.5  | 3         |
| 875 | Identification of antiparasitic drug targets using a multi-omics workflow in the acanthocephalan model. <i>BMC Genomics</i> , 2022, 23, .                                                                                                                                                                                                                                                                                 | 1.2  | 1         |
| 876 | Bacteriophages pass through candle-shaped porous ceramic filters: Application for the collection of viruses in soil water. <i>MicrobiologyOpen</i> , 2022, 11, .                                                                                                                                                                                                                                                          | 1.2  | 1         |
| 877 | <i>Apilactobacillus apisilvae</i> sp. nov., <i>Nicolia spurrieriana</i> gen. nov. sp. nov., <i>Bombilactobacillus folatiphilus</i> sp. nov. and <i>Bombilactobacillus thymidiniphilus</i> sp. nov., four new lactic acid bacterial isolates from stingless bees <i>Tetragonula carbonaria</i> and <i>Austroplebeia australis</i> . <i>International Journal of Systematic and Evolutionary Microbiology</i> , 2022, 72, . | 0.8  | 11        |
| 879 | LSQP-DB: a species-specific quantitative PCR primer database for 307 Lactobacillaceae species. <i>Systems Microbiology and Biomanufacturing</i> , 0, , .                                                                                                                                                                                                                                                                  | 1.5  | 2         |
| 880 | mobileOG-db: a Manually Curated Database of Protein Families Mediating the Life Cycle of Bacterial Mobile Genetic Elements. <i>Applied and Environmental Microbiology</i> , 2022, 88, .                                                                                                                                                                                                                                   | 1.4  | 64        |
| 882 | Deciphering microbial gene function using natural language processing. <i>Nature Communications</i> , 2022, 13, .                                                                                                                                                                                                                                                                                                         | 5.8  | 13        |
| 885 | McaA and McaB control the dynamic positioning of a bacterial magnetic organelle. <i>Nature Communications</i> , 2022, 13, .                                                                                                                                                                                                                                                                                               | 5.8  | 8         |
| 886 | A G358S mutation in the <i>Plasmodium falciparum</i> Na <sup>+</sup> pump PfATP4 confers clinically-relevant resistance to ciprofloxacin. <i>Nature Communications</i> , 2022, 13, .                                                                                                                                                                                                                                      | 5.8  | 15        |
| 889 | Robust deep learning-based protein sequence design using ProteinMPNN. <i>Science</i> , 2022, 378, 49-56.                                                                                                                                                                                                                                                                                                                  | 6.0  | 335       |
| 890 | Cryptic and ubiquitous aplastidic cryptophytes are key freshwater flagellated bacterivores. <i>ISME Journal</i> , 2023, 17, 84-94.                                                                                                                                                                                                                                                                                        | 4.4  | 3         |
| 891 | AlphaDrug: protein target specific de novo molecular generation. , 2022, 1, .                                                                                                                                                                                                                                                                                                                                             |      | 12        |

| #   | ARTICLE                                                                                                                                                                                        | IF  | CITATIONS |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 892 | The SMC-family Wadjet complex protects bacteria from plasmid transformation by recognition and cleavage of closed-circular DNA. <i>Molecular Cell</i> , 2022, 82, 4145-4159.e7.                | 4.5 | 44        |
| 893 | Development and Validation of a Bioinformatic Workflow for the Rapid Detection of Viruses in Biosecurity. <i>Viruses</i> , 2022, 14, 2163.                                                     | 1.5 | 4         |
| 894 | KAT2-mediated acetylation switches the mode of PALB2 chromatin association to safeguard genome integrity. <i>ELife</i> , 0, 11, .                                                              | 2.8 | 5         |
| 895 | MetaGT: A pipeline for de novo assembly of metatranscriptomes with the aid of metagenomic data. <i>Frontiers in Microbiology</i> , 0, 13, .                                                    | 1.5 | 1         |
| 897 | E-SNPs&GO: embedding of protein sequence and function improves the annotation of human pathogenic variants. <i>Bioinformatics</i> , 2022, 38, 5168-5174.                                       | 1.8 | 9         |
| 898 | UFCC: database of universal fungal core genes and pipeline for genome-wide phylogenetic analysis of fungi. <i>Nucleic Acids Research</i> , 2023, 51, D777-D784.                                | 6.5 | 7         |
| 900 | DPCfam: Unsupervised protein family classification by Density Peak Clustering of large sequence datasets. <i>PLoS Computational Biology</i> , 2022, 18, e1010610.                              | 1.5 | 2         |
| 902 | SETH predicts nuances of residue disorder from protein embeddings. <i>Frontiers in Bioinformatics</i> , 0, 2, .                                                                                | 1.0 | 14        |
| 903 | LiftoffTools: a toolkit for comparing gene annotations mapped between genome assemblies. <i>F1000Research</i> , 0, 11, 1230.                                                                   | 0.8 | 4         |
| 905 | Antibacterial T6SS effectors with a VRR-Nuc domain are structure-specific nucleases. <i>ELife</i> , 0, 11, .                                                                                   | 2.8 | 10        |
| 908 | WGD: A user-friendly toolkit for evolutionary analyses of whole-genome duplications and ancestral karyotypes. <i>Molecular Plant</i> , 2022, 15, 1841-1851.                                    | 3.9 | 112       |
| 909 | VEBA: a modular end-to-end suite for in silico recovery, clustering, and analysis of prokaryotic, microeukaryotic, and viral genomes from metagenomes. <i>BMC Bioinformatics</i> , 2022, 23, . | 1.2 | 7         |
| 910 | Multi-trial analysis of HIV-1 envelope gp41-reactive antibodies among global recipients of candidate HIV-1 vaccines. <i>Frontiers in Immunology</i> , 0, 13, .                                 | 2.2 | 4         |
| 911 | Systematic discovery of recombinases for efficient integration of large DNA sequences into the human genome. <i>Nature Biotechnology</i> , 2023, 41, 488-499.                                  | 9.4 | 53        |
| 914 | Structural Model of a <i>Porphyromonas gingivalis</i> type IX Secretion System Shuttle Complex. <i>Journal of Molecular Biology</i> , 2022, 434, 167871.                                       | 2.0 | 3         |
| 916 | Between Protein Fold and Nucleophile Identity: Multiscale Modeling of the TEV Protease Enzymeâ€™Substrate Complex. <i>ACS Omega</i> , 2022, 7, 40279-40292.                                    | 1.6 | 4         |
| 917 | Structural basis for HflXr-mediated antibiotic resistance in <i>Listeria monocytogenes</i> . <i>Nucleic Acids Research</i> , 2022, 50, 11285-11300.                                            | 6.5 | 9         |
| 918 | Elimination of Foreign Sequences in Eukaryotic Viral Reference Genomes Improves the Accuracy of Virome Analysis. <i>MSystems</i> , 2022, 7, .                                                  | 1.7 | 7         |



| #   | ARTICLE                                                                                                                                                                                                                                       | IF  | CITATIONS |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 919 | A sequence-based foldability score combined with <sc>AlphaFold2</sc> predictions to disentangle the protein order/disorder continuum. <i>Proteins: Structure, Function and Bioinformatics</i> , 2023, 91, 466-484.                            | 1.5 | 9         |
| 920 | Predicting the structure of large protein complexes using AlphaFold and Monte Carlo tree search. <i>Nature Communications</i> , 2022, 13, .                                                                                                   | 5.8 | 75        |
| 922 | Multi-model predictive analysis of RNA solvent accessibility based on modified residual attention mechanism. <i>Briefings in Bioinformatics</i> , 2022, 23, .                                                                                 | 3.2 | 1         |
| 923 | An expanded arsenal of immune systems that protect bacteria from phages. <i>Cell Host and Microbe</i> , 2022, 30, 1556-1569.e5.                                                                                                               | 5.1 | 137       |
| 925 | The <i>Venturia inaequalis</i> effector repertoire is dominated by expanded families with predicted structural similarity, but unrelated sequence, to avirulence proteins from other plant-pathogenic fungi. <i>BMC Biology</i> , 2022, 20, . | 1.7 | 13        |
| 927 | A chromosome-level phased genome enabling allele-level studies in sweet orange: a case study on citrus Huanglongbing tolerance. <i>Horticulture Research</i> , 2023, 10, .                                                                    | 2.9 | 6         |
| 928 | Biochemical, genomic and structural characteristics of the Acr3 pump in <i>Exiguobacterium</i> strains isolated from arsenic-rich Salar de Huasco sediments. <i>Frontiers in Microbiology</i> , 0, 13, .                                      | 1.5 | 4         |
| 930 | Landscape of infection enhancing antibodies in COVID-19 and healthy donors. <i>Computational and Structural Biotechnology Journal</i> , 2022, 20, 6033-6040.                                                                                  | 1.9 | 5         |
| 932 | Comparative Genomics of the Genus <i>Pseudomonas</i> Reveals Host- and Environment-Specific Evolution. <i>Microbiology Spectrum</i> , 2022, 10, .                                                                                             | 1.2 | 11        |
| 933 | Defining Composition and Function of the Rhizosphere Microbiota of Barley Genotypes Exposed to Growth-Limiting Nitrogen Supplies. <i>MSystems</i> , 2022, 7, .                                                                                | 1.7 | 12        |
| 934 | Prediction of inter-chain distance maps of protein complexes with 2D attention-based deep neural networks. <i>Nature Communications</i> , 2022, 13, .                                                                                         | 5.8 | 22        |
| 935 | <sc>NGSEP</sc> 4: Efficient and accurate identification of orthogroups and whole-genome alignment. <i>Molecular Ecology Resources</i> , 2023, 23, 712-724.                                                                                    | 2.2 | 7         |
| 939 | CoreGenes5.0: An Updated User-Friendly Webserver for the Determination of Core Genes from Sets of Viral and Bacterial Genomes. <i>Viruses</i> , 2022, 14, 2534.                                                                               | 1.5 | 6         |
| 942 | MTSv: rapid alignment-based taxonomic classification and high-confidence metagenomic analysis. <i>PeerJ</i> , 0, 10, e14292.                                                                                                                  | 0.9 | 1         |
| 943 | Improving protein structure prediction using templates and sequence embedding. <i>Bioinformatics</i> , 2023, 39, .                                                                                                                            | 1.8 | 4         |
| 946 | Nearest neighbor search on embeddings rapidly identifies distant protein relations. <i>Frontiers in Bioinformatics</i> , 0, 2, .                                                                                                              | 1.0 | 13        |
| 952 | Advanced prokaryotic systematics: the modern face of an ancient science. <i>New Microbes and New Infections</i> , 2022, 49-50, 101036.                                                                                                        | 0.8 | 4         |
| 953 | Deep learning for protein secondary structure prediction: Pre and post-AlphaFold. <i>Computational and Structural Biotechnology Journal</i> , 2022, 20, 6271-6286.                                                                            | 1.9 | 13        |

| #   | ARTICLE                                                                                                                                                                                                                                                                                                                     | IF  | CITATIONS |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 955 | <sc>BepiPred</sc> 3.0: Improved B-cell epitope prediction using protein language models. Protein Science, 2022, 31, .                                                                                                                                                                                                       | 3.1 | 39        |
| 956 | NLRscape: an atlas of plant NLR proteins. Nucleic Acids Research, 2023, 51, D1470-D1482.                                                                                                                                                                                                                                    | 6.5 | 9         |
| 957 | Protein structure prediction in the deep learning era. Current Opinion in Structural Biology, 2022, 77, 102495.                                                                                                                                                                                                             | 2.6 | 12        |
| 958 | Comparative computational study to augment UbiA prenyltransferases inherent in purple photosynthetic bacteria cultured from mangrove microbial mats in Qatar for coenzyme Q10 biosynthesis.. Biotechnology Reports (Amsterdam, Netherlands), 2022, 36, e00775.                                                              | 2.1 | 2         |
| 959 | The differences in carbohydrate utilization ability between six rounds of Sauce-flavor Daqu. Food Research International, 2023, 163, 112184.                                                                                                                                                                                | 2.9 | 6         |
| 961 | IMG/VR v4: an expanded database of uncultivated virus genomes within a framework of extensive functional, taxonomic, and ecological metadata. Nucleic Acids Research, 2023, 51, D733-D743.                                                                                                                                  | 6.5 | 80        |
| 962 | Soft disorder modulates the assembly path of protein complexes. PLoS Computational Biology, 2022, 18, e1010713.                                                                                                                                                                                                             | 1.5 | 5         |
| 964 | De novo modelling of HEV replication polyprotein: Five-domain breakdown and involvement of flexibility in functional regulation. Virology, 2023, 578, 128-140.                                                                                                                                                              | 1.1 | 11        |
| 965 | Monsoon affects the distribution of antibiotic resistome in Tibetan glaciers. Environmental Pollution, 2023, 317, 120809.                                                                                                                                                                                                   | 3.7 | 11        |
| 966 | Protein model refinement for cryo-EM maps using <i>AlphaFold</i>2 and the DAQ score. Acta Crystallographica Section D: Structural Biology, 2023, 79, 10-21.                                                                                                                                                                 | 1.1 | 5         |
| 967 | Profile and actual transmissibility of Carbapenem resistance genes: Intracellular and extracellular DNA in hospital wastewater. Journal of Environmental Management, 2023, 329, 117085.                                                                                                                                     | 3.8 | 3         |
| 968 | Sequence Analysis and Structural Predictions of Lipid Transfer Bridges in the Repeating Beta Groove (RBC) Superfamily Reveal Past and Present Domain Variations Affecting Form, Function and Interactions of VPS13, ATG2, SHIP164, Hobbit and Tweek. Contact (Thousand Oaks (Ventura County, Calif) Tj ETQq1 1 0.784314 rgB | 0.4 | 18        |
| 969 | Generating new protein sequences by using dense network and attention mechanism. Mathematical Biosciences and Engineering, 2022, 20, 4178-4197.                                                                                                                                                                             | 1.0 | 0         |
| 975 | Using deep-learning predictions of inter-residue distances for model validation. Acta Crystallographica Section D: Structural Biology, 2022, 78, 1412-1427.                                                                                                                                                                 | 1.1 | 6         |
| 976 | The <i>B. subtilis</i> Rok protein is an atypical H-NS-like protein irresponsive to physico-chemical cues. Nucleic Acids Research, 2022, 50, 12166-12185.                                                                                                                                                                   | 6.5 | 6         |
| 978 | Origins of transfer establish networks of functional dependencies for plasmid transfer by conjugation. Nucleic Acids Research, 2023, 51, 3001-3016.                                                                                                                                                                         | 6.5 | 20        |
| 979 | Functions and mechanisms of RNA helicases in plants. Journal of Experimental Botany, 2023, 74, 2295-2310.                                                                                                                                                                                                                   | 2.4 | 6         |
| 980 | AlphaPulldownâ€”a python package for proteinâ€”protein interaction screens using AlphaFold-Multimer. Bioinformatics, 2023, 39, .                                                                                                                                                                                            | 1.8 | 42        |

| #    | ARTICLE                                                                                                                                                                                                                                                 | IF  | CITATIONS |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 981  | RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. <i>Nucleic Acids Research</i> , 2023, 51, D488-D508. | 6.5 | 152       |
| 982  | DeepHomo2.0: improved protein-protein contact prediction of homodimers by transformer-enhanced deep learning. <i>Briefings in Bioinformatics</i> , 2023, 24, .                                                                                          | 3.2 | 14        |
| 983  | A large-scale benchmark study of tools for the classification of protein-coding and non-coding RNAs. <i>Nucleic Acids Research</i> , 2022, 50, 12094-12111.                                                                                             | 6.5 | 2         |
| 989  | Plant microbiomes harbor potential to promote nutrient turnover in impoverished substrates of a Brazilian biodiversity hotspot. <i>ISME Journal</i> , 2023, 17, 354-370.                                                                                | 4.4 | 14        |
| 991  | Adaptive lifestyle of bacteria determines phage-bacteria interaction. <i>Frontiers in Microbiology</i> , 0, 13, .                                                                                                                                       | 1.5 | 1         |
| 992  | Genome-Wide Analysis and Characterization of Eggplant F-Box Gene Superfamily: Gene Evolution and Expression Analysis under Stress. <i>International Journal of Molecular Sciences</i> , 2022, 23, 16049.                                                | 1.8 | 2         |
| 995  | Exploring a diverse world of effector domains and amyloid signaling motifs in fungal NLR proteins. <i>PLoS Computational Biology</i> , 2022, 18, e1010787.                                                                                              | 1.5 | 5         |
| 996  | Targeted single-cell genomics reveals novel host adaptation strategies of the symbiotic bacteria <i>Endozoicomonas</i> in <i>Acropora tenuis</i> coral. <i>Microbiome</i> , 2022, 10, .                                                                 | 4.9 | 6         |
| 998  | <sc>COLLAPSE</sc>: A representation learning framework for identification and characterization of protein structural sites. <i>Protein Science</i> , 2023, 32, .                                                                                        | 3.1 | 4         |
| 1000 | Prokaryotic and viral genomes recovered from 787 Japanese gut metagenomes revealed microbial features linked to diets, populations, and diseases. <i>Cell Genomics</i> , 2022, 2, 100219.                                                               | 3.0 | 8         |
| 1001 | Coastal water bacteriophages infect various sets of <i>Vibrio parahaemolyticus</i> sequence types. <i>Frontiers in Microbiology</i> , 0, 13, .                                                                                                          | 1.5 | 3         |
| 1002 | Artificial intelligence for template-free protein structure prediction: a comprehensive review. <i>Artificial Intelligence Review</i> , 2023, 56, 7665-7732.                                                                                            | 9.7 | 2         |
| 1003 | Single-sequence protein structure prediction using supervised transformer protein language models. <i>Nature Computational Science</i> , 2022, 2, 804-814.                                                                                              | 3.8 | 46        |
| 1004 | Structure-guided isoform identification for the human transcriptome. <i>ELife</i> , 0, 11, .                                                                                                                                                            | 2.8 | 14        |
| 1005 | Allosteric Activation of 15-lipoxygenase-1 by Boswellic Acid Induces the Lipid Mediator Class Switch to Promote Resolution of Inflammation. <i>Advanced Science</i> , 2023, 10, .                                                                       | 5.6 | 10        |
| 1006 | <sc>DPAM</sc>: A domain parser for <sc>AlphaFold</sc> models. <i>Protein Science</i> , 2023, 32, .                                                                                                                                                      | 3.1 | 17        |
| 1008 | Pharokka: a fast scalable bacteriophage annotation tool. <i>Bioinformatics</i> , 2023, 39, .                                                                                                                                                            | 1.8 | 53        |
| 1010 | Deep embedding and alignment of protein sequences. <i>Nature Methods</i> , 2023, 20, 104-111.                                                                                                                                                           | 9.0 | 13        |

| #    | ARTICLE                                                                                                                                                                                                                                         | IF   | CITATIONS |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 1016 | Multi-omics analyses reveal that the gut microbiome and its metabolites promote milk fat synthesis in Zhongdian yak cows. <i>PeerJ</i> , 0, 10, e14444.                                                                                         | 0.9  | 6         |
| 1018 | Global phylogenomic novelty of the Cas1 gene from hot spring microbial communities. <i>Frontiers in Microbiology</i> , 0, 13, .                                                                                                                 | 1.5  | 0         |
| 1019 | Diversity, duplication, and genomic organization of homeobox genes in Lepidoptera. <i>Genome Research</i> , 2023, 33, 32-44.                                                                                                                    | 2.4  | 11        |
| 1020 | Immunogenetic losses co-occurred with seahorse male pregnancy and mutation in <i>tlx1</i> accompanied functional asplenia. <i>Nature Communications</i> , 2022, 13, .                                                                           | 5.8  | 9         |
| 1021 | Characteristics of the GlnH and GlnX Signal Transduction Proteins Controlling PknG-Mediated Phosphorylation of OdhI and 2-Oxoglutarate Dehydrogenase Activity in <i>Corynebacterium glutamicum</i> . <i>Microbiology Spectrum</i> , 2022, 10, . | 1.2  | 3         |
| 1022 | Neural network-derived Potts models for structure-based protein design using backbone atomic coordinates and tertiary motifs. <i>Protein Science</i> , 2023, 32, .                                                                              | 3.1  | 3         |
| 1025 | Flavofun: Exploration of fungal flavoproteomes. <i>Frontiers in Catalysis</i> , 0, 2, .                                                                                                                                                         | 1.8  | 0         |
| 1026 | Reekekee- and roodoodooviruses, two different <i>Microviridae</i> clades constituted by the smallest DNA phages. <i>Virus Evolution</i> , 2023, 9, .                                                                                            | 2.2  | 5         |
| 1028 | Protein-to-genome alignment with miniprot. <i>Bioinformatics</i> , 2023, 39, .                                                                                                                                                                  | 1.8  | 56        |
| 1030 | Inside-Out Design of Zinc-Binding Proteins with Non-Native Backbones. <i>Biochemistry</i> , 2023, 62, 770-781.                                                                                                                                  | 1.2  | 4         |
| 1031 | Novel integrative elements and genomic plasticity in ocean ecosystems. <i>Cell</i> , 2023, 186, 47-62.e16.                                                                                                                                      | 13.5 | 25        |
| 1032 | CATHe: detection of remote homologues for CATH superfamilies using embeddings from protein language models. <i>Bioinformatics</i> , 2023, 39, .                                                                                                 | 1.8  | 8         |
| 1033 | A widespread family of phage-inducible chromosomal islands only steals bacteriophage tails to spread in nature. <i>Cell Host and Microbe</i> , 2023, 31, 69-82.e5.                                                                              | 5.1  | 9         |
| 1034 | ISPRED-SEQ: Deep Neural Networks and Embeddings for Predicting Interaction Sites in Protein Sequences. <i>Journal of Molecular Biology</i> , 2023, 435, 167963.                                                                                 | 2.0  | 9         |
| 1035 | The person-to-person transmission landscape of the gut and oral microbiomes. <i>Nature</i> , 2023, 614, 125-135.                                                                                                                                | 13.7 | 111       |
| 1036 | Genome-wide signatures of adaptation to extreme environments in red algae. <i>Nature Communications</i> , 2023, 14, .                                                                                                                           | 5.8  | 17        |
| 1038 | The first gapless, reference-quality, fully annotated genome from a Southern Han Chinese individual. <i>G3: Genes, Genomes, Genetics</i> , 2023, 13, .                                                                                          | 0.8  | 8         |
| 1039 | Transformer-based deep learning for predicting protein properties in the life sciences. <i>ELife</i> , 0, 12, .                                                                                                                                 | 2.8  | 29        |

| #    | ARTICLE                                                                                                                                                                                          | IF  | CITATIONS |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 1040 | A structural dendrogram of the actinobacteriophage major capsid proteins provides important structural insights into the evolution of capsid stability. <i>Structure</i> , 2023, 31, 282-294.e5. | 1.6 | 3         |
| 1041 | 3D-equivariant graph neural networks for protein model quality assessment. <i>Bioinformatics</i> , 2023, 39, .                                                                                   | 1.8 | 11        |
| 1042 | <i>coth</i> Genes Are Necessary for Normal Spore Formation and Virulence in <i>Mucor lusitanicus</i> . <i>MBio</i> , 2023, 14, .                                                                 | 1.8 | 10        |
| 1043 | HLA Class I Supertype Classification Based on Structural Similarity. <i>Journal of Immunology</i> , 2023, 210, 103-114.                                                                          | 0.4 | 3         |
| 1044 | Detecting macroevolutionary genotype–phenotype associations using error-corrected rates of protein convergence. <i>Nature Ecology and Evolution</i> , 2023, 7, 155-170.                          | 3.4 | 10        |
| 1045 | The META tool optimizes metagenomic analyses across sequencing platforms and classifiers. <i>Frontiers in Bioinformatics</i> , 0, 2, .                                                           | 1.0 | 0         |
| 1046 | CAT-Site: Predicting Protein Binding Sites Using a Convolutional Neural Network. <i>Pharmaceutics</i> , 2023, 15, 119.                                                                           | 2.0 | 3         |
| 1054 | Design of a pan-betacoronavirus vaccine candidate through a phylogenetically informed approach. <i>Science Advances</i> , 2023, 9, .                                                             | 4.7 | 2         |
| 1055 | High-resolution metagenomic reconstruction of the freshwater spring bloom. <i>Microbiome</i> , 2023, 11, .                                                                                       | 4.9 | 9         |
| 1056 | ProtRAP: Predicting Lipid Accessibility Together with Solvent Accessibility of Proteins in One Run. <i>Journal of Chemical Information and Modeling</i> , 2023, 63, 1058-1065.                   | 2.5 | 0         |
| 1057 | Nanopore-based enrichment of antimicrobial resistance genes – a case-based study. <i>GigaByte</i> , 0, 2023, 1-15.                                                                               | 0.0 | 3         |
| 1058 | Functional metagenomics profiling of symbiotic microbiome. , 2023, , 691-713.                                                                                                                    |     | 1         |
| 1060 | Viruses Regulate Microbial Community Assembly Together with Environmental Factors in Acid Mine Drainage. <i>Applied and Environmental Microbiology</i> , 2023, 89, .                             | 1.4 | 1         |
| 1061 | Lessons on fruiting body morphogenesis from genomes and transcriptomes of <i>Agaricomycetes</i> . <i>Studies in Mycology</i> , 2023, 104, 1-85.                                                  | 4.5 | 9         |
| 1063 | Isolation and infection cycle of a polinton-like virus virophage in an abundant marine alga. <i>Nature Microbiology</i> , 2023, 8, 332-346.                                                      | 5.9 | 15        |
| 1065 | A conserved motif suggests a common origin for a group of proteins involved in the cell division of Gram-positive bacteria. <i>PLoS ONE</i> , 2023, 18, e0273136.                                | 1.1 | 1         |
| 1068 | Using evolutionary data to make sense of macromolecules with a –ConSurf. <i>Protein Science</i> , 2023, 32, .                                                                                    | 3.1 | 73        |
| 1071 | Multienzyme deep learning models improve peptide de novo sequencing by mass spectrometry proteomics. <i>PLoS Computational Biology</i> , 2023, 19, e1010457.                                     | 1.5 | 2         |

| #    | ARTICLE                                                                                                                                                                                                                              | IF   | CITATIONS |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 1072 | Phylogeny and cross-regulation of the YjiM and LeuO transcription factors translated as multiple protein forms from one gene in <i>Escherichia coli</i> . <i>Mathematical Biology and Bioinformatics</i> , 2023, 18, 1-14.           | 0.1  | 0         |
| 1074 | Genetic and Structural Diversity of Prokaryotic Ice-Binding Proteins from the Central Arctic Ocean. <i>Genes</i> , 2023, 14, 363.                                                                                                    | 1.0  | 5         |
| 1076 | Snekmer: a scalable pipeline for protein sequence fingerprinting based on amino acid recoding. <i>Bioinformatics Advances</i> , 2023, 3, .                                                                                           | 0.9  | 1         |
| 1077 | The FANCCâ€“FANCEâ€“FANCF complex is evolutionarily conserved and regulates meiotic recombination. <i>Nucleic Acids Research</i> , 2023, 51, 2516-2528.                                                                              | 6.5  | 5         |
| 1078 | RNA sequencing indicates widespread conservation of circadian clocks in marine zooplankton. <i>NAR Genomics and Bioinformatics</i> , 2023, 5, .                                                                                      | 1.5  | 1         |
| 1079 | A chromosome-scale reference genome assembly of the great sand eel, <i>Hyperoplus lanceolatus</i> . <i>Journal of Heredity</i> , 2023, 114, 189-194.                                                                                 | 1.0  | 2         |
| 1080 | Extreme-Scale Many-against-Many Protein Similarity Search. , 2022, , .                                                                                                                                                               |      | 0         |
| 1081 | Comparison of methods for biological sequence clustering. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2023, , 1-24.                                                                                   | 1.9  | 0         |
| 1082 | Protein Loop Modeling Using AlphaFold2. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2023, , 1-8.                                                                                                      | 1.9  | 1         |
| 1083 | Revised classification of the Cyanidiophyceae based on plastid genome data with descriptions of the <i>Cavernicolales</i> ord. nov. and <i>Galdieriales</i> ord. nov. (Rhodophyta). <i>Journal of Phycology</i> , 2023, 59, 444-466. | 1.0  | 11        |
| 1084 | Reverse engineering environmental metatranscriptomes clarifies best practices for eukaryotic assembly. <i>BMC Bioinformatics</i> , 2023, 24, .                                                                                       | 1.2  | 2         |
| 1086 | Effect of Temperature on Microorganisms and Nitrogen Removal in a Multi-Stage Surface Flow Constructed Wetland. <i>Water (Switzerland)</i> , 2023, 15, 1256.                                                                         | 1.2  | 2         |
| 1087 | Chromosome-Level Genome Assembly of <i>Papilio elwesi</i> Leech, 1889 (Lepidoptera: Papilionidae). <i>Insects</i> , 2023, 14, 304.                                                                                                   | 1.0  | 0         |
| 1088 | Competition between lysogenic and sensitive bacteria is determined by the fitness costs of the different emerging phage-resistance strategies. <i>ELife</i> , 0, 12, .                                                               | 2.8  | 9         |
| 1090 | The structure of biodegradable surfactants shaped the microbial community, antimicrobial resistance, and potential for horizontal gene transfer. <i>Water Research</i> , 2023, 236, 119944.                                          | 5.3  | 2         |
| 1093 | The little skate genome and the evolutionary emergence of wing-like fins. <i>Nature</i> , 2023, 616, 495-503.                                                                                                                        | 13.7 | 14        |
| 1094 | Orthogonal analysis of variants in APOE gene using in-silico approaches reveals novel disrupting variants. <i>Frontiers in Bioinformatics</i> , 0, 3, .                                                                              | 1.0  | 0         |
| 1095 | Towards the accurate alignment of over a million protein sequences: Current state of the art. <i>Current Opinion in Structural Biology</i> , 2023, 80, 102577.                                                                       | 2.6  | 1         |

| #    | ARTICLE                                                                                                                                                                                                                               | IF   | CITATIONS |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 1096 | Progress at protein structure prediction, as seen in CASP15. <i>Current Opinion in Structural Biology</i> , 2023, 80, 102594.                                                                                                         | 2.6  | 26        |
| 1097 | Susceptibility and transcriptomic response to plasma-activated water of <i>Listeria monocytogenes</i> planktonic and sessile cells. <i>Food Microbiology</i> , 2023, 113, 104252.                                                     | 2.1  | 3         |
| 1098 | The defense island repertoire of the <i>Escherichia coli</i> pan-genome. <i>PLoS Genetics</i> , 2023, 19, e1010694.                                                                                                                   | 1.5  | 19        |
| 1100 | Mining metatranscriptomes reveals a vast world of viroid-like circular RNAs. <i>Cell</i> , 2023, 186, 646-661.e4.                                                                                                                     | 13.5 | 23        |
| 1101 | Closed genomes uncover a saltwater species of <i>Candidatus</i> <i>Electronema</i> and shed new light on the boundary between marine and freshwater cable bacteria. <i>ISME Journal</i> , 2023, 17, 561-569.                          | 4.4  | 10        |
| 1102 | Annelid functional genomics reveal the origins of bilaterian life cycles. <i>Nature</i> , 2023, 615, 105-110.                                                                                                                         | 13.7 | 34        |
| 1105 | Epitopedia: identifying molecular mimicry between pathogens and known immune epitopes. <i>Immunoinformatics</i> , 2023, 9, 100023.                                                                                                    | 1.2  | 5         |
| 1106 | RCSB Protein Data Bank: Efficient Searching and Simultaneous Access to One Million Computed Structure Models Alongside the PDB Structures Enabled by Architectural Advances. <i>Journal of Molecular Biology</i> , 2023, 435, 167994. | 2.0  | 19        |
| 1107 | Mining <i>Chromodoris quadricolor</i> symbionts for biosynthesis of novel secondary metabolites. <i>Marine Genomics</i> , 2023, 68, 101017.                                                                                           | 0.4  | 1         |
| 1109 | Phylogeny analysis of whole protein-coding genes in metagenomic data detected an environmental gradient for the microbiota. <i>PLoS ONE</i> , 2023, 18, e0281288.                                                                     | 1.1  | 1         |
| 1110 | Cytidine deaminases catalyze the conversion of <i>N</i> ( <i>S</i> , <i>O</i> ) <sup>4</sup> -substituted pyrimidine nucleosides. <i>Science Advances</i> , 2023, 9, .                                                                | 4.7  | 1         |
| 1116 | Sequence and structure alignments in post-AlphaFold era. <i>Current Opinion in Structural Biology</i> , 2023, 79, 102539.                                                                                                             | 2.6  | 2         |
| 1117 | Bacteriophages inhibit and evade cGAS-like immune function in bacteria. <i>Cell</i> , 2023, 186, 864-876.e21.                                                                                                                         | 13.5 | 49        |
| 1119 | A nontuberculous mycobacterium could solve the mystery of the lady from the Franciscan church in Basel, Switzerland. <i>BMC Biology</i> , 2023, 21, .                                                                                 | 1.7  | 1         |
| 1120 | Comple+: a computationally scalable method to improve completeness of large-scale protein sequence clustering. <i>PeerJ</i> , 0, 11, e14779.                                                                                          | 0.9  | 0         |
| 1121 | Metagenomic insights into the effects of benzyl dodecyl dimethyl ammonium bromide (BDAB) shock on bacteria-driven nitrogen removal in a moving-bed biofilm reactor (MBBR). <i>Chemosphere</i> , 2023, 320, 138098.                    | 4.2  | 6         |
| 1123 | Machine Learning Advances in Predicting Peptide/Protein-Protein Interactions Based on Sequence Information for Lead Peptides Discovery. <i>Advanced Biology</i> , 2023, 7, .                                                          | 1.4  | 5         |
| 1124 | gutSMASH predicts specialized primary metabolic pathways from the human gut microbiota. <i>Nature Biotechnology</i> , 2023, 41, 1416-1423.                                                                                            | 9.4  | 24        |

| #    | ARTICLE                                                                                                                                                                              | IF  | CITATIONS |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 1126 | Dataset of 143 metagenome-assembled genomes from the Arctic and Atlantic Oceans, including 21 for eukaryotic organisms. <i>Data in Brief</i> , 2023, 47, 108990.                     | 0.5 | 1         |
| 1129 | MGnify Genomes: A Resource for Biome-specific Microbial Genome Catalogues. <i>Journal of Molecular Biology</i> , 2023, 435, 168016.                                                  | 2.0 | 6         |
| 1131 | VirBot: an RNA viral contig detector for metagenomic data. <i>Bioinformatics</i> , 2023, 39, .                                                                                       | 1.8 | 5         |
| 1132 | The NagY regulator: A member of the BglG/SacY antiterminator family conserved in <i>Enterococcus faecalis</i> and involved in virulence. <i>Frontiers in Microbiology</i> , 0, 13, . | 1.5 | 0         |
| 1134 | Enzyme Substrate Prediction from Three-Dimensional Feature Representations Using Space-Filling Curves. <i>Journal of Chemical Information and Modeling</i> , 2023, 63, 1637-1648.    | 2.5 | 1         |
| 1135 | VPAGs-Dataset4ML: A Dataset to Predict Viral Protective Antigens for Machine Learning-Based Reverse Vaccinology. <i>Data</i> , 2023, 8, 41.                                          | 1.2 | 1         |
| 1136 | Structural screens identify candidate human homologs of insect chemoreceptors and cryptic <i>Drosophila</i> gustatory receptor-like proteins. <i>ELife</i> , 0, 12, .                | 2.8 | 4         |
| 1137 | Variation in microbial CAZyme families across degradation severity in a steppe grassland in northern China. <i>Frontiers in Environmental Science</i> , 0, 11, .                     | 1.5 | 3         |
| 1141 | Giant proteins in a giant cell: Molecular basis of ultrafast Ca <sup>2+</sup> -dependent cell contraction. <i>Science Advances</i> , 2023, 9, .                                      | 4.7 | 4         |
| 1145 | Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlan 4. <i>Nature Biotechnology</i> , 2023, 41, 1633-1644.                           | 9.4 | 132       |
| 1146 | Multiomic spatial analysis reveals a distinct mucosa-associated virome. <i>Gut Microbes</i> , 2023, 15, .                                                                            | 4.3 | 5         |
| 1147 | Population-level impacts of antibiotic usage on the human gut microbiome. <i>Nature Communications</i> , 2023, 14, .                                                                 | 5.8 | 15        |
| 1148 | Cobalamin producers and prokaryotic consumers in the Northwest Atlantic. <i>Environmental Microbiology</i> , 2023, 25, 1300-1313.                                                    | 1.8 | 0         |
| 1149 | Exploring microbial functional biodiversity at the protein family level—From metagenomic sequence reads to annotated protein clusters. <i>Frontiers in Bioinformatics</i> , 0, 3, .  | 1.0 | 2         |
| 1151 | A chromosome-scale high-contiguity genome assembly of the cheetah ( <i>Acinonyx jubatus</i> ). <i>Journal of Heredity</i> , 0, , .                                                   | 1.0 | 0         |
| 1152 | Identification and characterization of thousands of bacteriophage satellites across bacteria. <i>Nucleic Acids Research</i> , 2023, 51, 2759-2777.                                   | 6.5 | 19        |
| 1157 | TopEnzyme: a framework and database for structural coverage of the functional enzyme space. <i>Bioinformatics</i> , 2023, 39, .                                                      | 1.8 | 2         |
| 1158 | A hydrogenotrophic <i>Sulfurimonas</i> is globally abundant in deep-sea oxygen-saturated hydrothermal plumes. <i>Nature Microbiology</i> , 2023, 8, 651-665.                         | 5.9 | 5         |



| #    | ARTICLE                                                                                                                                                                                                                                                                                                                                                                                                 | IF  | CITATIONS |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 1161 | Acinetobacter nematophilus sp. nov., Alcaligenes nematophilus sp. nov., Enterobacter nematophilus sp. nov., and Kaistia nematophila sp. nov., Isolated from Soil-Borne Nematodes and Proposal for the Elevation of Alcaligenes faecalis subsp. faecalis, Alcaligenes faecalis subsp. parafaecalis, and Alcaligenes faecalis subsp. phenolicus to the Species Level. <i>Taxonomy</i> , 2023, 3, 148-168. | 0.4 | 5         |
| 1162 | Widespread Distribution and Evolution of Poxviral Entry-Fusion Complex Proteins in Giant Viruses. <i>Microbiology Spectrum</i> , 2023, 11, .                                                                                                                                                                                                                                                            | 1.2 | 8         |
| 1163 | <scpiGDP</scpi>: An integrated genome decontamination pipeline for wild ciliated microeukaryotes. <i>Molecular Ecology Resources</i> , 2023, 23, 1182-1193.                                                                                                                                                                                                                                             | 2.2 | 2         |
| 1164 | Seasonality of biogeochemically relevant microbial genes in a coastal ocean microbiome. <i>Environmental Microbiology</i> , 2023, 25, 1465-1483.                                                                                                                                                                                                                                                        | 1.8 | 1         |
| 1165 | Site-specific analysis reveals candidate cross-kingdom small RNAs, tRNA and rRNA fragments, and signs of fungal RNA phasing in the barley-powdery mildew interaction. <i>Molecular Plant Pathology</i> , 2023, 24, 570-587.                                                                                                                                                                             | 2.0 | 11        |
| 1169 | Evolutionary-scale prediction of atomic-level protein structure with a language model. <i>Science</i> , 2023, 379, 1123-1130.                                                                                                                                                                                                                                                                           | 6.0 | 623       |
| 1170 | Large-Scale Identification of Known and Novel RRNPP Quorum-Sensing Systems by RRNPP_Detector Captures Novel Features of Bacterial, Plasmidic, and Viral Coevolution. <i>Molecular Biology and Evolution</i> , 2023, 40, .                                                                                                                                                                               | 3.5 | 1         |
| 1171 | An informatic workflow for the enhanced annotation of excretory/secretory proteins of <i>Haemonchus contortus</i> . <i>Computational and Structural Biotechnology Journal</i> , 2023, 21, 2696-2704.                                                                                                                                                                                                    | 1.9 | 1         |
| 1172 | <i>Treponema pallidum</i> subsp. <i>pallidum</i> with an Artificially impaired TprK antigenic variation system is attenuated in the Rabbit model of syphilis. <i>PLoS Pathogens</i> , 2023, 19, e1011259.                                                                                                                                                                                               | 2.1 | 6         |
| 1174 | Poincaré maps for visualization of large protein families. <i>Briefings in Bioinformatics</i> , 2023, 24, .                                                                                                                                                                                                                                                                                             | 3.2 | 3         |
| 1175 | Phylogroup-specific variation shapes the clustering of antimicrobial resistance genes and defence systems across regions of genome plasticity in <i>Pseudomonas aeruginosa</i> . <i>EBioMedicine</i> , 2023, 90, 104532.                                                                                                                                                                                | 2.7 | 11        |
| 1177 | Uncovering gene-family founder events during major evolutionary transitions in animals, plants and fungi using GenEra. <i>Genome Biology</i> , 2023, 24, .                                                                                                                                                                                                                                              | 3.8 | 9         |
| 1178 | Foldcomp: a library and format for compressing and indexing large protein structure sets. <i>Bioinformatics</i> , 2023, 39, .                                                                                                                                                                                                                                                                           | 1.8 | 9         |
| 1179 | Effective Molecular Dynamics from Neural Network-Based Structure Prediction Models. <i>Journal of Chemical Theory and Computation</i> , 2023, 19, 1965-1975.                                                                                                                                                                                                                                            | 2.3 | 10        |
| 1181 | An OmpW-dependent T4-like phage infects <i>Serratia</i> sp. ATCC 39006. <i>Microbial Genomics</i> , 2023, 9, .                                                                                                                                                                                                                                                                                          | 1.0 | 0         |
| 1182 | Enzyme function prediction using contrastive learning. <i>Science</i> , 2023, 379, 1358-1363.                                                                                                                                                                                                                                                                                                           | 6.0 | 59        |
| 1183 | ViroProfiler: a containerized bioinformatics pipeline for viral metagenomic data analysis. <i>Gut Microbes</i> , 2023, 15, .                                                                                                                                                                                                                                                                            | 4.3 | 10        |
| 1186 | Cellular differentiation into hyphae and spores in halophilic archaea. <i>Nature Communications</i> , 2023, 14, .                                                                                                                                                                                                                                                                                       | 5.8 | 5         |

| #    | ARTICLE                                                                                                                                                                                     | IF   | CITATIONS |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 1189 | Plant domestication shapes rhizosphere microbiome assembly and metabolic functions. <i>Microbiome</i> , 2023, 11, .                                                                         | 4.9  | 33        |
| 1191 | Assessment of plasmids for relating the 2020 <i>Salmonella enterica</i> serovar Newport onion outbreak to farms implicated by the outbreak investigation. <i>BMC Genomics</i> , 2023, 24, . | 1.2  | 0         |
| 1193 | New carbohydrate binding domains identified by phage display based functional metagenomic screens of human gut microbiota. <i>Communications Biology</i> , 2023, 6, .                       | 2.0  | 2         |
| 1194 | TnpB structure reveals minimal functional core of Cas12 nuclease family. <i>Nature</i> , 2023, 616, 384-389.                                                                                | 13.7 | 27        |
| 1196 | Viruses interact with hosts that span distantly related microbial domains in dense hydrothermal mats. <i>Nature Microbiology</i> , 2023, 8, 946-957.                                        | 5.9  | 15        |
| 1197 | Omics data integration facilitates target selection for new antiparasitic drugs against <i>TriTryp</i> infections. <i>Frontiers in Pharmacology</i> , 0, 14, .                              | 1.6  | 1         |
| 1198 | Quantifying Shared and Unique Gene Content across 17 Microbial Ecosystems. <i>MSystems</i> , 0, , .                                                                                         | 1.7  | 0         |
| 1200 | Analysis of the <i>P.Âlividus</i> sea urchin genome highlights contrasting trends of genomic and regulatory evolution in deuterostomes. <i>Cell Genomics</i> , 2023, 3, 100295.             | 3.0  | 11        |
| 1201 | Metagenomic insights into the composition and function of the gut microbiota of mice infected with <i>Toxoplasma gondii</i> . <i>Frontiers in Immunology</i> , 0, 14, .                     | 2.2  | 2         |
| 1202 | Multilevel interrogation of H3.3 reveals a primordial role in transcription regulation. <i>Epigenetics and Chromatin</i> , 2023, 16, .                                                      | 1.8  | 1         |
| 1207 | In Silico and In Vivo Studies of a Tumor-Penetrating and Interfering Peptide with Antitumoral Effect on Xenograft Models of Breast Cancer. <i>Pharmaceutics</i> , 2023, 15, 1180.           | 2.0  | 0         |
| 1208 | Large-scale invasion of unicellular eukaryotic genomes by integrating DNA viruses. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2023, 120, .    | 3.3  | 19        |
| 1209 | Total escape of SARS-CoV-2 from dual monoclonal antibody therapy in an immunocompromised patient. <i>Nature Communications</i> , 2023, 14, .                                                | 5.8  | 5         |
| 1211 | Identifying B-cell epitopes using AlphaFold2 predicted structures and pretrained language model. <i>Bioinformatics</i> , 2023, 39, .                                                        | 1.8  | 4         |
| 1213 | Staphylococcal diversity in atopic dermatitis from an individual to a global scale. <i>Cell Host and Microbe</i> , 2023, 31, 578-592.e6.                                                    | 5.1  | 9         |
| 1214 | Mutation-induced infections of phage-plasmids. <i>Nature Communications</i> , 2023, 14, .                                                                                                   | 5.8  | 5         |
| 1215 | Chromosome-length genome assemblies and cytogenomic analyses of pangolins reveal remarkable chromosome counts and plasticity. <i>Chromosome Research</i> , 2023, 31, .                      | 1.0  | 3         |
| 1217 | Profiling the disintegration of BRPs released by massive wasp stings using serratiopeptidase: An in-silico insight. <i>Computers in Biology and Medicine</i> , 2023, 159, 106951.           | 3.9  | 3         |

| #    | ARTICLE                                                                                                                                                                                                            | IF   | CITATIONS |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------|
| 1218 | Domesticating <i>Vigna stipulacea</i> : Chromosome-Level genome assembly reveals VsPSAT1 as a candidate gene decreasing hard-seededness. <i>Frontiers in Plant Science</i> , 0, 14, .                              | 1.7  | 2         |
| 1221 | Plant-on-chip: Core morphogenesis processes in the tiny plant <i>Wolffia australiana</i> . , 2023, 2, .                                                                                                            |      | 2         |
| 1222 | Defense systems are pervasive across chromosomally integrated mobile genetic elements and are inversely correlated to virulence and antimicrobial resistance. <i>Nucleic Acids Research</i> , 2023, 51, 4385-4397. | 6.5  | 10        |
| 1223 | FAS: assessing the similarity between proteins using multi-layered feature architectures. <i>Bioinformatics</i> , 2023, 39, .                                                                                      | 1.8  | 2         |
| 1224 | Comparative genomics identifies conserved and variable TAL effectors in African strains of the cotton pathogen <i>Xanthomonas citri</i> pv. <i>malvacearum</i> . <i>Phytopathology</i> , 0, , .                    | 1.1  | 0         |
| 1233 | Fast and accurate protein structure search with Foldseek. <i>Nature Biotechnology</i> , 2024, 42, 243-246.                                                                                                         | 9.4  | 254       |
| 1270 | Viabin: A novel method for Overlapped Binning of Metagenomic Contigs using ZINB-autodecoder and Assembly Graphs. , 2023, , .                                                                                       |      | 0         |
| 1368 | Open-Source Machine Learning in Computational Chemistry. <i>Journal of Chemical Information and Modeling</i> , 2023, 63, 4505-4532.                                                                                | 2.5  | 3         |
| 1423 | Pre-training Antibody Language Models for Antigen-Specific Computational Antibody Design. , 2023, , .                                                                                                              |      | 0         |
| 1446 | Machine Learning for Protein Engineering. Challenges and Advances in Computational Chemistry and Physics, 2023, , 277-311.                                                                                         | 0.6  | 1         |
| 1479 | Stem Cells Application in Eye Regeneration and Restoration of Vision. , 2023, , 1-31.                                                                                                                              |      | 0         |
| 1548 | Machine learning for microbiologists. <i>Nature Reviews Microbiology</i> , 2024, 22, 191-205.                                                                                                                      | 13.6 | 6         |
| 1670 | De Novo Design of Target-Specific Ligands Using BERT-Pretrained Transformer. <i>Lecture Notes in Computer Science</i> , 2024, , 311-322.                                                                           | 1.0  | 0         |
| 1693 | Semi-Supervised and Incremental Sequence Analysis for Taxonomic Classification. , 2023, , .                                                                                                                        |      | 0         |
| 1699 | A computational pipeline elucidating functions of conserved hypothetical <i>Trypanosoma cruzi</i> proteins based on public proteomic data. <i>Advances in Protein Chemistry and Structural Biology</i> , 2024, , . | 1.0  | 0         |
| 1723 | SLPFA: Protein Structure-Label Embedding Attention Network for Protein Function Annotation. , 2023, , .                                                                                                            |      | 0         |
| 1779 | Identification of drought stress genes expressed in <i>Jatropha curcas</i> by using RNA sequencing. <i>AIP Conference Proceedings</i> , 2024, , .                                                                  | 0.3  | 0         |
| 1808 | Shotgun Metagenomic Approaches. , 2024, , 43-54.                                                                                                                                                                   |      | 0         |

| #    | ARTICLE                                                                                                                         | IF  | CITATIONS |
|------|---------------------------------------------------------------------------------------------------------------------------------|-----|-----------|
| 1843 | Structural Modeling of T9SS Outer Membrane Proteins and Their Complexes. <i>Methods in Molecular Biology</i> , 2024, , 331-344. | 0.4 | 0         |