



Comparative Genomics and Metagenomics Lab (CGMLab)

<https://cgmlab.org>

The **CGMLab** is a bioinformatics research group based at the [CBGP](#), dedicated to deciphering the unique characteristics of organisms and microbial ecosystems. Our team adopts an interdisciplinary approach, combining expertise in evolutionary biology, computer science, ecology, and molecular biology with high-performance computing to analyze complex sequencing data.

Our primary research objective is to characterize the **Unknown Microbial Biosphere**. We address this challenge by building computational frameworks for comparative metagenomics, developing novel genomic tools, and analyzing the microbiome at a global scale to uncover functional diversity that remains hidden in standard analyses.

Key Resources & Tools The lab is a prolific developer of open-source bioinformatics software and databases that are widely used in the scientific community. Our key resources include:

- **eggNOG & eggNOG-mapper:** A leading resource for orthology prediction and functional annotation at the metagenomic scale (currently in v6.0).
- **ETE Toolkit (ETE 3):** A comprehensive Python framework for the reconstruction, analysis, and visualization of phylogenomic data.
- **FESNOV:** A resource dedicated to the exploration of novel microbial gene families and the functional/evolutionary significance of unknown genes.
- **GeCoViz & PhyloCloud:** Specialized online platforms for genomic context visualization and the management/exploration of phylogenomic trees.

Selected Publications

Research Articles & Biological Discovery

- **Unveiling the hidden biodiversity of microbial chemoreceptor genes in the rare biosphere.** Sanchis-López C, Santamaría-Hernando S, Cantalapiedra CP, Pokorny L, González-Recio O, Sunagawa S, López-Solanilla E, Huerta-Cepas J. *bioRxiv*. 2025. DOI: [10.1101/2025.06.24.661292](https://doi.org/10.1101/2025.06.24.661292)
- **Functional and evolutionary significance of unknown genes from uncultivated taxa.** Rodríguez Del Río Á, Giner-Lamia J, Cantalapiedra CP, Botas J, Deng Z, Hernández-Plaza A, Munar-Palmer M, Santamaría-Hernando S, Rodríguez-Herva JJ, Ruscheweyh HJ, Paoli L, Schmidt TSB, Sunagawa S, Bork P, López-Solanilla E, Coelho LP, Huerta-Cepas J. *Nature*. 2023 Dec 18. DOI: [10.1038/s41586-023-06955-z](https://doi.org/10.1038/s41586-023-06955-z)

- **Comparison of gene clustering criteria reveals intrinsic uncertainty in pangenome analyses.** Manzano-Morales S, Liu Y, González-Bodí S, Huerta-Cepas J, Iranzo J. *Genome Biology*. 2023 Oct 30;24(1):250. [DOI: 10.1186/s13059-023-03089-3](https://doi.org/10.1186/s13059-023-03089-3)
- **Towards the biogeography of prokaryotic genes.** Coelho LP, Alves R, Del Río ÁR, Myers PN, Cantalapiedra CP, Giner-Lamia J, Schmidt TS, Mende DR, Orakov A, Letunic I, Hildebrand F, Van Rossum T, Forslund SK, Khedkar S, Maistrenko OM, Pan S, Jia L, Ferretti P, Sunagawa S, Zhao XM, Nielsen HB, Huerta-Cepas J, Bork P. *Nature*. 2022 Jan;601(7892):252-256. [DOI: 10.1038/s41586-021-04233-4](https://doi.org/10.1038/s41586-021-04233-4)
- **Prevalence and Specificity of Chemoreceptor Profiles in Plant-Associated Bacteria.** Sanchis-López C, Cerna-Vargas JP, Santamaría-Hernando S, Ramos C, Krell T, Rodríguez-Palenzuela P, López-Solanilla E, Huerta-Cepas J, Rodríguez-Herva JJ. *mSystems*. 2021 Oct 26;6(5):e0095121. [DOI: 10.1128/mSystems.00951-21](https://doi.org/10.1128/mSystems.00951-21)

Tools and Resources

- **eggNOG 6.0: enabling comparative genomics across 12,535 organisms.** Hernández-Plaza A, Szklarczyk D, Botas J, Cantalapiedra CP, Giner-Lamia J, Mende DR, Kirsch R, Rattei T, Letunic I, Jensen LJ, Bork P, von Mering C, Huerta-Cepas J. *Nucleic Acids Research*. 2022 Nov 18. [DOI: 10.1093/nar/gkac1022](https://doi.org/10.1093/nar/gkac1022)
- **GeCoViz: genomic context visualisation of prokaryotic genes from a functional and evolutionary perspective.** Botas J, Rodríguez Del Río Á, Giner-Lamia J, Huerta-Cepas J. *Nucleic Acids Research*. 2022 May 26;50(W1):W352-7. [DOI: 10.1093/nar/gkac367](https://doi.org/10.1093/nar/gkac367)
- **eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale.** Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. *Molecular Biology and Evolution*. 2021 Dec 9;38(12):5825-5829. [DOI: 10.1093/molbev/msab293](https://doi.org/10.1093/molbev/msab293)
- **ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data.** Huerta-Cepas J, Serra F, Bork P. *Molecular Biology and Evolution*. 2016 Jun;33(6):1635-8. [DOI: 10.1093/molbev/msw046](https://doi.org/10.1093/molbev/msw046)