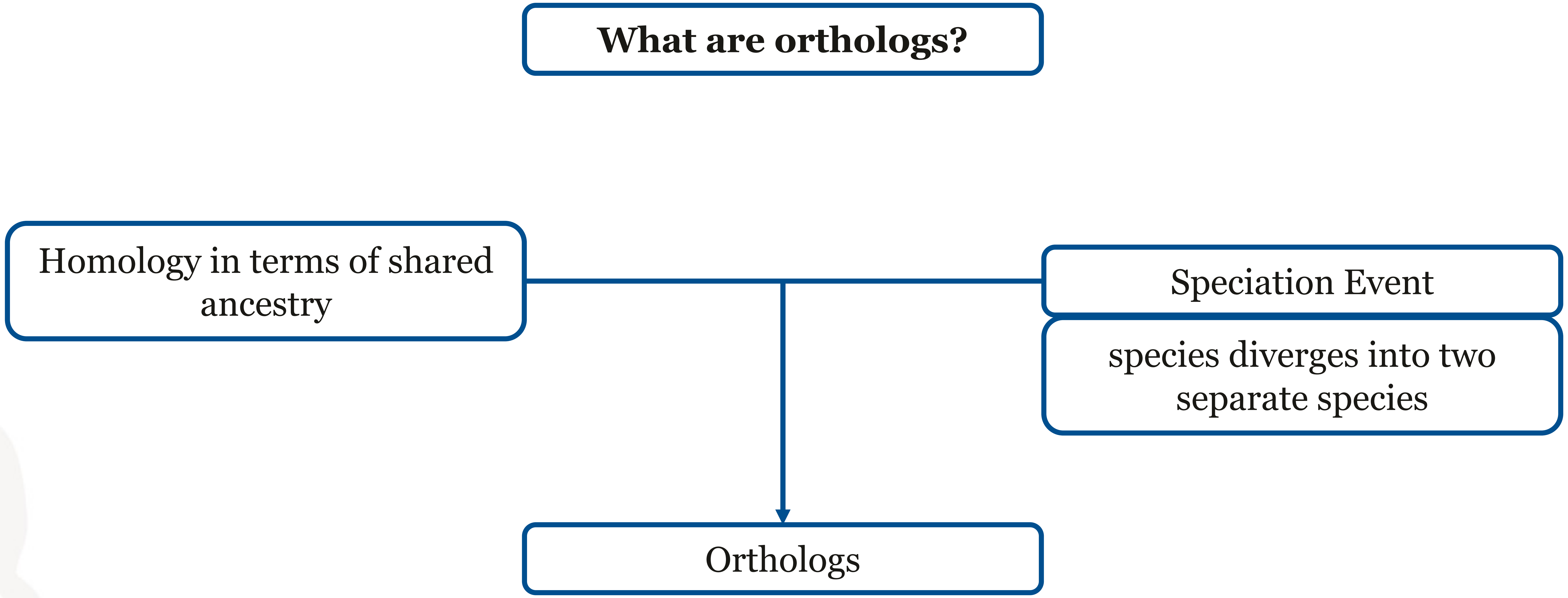
Jonas Elpelt & Julian Rummel

Standardized benchmarking in the quest for orthologs

Altenhoff, A.[1], Boeckmann, B.[2], Capella-Gutierrez, S.[3] *et al.*
in *Nat Methods* **13,** 425–430 (2016). doi: https://doi.org/10.1038/nmeth.3830

[1]Department of Computer Science & Computational Biochemistry Research Group, ETH Zurich & Swiss Institute of Bioinformatics, Switzerland
[2]Swiss-Prot Group, Swiss Institute of Bioinformatics, Geneva, Switzerland
[3]Bioinformatics and Genomics Programme, Centre for Genomic Regulation, Barcelona, Spain

The quest for orthologs



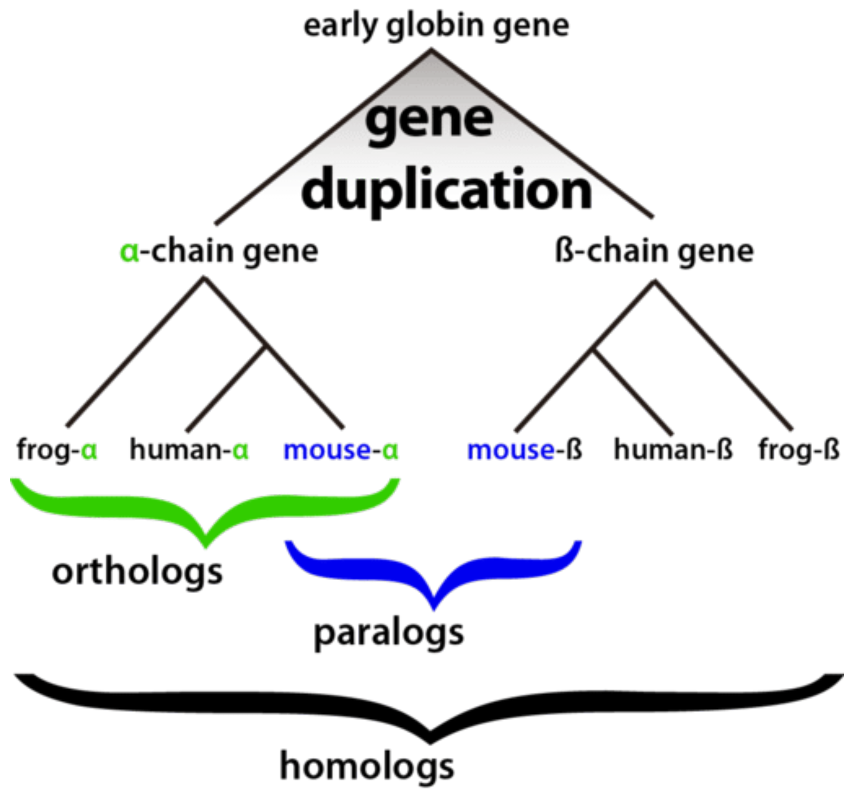
What are orthologs?

Homology in terms of shared ancestry

Speciation Event

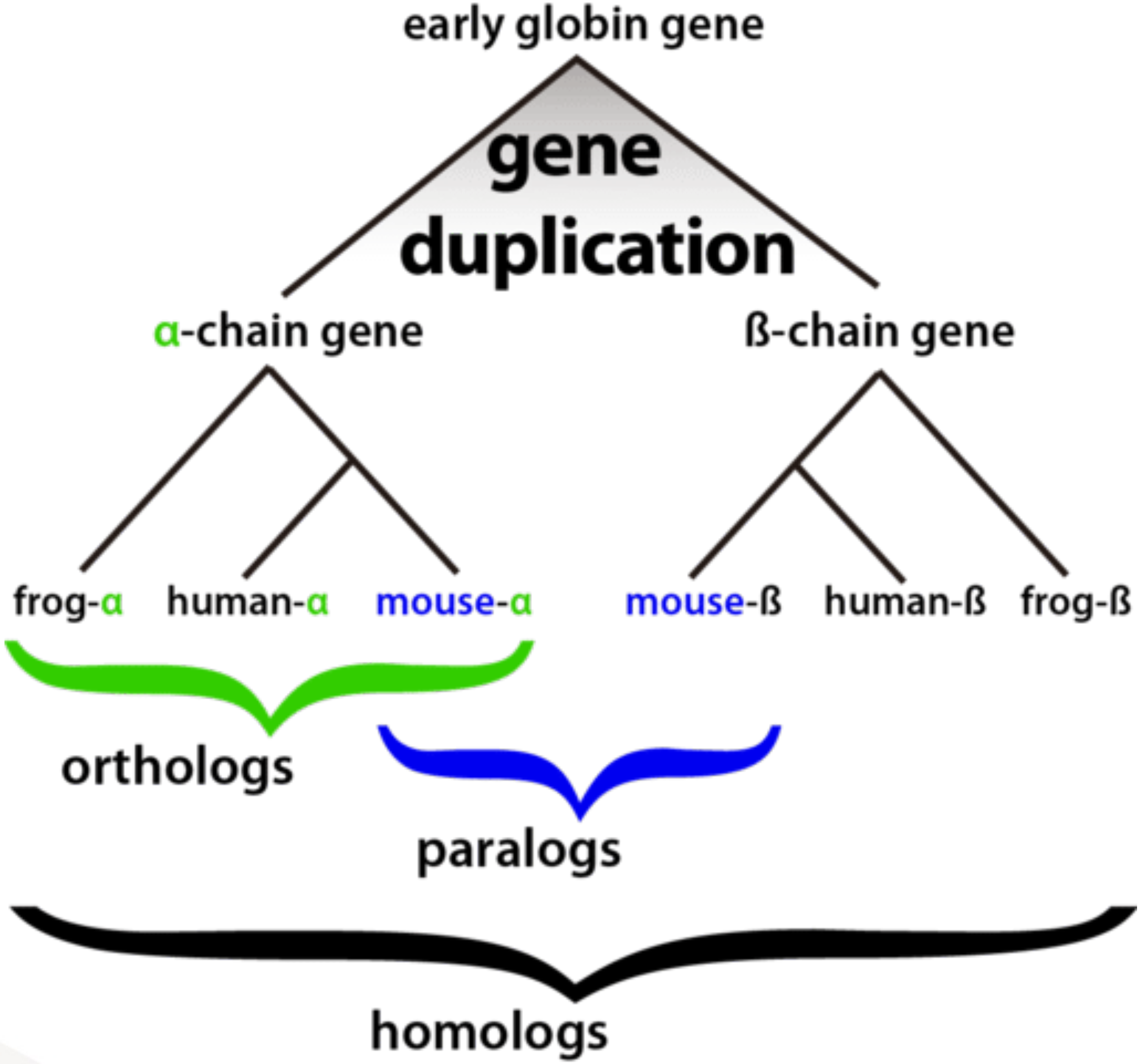
species diverges into two separate species

Orthologs



Quest for orthologs (https://bitesizebio.com/26762/homology-terminology-never-say-wrong-word/)

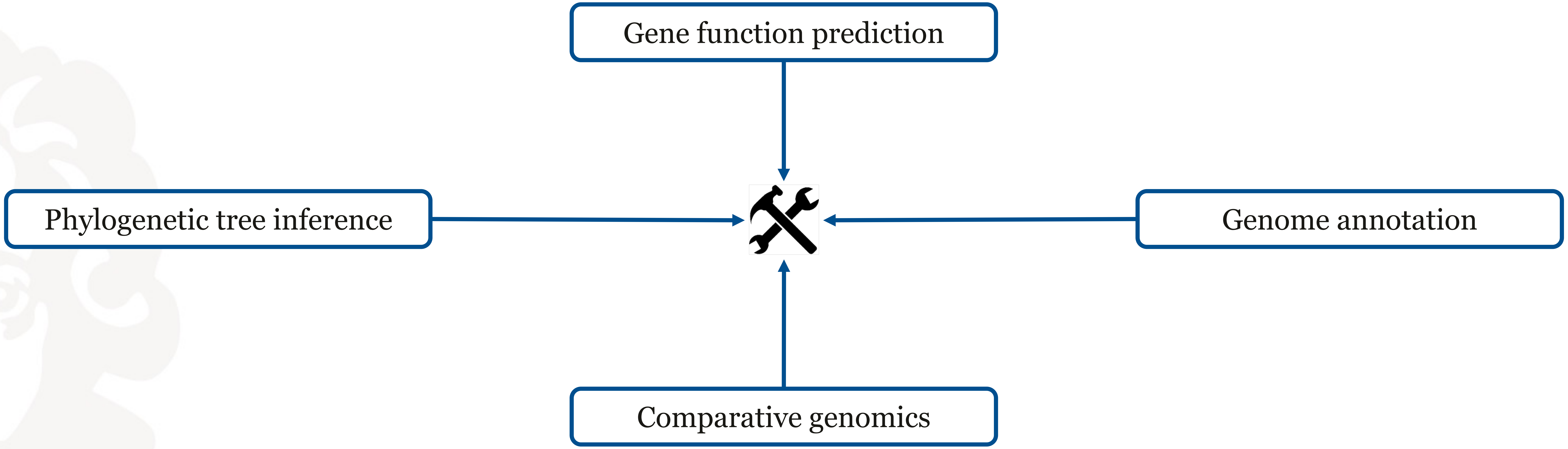
The quest for orthologs



Quest for orthologs (https://bitesizebio.com/26762/homology-terminology-never-say-wrong-word/)

The quest for orthologs

Why is it important?



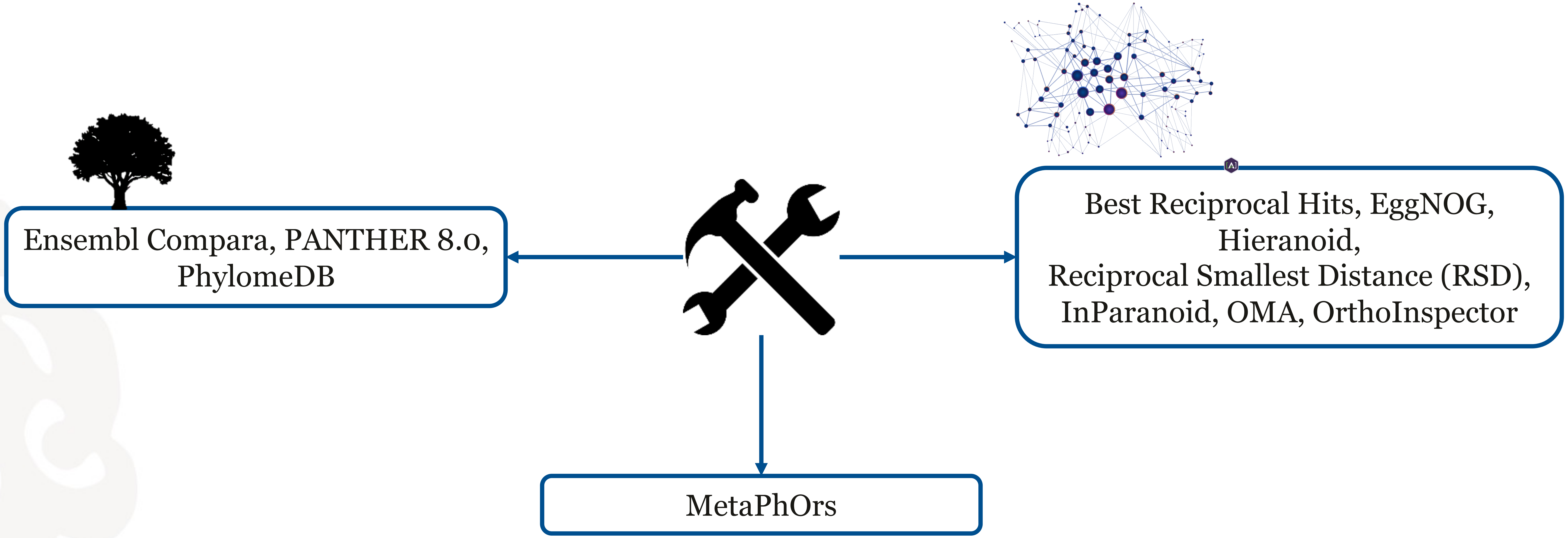
Gene function prediction

Phylogenetic tree inference

Genome annotation

Comparative genomics

The quest for orthologs



Ensembl Compara, PANTHER 8.0, PhylomeDB

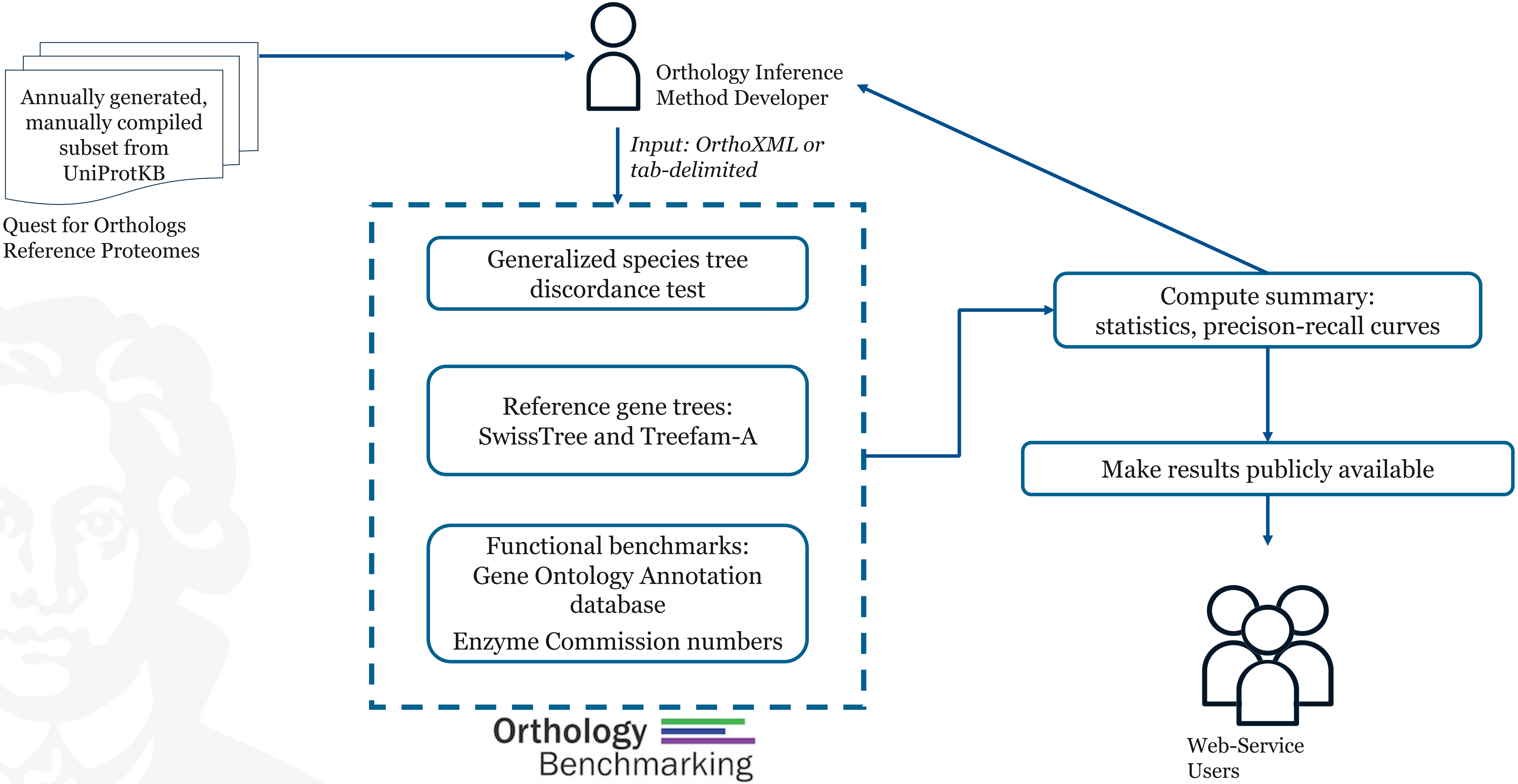
Best Reciprocal Hits, EggNOG, Hieranoid, Reciprocal Smallest Distance (RSD), InParanoid, OMA, OrthoInspector

MetaPhOrs

The quest for orthologs

How can we create a comprehensive assessment of state-of-the-art orthology inference tools?

Orthology Benchmarking



Annually generated, manually compiled subset from UniProtKB

Quest for Orthologs Reference Proteomes

Orthology Inference Method Developer

Input: OrthoXML or tab-delimited

Generalized species tree discordance test

Reference gene trees: SwissTree and Treefam-A

Functional benchmarks: Gene Ontology Annotation database

Enzyme Commission numbers

Orthology Benchmarking

Compute summary: statistics, precison-recall curves

Make results publicly available

Web-Service Users

Orthology Benchmarking (Illustrations based on Figure 1 in paper, logo taken from orthology.benchmarkservice.org)

Conclusion Benchmarking

Generalized species tree discordance test:

- highest precision, lowest recall = OMA; lowest precision, highest recall = PANTHER 8.0. Overall good performance of Orthoinspector, In Paranoid, PANTHER (LDO only). No obvious performance difference between tree-based and graph-based methods.

Reference gene trees:

- Predictions made with SwissTree and TreeFam-A give quite similar results
- Balanced precision-recall strategies performed best (especially MetaPhOrs)
- Skewed precision-recall strategy (OMA or PANTHER) might perform better on ambiguous phylogenies

Functional benchmarks:

- Consistent results for UniProt-GOA and ENZYME database (except for MetaPhOrs, whose missing taxa have a negative effect on recall)

➔ Trade-off between precision and recall

Conclusion

- Systematic comparison of multiple benchmarks
- Basis for meta-prediction methods
- Comparative test for new methods
- Over 1000 user submissions so far
- Open Challenges: Circularity, individual protein domains, orthogroups instead of paired orthologs, confidence scores

Conclusion (Altenhoff AM et al., The Quest for Orthologs benchmark service and consensus calls in 2020 Nucleic Acids Res, (2020), 48:W1, W538–W545 (doi: 10.1093/nar/gkaa308))

References

(1) Altenhoff AM et al., *The Quest for Orthologs benchmark service and consensus calls in 2020* Nucleic Acids Res, (2020), 48:W1, W538–W545 (doi: 10.1093/nar/gkaa308)

Poster