Supplementary Methods

In order for DIAMOND to be added to the tool as an alternative sequence alignment method, a new parser script was built, and the possibility to run DIAMOND from the tool was added to the code. In addition to this, the program has been made more user friendly by adding an extensive help-function, a possibility to handle input data in bulk i.e. running all versus all proteomes given in a directory, parallelization of the runs, and the possibility to take command line parameters as input. InParanoid is now available to run in a Docker, as well as a Singularity container including all necessary dependencies, to make it easier to install, and use on different systems.

To evaluate the speedup and ortholog quality, InParanoid-DIAMOND run with different sensitivity and composition based statistics options was compared to InParanoid-BLAST 4.2 (Remm *et al.*, 2001; Ostlund *et al.*, 2010). Furthermore, InParanoid-DIAMOND and InParanoid-BLAST were compared to SonicParanoid (Cosentino and Iwasaki, 2019), an orthology detection algorithm very similar to InParanoid, developed with focus on high speed. In addition to this, comparisons were made to Proteinortho (Lechner *et al.*, 2011) and OrthoFinder (Emms and Kelly, 2019), both orthology detection tools having the option to run DIAMOND for sequence search. For OrthoFInder, two variants of the tools, with and without the MSA option were used in the comparisons.

To evaluate the speed of the tools, the 78 Quest for Orthologs (Altenhoff *et al.*, 2020) 2018 reference proteomes were run against each other on an Intel Xeon CPU E5-2690 v3 @2.60GHz cluster node with 24 cores and 48 threads. The runtimes in hours and CPU hours were measured using the Ubuntu time command. The runtime of the tools was measured using versions 4.2 for InParanoid-BLAST, 1.3.5 for SonicParanoid, 6.0.31 for Proteinortho and 2.5.4 for OrthoFinder. For InParanoid-DIAMOND, all cores were used for each of the DIAMOND runs, which were executed in sequence. For InParanoid-BLAST, a wrapper bash script was created to handle running all vs all species, where the runs were distributed over the threads so that multiple instances of InParanoid were run simultaneously.

Evaluation of the quality of the orthologs was made using the Orthology benchmark service (Altenhoff *et al.*, 2020) for the QFO reference proteomes 2018, in OpenEBench (Capella-Gutierrez *et al.*) run on 2021-04-16. This is a web-based service to estimate the performance of ortholog predictions in eleven different tests, comparing the predictions to other orthology analysis tools or variants (currently 21). The tests measure the performance in terms of precision and recall, and a tool is considered to be a top performer if it appears on the Pareto frontier, hence having a tradeoff between precision and recall that is not outperformed by any other method (Altenhoff *et al.*, 2016). Comparisons of the performance in the benchmark was made with results publicly available in the benchmark server from the tools InParanoid (v. 4.1), SonicParanoid (default and sensitive v. 1.0.9 and most-sensitive v. 1.2.6), Proteinortho (v. 6.0.13) and OrthoFinder (v. 2.0). Version information was obtained from https://orthology.benchmarkservice.org/proxy/projects/2018/. Since public results for the QFO reference proteomes 2020 were not available for all methods in the comparison at the time of the analysis, the 2018 reference proteomes were used for the runtime comparisons as well as benchmark performance comparisons. InParanoid-DIAMOND was however benchmarked on the QFO 2020 data on the OpenEBench web server (on 2021-12-15), and the results of this, together with the compared tools having publically available results can be found in Supplementary Table 4. In order to be able to use publicly available benchmark results submitted by the authors of the methods for benchmark comparisons while still using the most recent version of the tools for runtime comparisons, different versions of the tools were used for timing and for benchmark performance.

Supplementary Results

Assessing the differences between the default version of InParanoid-DIAMOND and InParanoid-BLAST, we could see that the BLAST version of the tool in general identified more orthologs than the DIAMOND version, which agrees with the lower sensitivity of DIAMOND even when using the very-sensitive setting (Buchfink *et al.*, 2021). As an example, for the species pair *Homo sapiens* and *Echerichia coli*, InParanoid-DIAMOND and InParanoid-BLAST identify the same number of ortholog groups, 537. From these groups, 1220 orthologs pairs with an average score of 195 and an average sequence length of 460 were generated for InParanoid-BLAST, and 1218 pairs with an average score of 193 and an average sequence length of 446 for InParanoid-DIAMOND. Out of these pairs, 167 ortholog pairs were found exclusively by InParanoid-BLAST and 165 exclusively by InParanoid-DIAMOND. These pairs possess a significantly lower average score than the complete set of pairs, slightly above 100 for both tools, indicating that the orthologs detected are more prone to disagree for lower scoring hits. For these pairs, the average sequence length was higher than in the complete set for both tools, 661 for InParanoid-BLAST and 570 for InParanoid-DIAMOND. This could indicate that the tools tend to diverge more for proteins with longer sequences, and that InParanoid-BLAST detects a higher number of ortholog pairs including proteins with longer sequences. For ortholog pairs identified by InParanoid-BLAST where none of the proteins were present in the InParanoid-DIAMOND output, many were found to be low-scoring homologs not detected by DIAMOND at all, while some were excluded due to the matches being shorter in DIAMOND, or by the InParanoid algorithm prioritizing other, higher-scoring homologs to one of the proteins in the pair as stronger ortholog candidates.

In 21% of the species pairs among the Quest for orthologs proteomes, InParanoid-DIAMOND inferred more ortholog pairs than InParanoid-BLAST. An example of this is the species pair *Nematostella vectensis/Rattus norvegicus* where InParanoid-BLAST generated 11052 ortholog pairs from 5509 ortholog groups, and InParanoid-DIAMOND generated 16674 pairs from 5599 groups. Although the numbers of ortholog groups are relatively similar, the numbers of pairs generated differ more drastically. This can be explained by the sizes of the groups that on average are 4.8% larger in InParanoid-DIAMOND compared to InParanoid-BLAST, and the ortholog groups detected exclusively by InParanoid-DIAMOND consisted to a larger extent of more than one protein from each species.

For some species pairs InParanoid-DIAMOND and InParanoid-BLAST have very large differences in the number of ortholog pairs identified (see Supplementary figure 2). One example of this is the species pair *Giardia intestinalis* and *Monodelphis domestica* generating 52564 ortholog pairs by InParanoid-BLAST and 2046 ortholog pairs by InParanoid-DIAMOND. This very large number of ortholog pairs for InParanoid-BLAST can to a large extent be traced back to one very large ortholog group, which alone generated 49959 of the pairs. The seed orthologs responsible for this large ortholog group could not be detected by InParanoid-DIAMOND, and therefore the pairs for that group could not be identified. We could identify several species pairs displaying a similar behavior, where InParanoid-BLAST detected very large ortholog groups with several hundreds of in-paralogs where an equally large group was not detected by InParanoid-DIAMOND, either because the seed ortholog pair was not found as a homolog, or because the corresponding ortholog groups were significantly smaller.

Running InParanoid-DIAMOND on the 273 proteomes used for the InParanoid 8 database (Sonnhammer and Östlund, 2015) took 7 hours on an AMD EPYC 7742 128-core cluster node with 256 threads and 256 GB RAM, using 4 threads for each DIAMOND run. The run used a total of 1311 CPU hours, showing that InParanoid-DIAMOND achieves a good level of parallelization and scales well to large sets of proteomes on machines with a high number of threads.

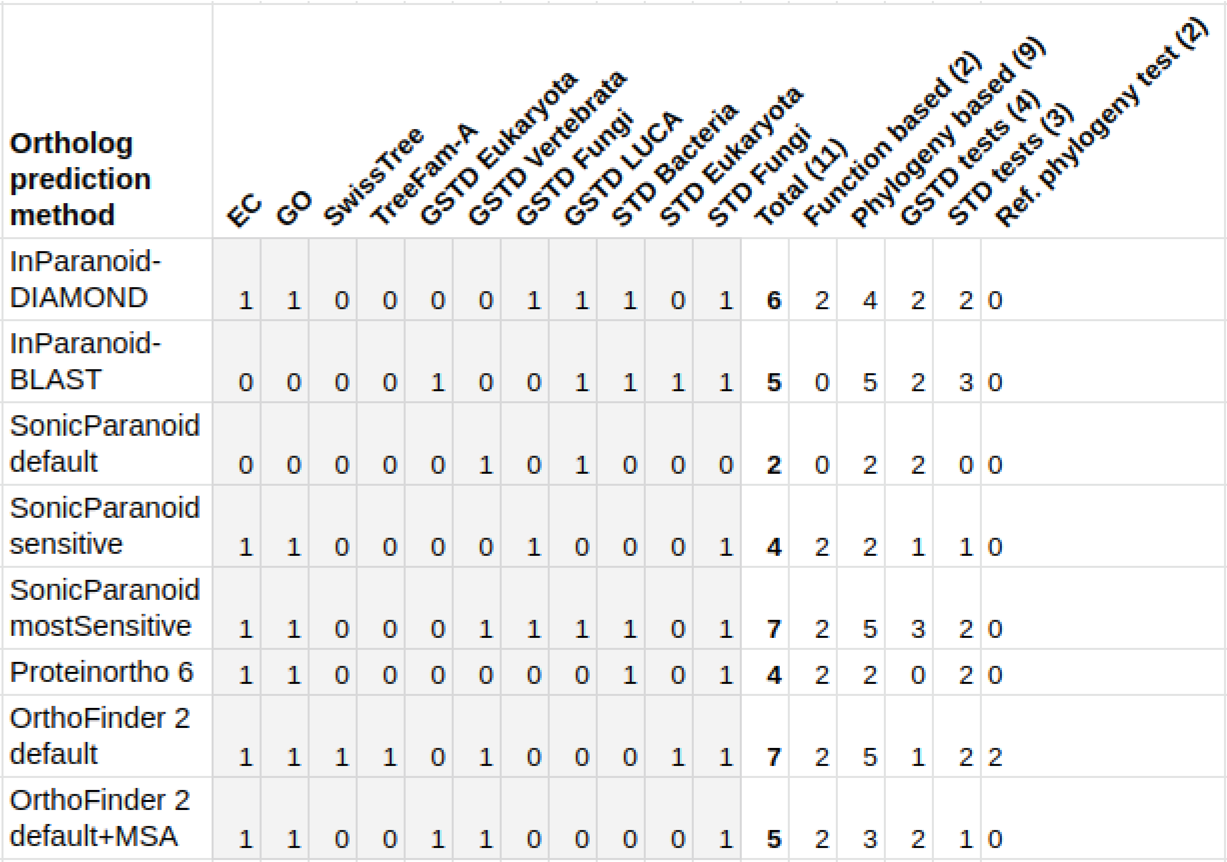
Supplementary Table 1. Total number of orthologs predicted for the Quest for orthologs (2018) reference proteomes and the runtime for the dataset in hours on a 48 thread node as well as in CPU hours, for InParanoid-BLAST and various sensitivity settings for InParanoid-DIAMOND .

Ortholog prediction method	Number of ortholog pairs	Average Jaccard	Average Unique InParanoid-BLAST	Average Unique InParanoid-DIAMOND	Runtime (hours)	Runtime (CPU hours)
InParanoid-BLAST	12,576,950	-	-	-	166.50	5940.12
InParanoid-DIAMOND sensitive	9,594,028	0.698	22.05%	8.17%	11.42	90.37
InParanoid-DIAMOND more sensitive	9,661,846	0.703	21.41%	8.30%	11.98	108.38
InParanoid-DIAMOND very sensitive	9,946,816	0.743	17.18%	8.50%	10.48	135.82
InParanoid-DIAMOND very sensitive with composition based statistics (option 4)	9,640,672	0.683	22.22%	9.48%	16.13	424.42
InParanoid-DIAMOND ultra sensitive	10,065,934	0.751	16.00%	8.86%	19.92	386.32

Supplementary Table 2. Runtimes in CPU minutes for Homo sapiens/Escherichia coli and number of orthologs predicted for InParanoid-BLAST and InParanoid-DIAMOND (InParaDiam) run with 1-pass and 2-pass for the same species.

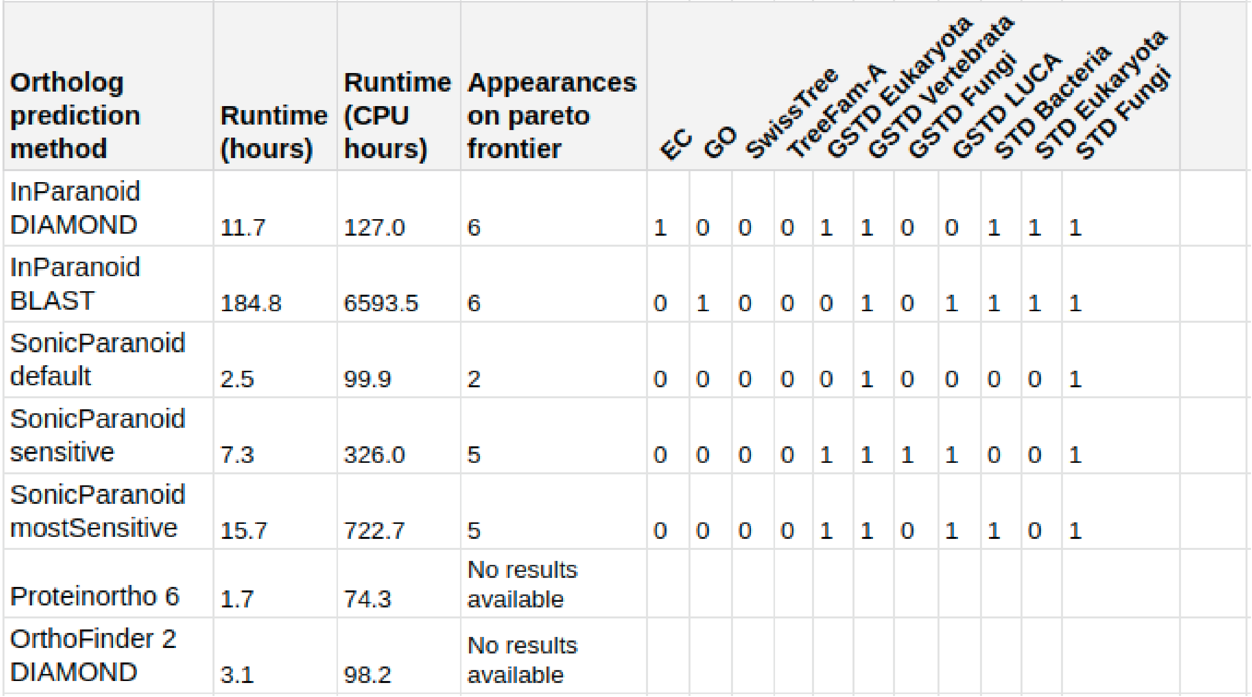
Ortholog prediction method	Runtime (CPU minutes)	Number of ortholog pairs
InParanoid-BLAST	238	1220
InParanoid-DIAMOND 2-pass	84	1016
InParanoid-DIAMOND 1-pass	9	1218

Supplementary Table 3. Summary of the appearances on the Pareto frontier in the 11 benchmark tests of the orthology benchmark service in OpenEBench for the QFO18 reference proteomes. Appearance on the Pareto frontier for each test, Enzyme Classification conservation (EC), Gene Ontology conservation (GO), Agreement with reference gene phylogenies: SwissTree (SwissTree), Agreement with reference gene phylogenies: TreeFam-A (TreeFam-A), Generalized Species tree discordance test (GSTD) for Eukaryota, Vertebrata, Fungi and LUCA, and Species tree discordance test (STD) for Bacteria, Eukaryota and Fungi, is represented by 1, and not on the Pareto frontier by 0. Total represents the total number of appearances on the Pareto frontier out of the 11 tests. Function based represents the number of appearances on the Pareto frontier for the function based tests, EC and GO. Phylogeny based represents the number of appearances on the Pareto frontier in the 9 remaining phylogeny-based tests. The GSTD tests represent the number of appearances on the Pareto frontier in the 4 GSTD tests, STD tests represent the number of appearances on the Pareto frontier in the 3 STD tests, and Ref. phylogeny test represents the number of appearances on the Pareto frontier in the 2 reference phylogeny tests.

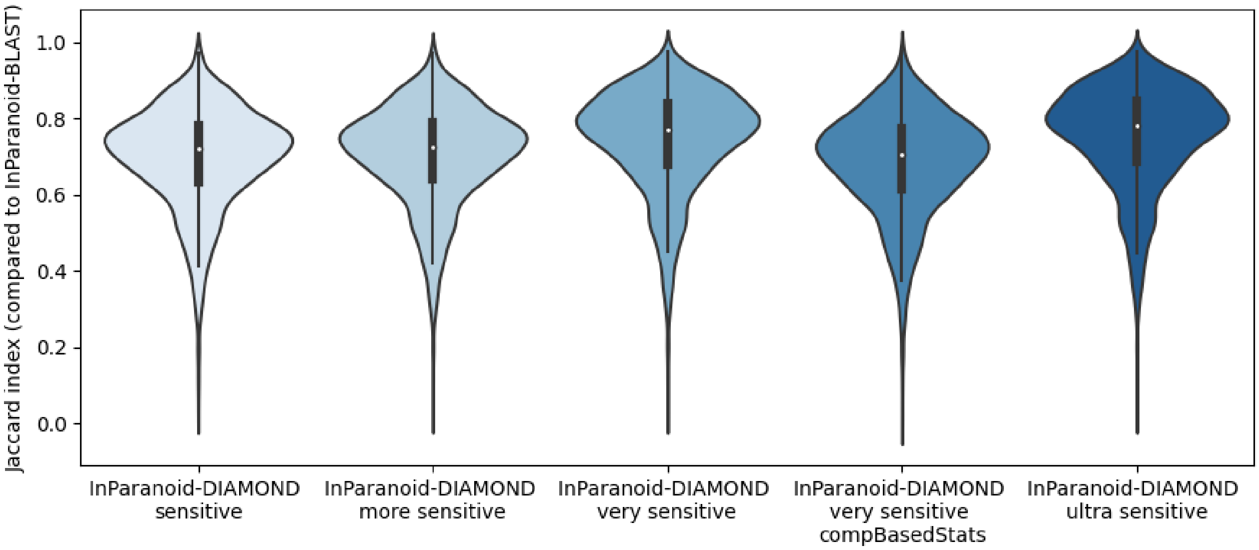


Ortholog prediction method	**EC**	**GO**	**SwissTree**	**TreeFam-A**	**GSTD Eukaryota**	**GSTD Vertebrata**	**GSTD Fungi**	**GSTD LUCA**	**STD Bacteria**	**STD Eukaryota**	**STD Fungi**	**Total (11)**	**Function based (2)**	**Phylogeny based (9)**	**GSTD tests (4)**	**STD tests (3)**	**Ref. phylogeny test (2)**
InParanoid-DIAMOND	1	1	0	0	0	0	1	1	1	0	1	**6**	2	4	2	2	0
InParanoid-BLAST	0	0	0	0	1	0	0	1	1	1	1	**5**	0	5	2	3	0
SonicParanoid default	0	0	0	0	0	1	0	1	0	0	0	**2**	0	2	2	0	0
SonicParanoid sensitive	1	1	0	0	0	0	1	0	0	0	1	**4**	2	2	1	1	0
SonicParanoid mostSensitive	1	1	0	0	0	1	1	1	1	0	1	**7**	2	5	3	2	0
Proteinortho 6	1	1	0	0	0	0	0	0	1	0	1	**4**	2	2	0	2	0
OrthoFinder 2 default	1	1	1	1	0	1	0	0	0	1	1	**7**	2	5	1	2	2
OrthoFinder 2 default+MSA	1	1	0	0	1	1	0	0	0	0	1	**5**	2	3	2	1	0

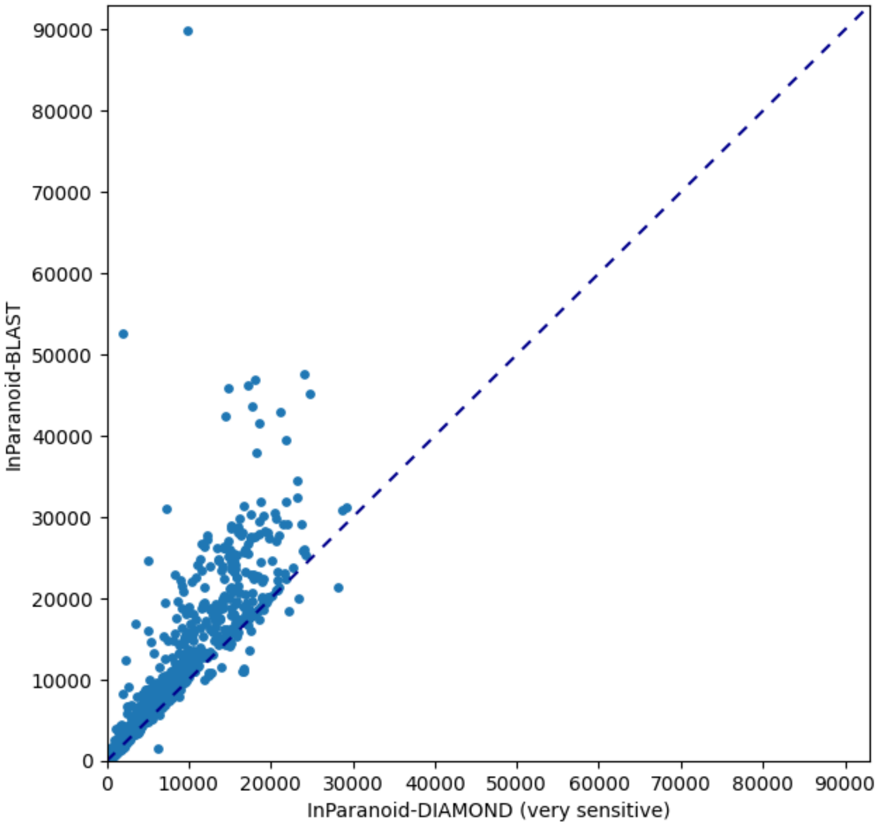
Supplementary Table 4. Summary of the appearances on the Pareto frontier in the 11 benchmark tests of the orthology benchmark service in OpenEBench for the QFO20 reference proteomes. Appearance on the Pareto frontier for each test, Enzyme Classification conservation (EC), Gene Ontology conservation (GO), Agreement with reference gene phylogenies: SwissTree (SwissTree), Agreement with reference gene phylogenies: TreeFam-A (TreeFam-A), Generalized Species tree discordance test (GSTD) for Eukaryota, Vertebrata, Fungi and LUCA, and Species tree discordance test (STD) for Bacteria, Eukaryota and Fungi, is represented by 1, and not on the Pareto frontier by 0. The results from the benchmark display publicly available results using versions 4.1 for InParanoid-BLAST, 1.0.9 for SonicParanoid default and sensitive, and 1.2.6 for SonicParanoid most-sensitive. Version information was obtained from https://orthology.benchmarkservice.org/proxy/projects/2020/. Runtimes displayed are measured when running version 4.1 of InParanoid-BLAST, version 1.3.5 of SonicParanoid, version 6.0.31 of Proteinortho and version 2.5.4 of OrthoFinder on the QFO 2020 reference proteomes.



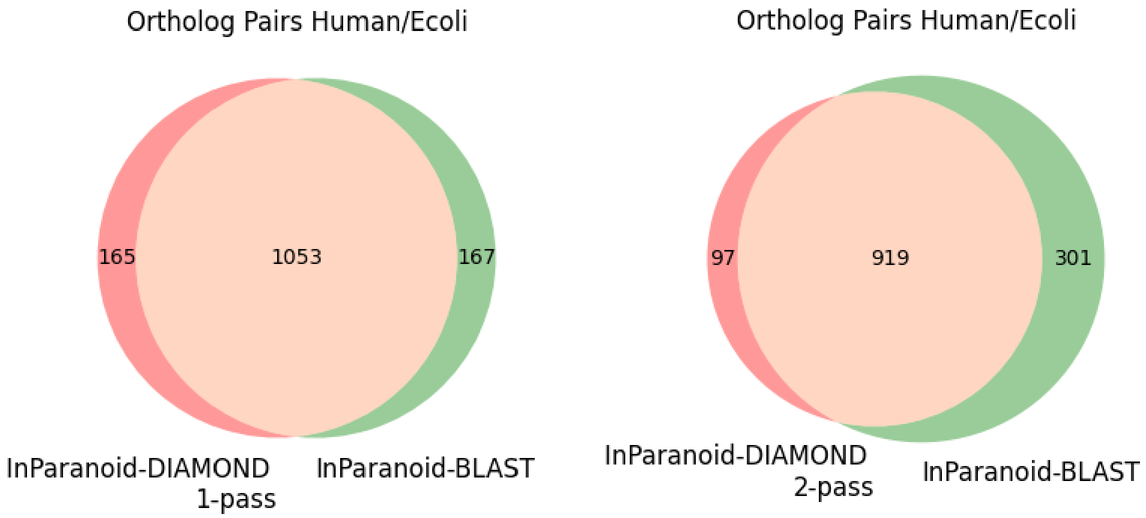
Ortholog prediction method	**Runtime (hours)**	**Runtime (CPU hours)**	**Appearances on pareto frontier**	**EC**	**GO**	**SwissTree**	**TreeFam-A**	**GSTD Eukaryota**	**GSTD Vertebrata**	**GSTD Fungi**	**GSTD LUCA**	**STD Bacteria**	**STD Eukaryota**	**STD Fungi**
InParanoid DIAMOND	11.7	127.0	6	1	0	0	0	1	1	0	0	1	1	1
InParanoid BLAST	184.8	6593.5	6	0	1	0	0	0	1	0	1	1	1	1
SonicParanoid default	2.5	99.9	2	0	0	0	0	0	1	0	0	0	0	1
SonicParanoid sensitive	7.3	326.0	5	0	0	0	0	1	1	1	1	0	0	1
SonicParanoid mostSensitive	15.7	722.7	5	0	0	0	0	1	1	0	1	1	0	1
Proteinortho 6	1.7	74.3	No results available											
OrthoFinder 2 DIAMOND	3.1	98.2	No results available											



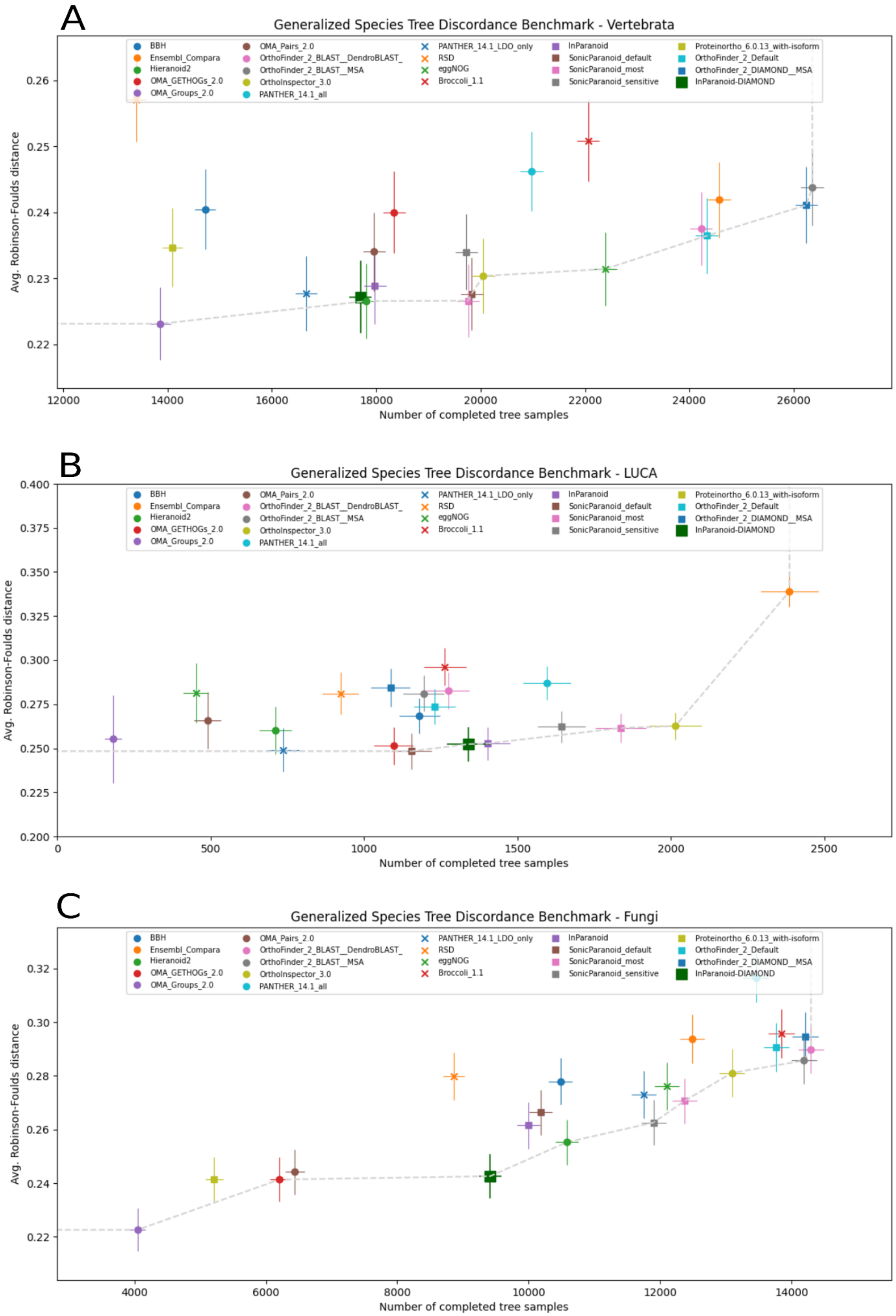
Supplementary Figure 1. Violin plot showing the distribution of jaccard indices between InParanoid-DIAMOND, and InParanoid-BLAST over all pairs of species in the Quest for orthologs (2018) reference proteomes, for sensitivity settings: sensitive, more sensitive, very sensitive, very sensitive with composition-based statistics (option 4), and ultra sensitive

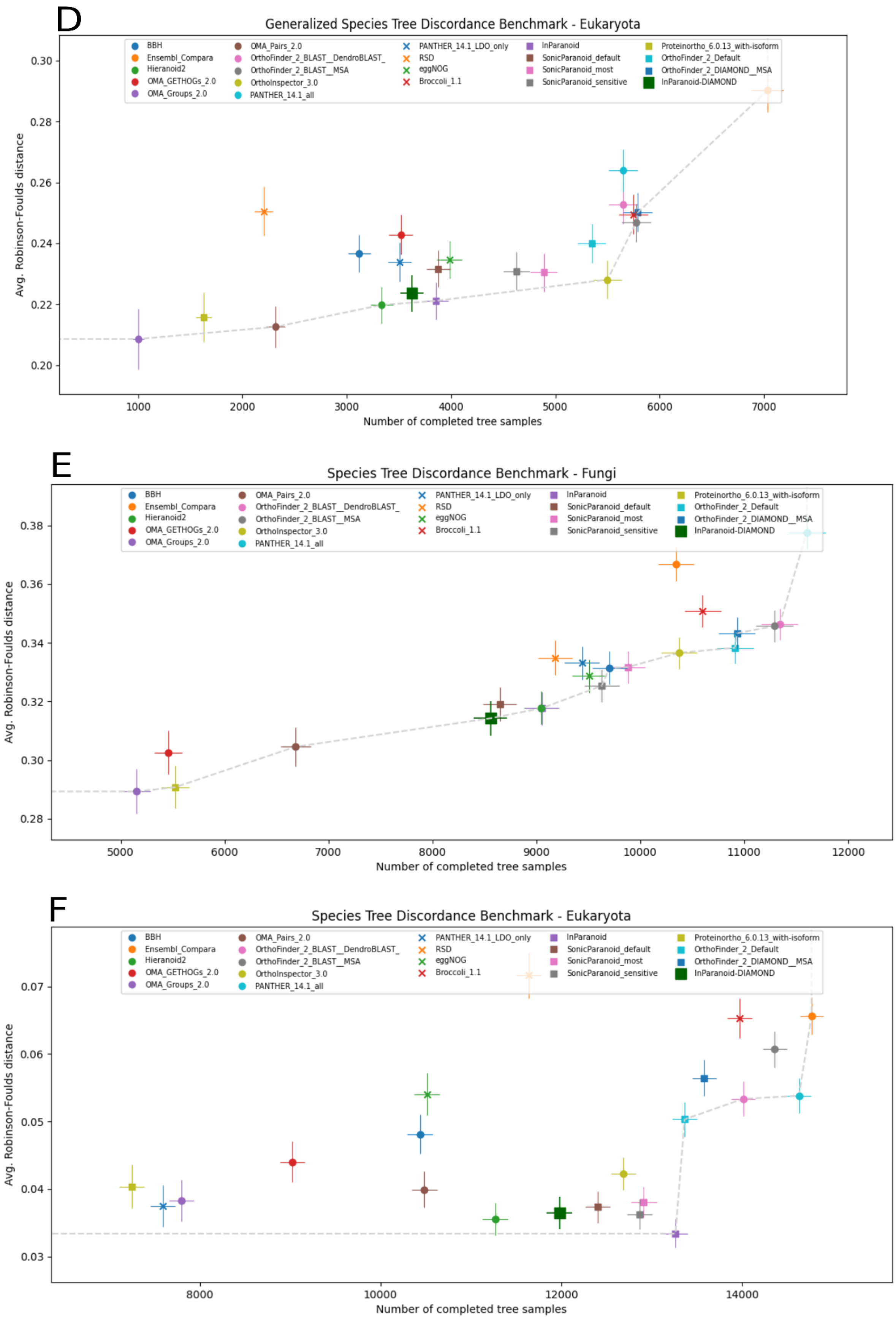


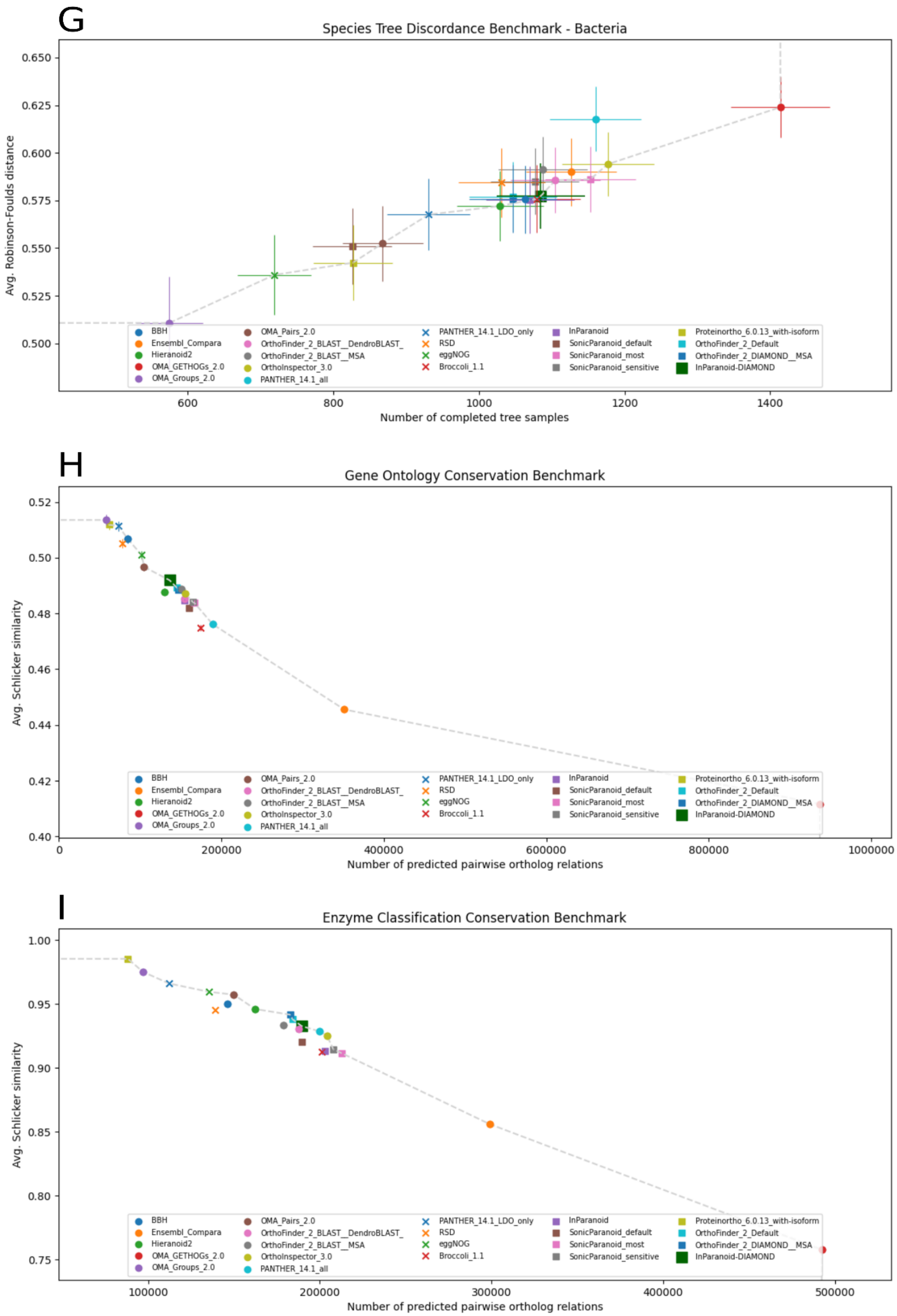
Supplementary Figure 2. Number or ortholog pairs detected for each species pair in the quest for orthologs reference proteomes for InParanoid-BLAST vs InParanoid-DIAMOND. The dotted line represents the diagonal, where the number of pairs is the same for both tools**.**

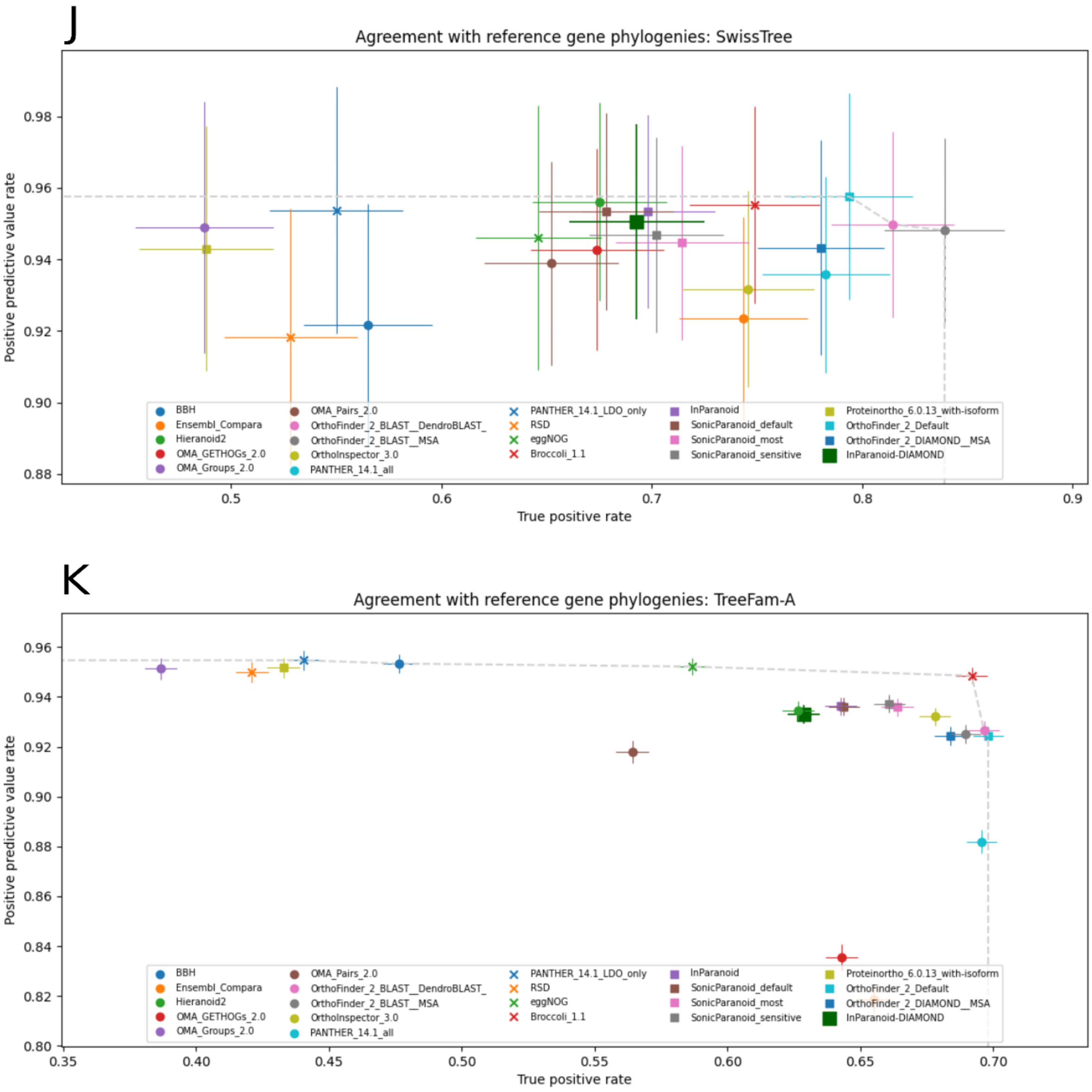


Supplementary Figure 3. Overlap of ortholog pairs detected by InParanoid-DIAMOND using 1-pass (left) and 2-pass (right) with InParanoid-BLAST pairs for *Homo sapiens* versus *Escherichia coli*.


A
Generalized Species Tree Discordance Benchmark - Vertebrata
BBH
Ensembl_Compara
Hieranoid2
OMA_GETHOGs_2.0
OMA_Groups_2.0
OMA_Pairs_2.0
OrthoFinder_2_BLAST__DendroBLAST_
OrthoFinder_2_BLAST__MSA
OrthoInspector_3.0
PANTHER_14.1_all
PANTHER_14.1_LDO_only
RSD
eggNOG
Broccoli_1.1
InParanoid
SonicParanoid_default
SonicParanoid_most
SonicParanoid_sensitive
Proteinortho_6.0.13_with-isoform
OrthoFinder_2_Default
OrthoFinder_2_DIAMOND__MSA
InParanoid-DIAMOND
Avg. Robinson-Foulds distance
0.22
0.23
0.24
0.25
0.26
Number of completed tree samples
12000
14000
16000
18000
20000
22000
24000
26000
B
Generalized Species Tree Discordance Benchmark - LUCA
Avg. Robinson-Foulds distance
0.200
0.225
0.250
0.275
0.300
0.325
0.350
0.375
0.400
Number of completed tree samples
0
500
1000
1500
2000
2500
C
Generalized Species Tree Discordance Benchmark - Fungi
Avg. Robinson-Foulds distance
0.22
0.24
0.26
0.28
0.30
0.32
Number of completed tree samples
4000
6000
8000
10000
12000
14000


D
Generalized Species Tree Discordance Benchmark - Eukaryota
BBH
Ensembl_Compara
Hieranoid2
OMA_GETHOGs_2.0
OMA_Groups_2.0
OMA_Pairs_2.0
OrthoFinder_2_BLAST__DendroBLAST_
OrthoFinder_2_BLAST__MSA
OrthoInspector_3.0
PANTHER_14.1_all
PANTHER_14.1_LDO_only
RSD
eggNOG
Broccoli_1.1
InParanoid
SonicParanoid_default
SonicParanoid_most
SonicParanoid_sensitive
Proteinortho_6.0.13_with-isoform
OrthoFinder_2_Default
OrthoFinder_2_DIAMOND__MSA
InParanoid-DIAMOND
Avg. Robinson-Foulds distance
0.30
0.28
0.26
0.24
0.22
0.20
1000
2000
3000
4000
5000
6000
7000
Number of completed tree samples
E
Species Tree Discordance Benchmark - Fungi
BBH
Ensembl_Compara
Hieranoid2
OMA_GETHOGs_2.0
OMA_Groups_2.0
OMA_Pairs_2.0
OrthoFinder_2_BLAST__DendroBLAST_
OrthoFinder_2_BLAST__MSA
OrthoInspector_3.0
PANTHER_14.1_all
PANTHER_14.1_LDO_only
RSD
eggNOG
Broccoli_1.1
InParanoid
SonicParanoid_default
SonicParanoid_most
SonicParanoid_sensitive
Proteinortho_6.0.13_with-isoform
OrthoFinder_2_Default
OrthoFinder_2_DIAMOND__MSA
InParanoid-DIAMOND
Avg. Robinson-Foulds distance
0.38
0.36
0.34
0.32
0.30
0.28
5000
6000
7000
8000
9000
10000
11000
12000
Number of completed tree samples
F
Species Tree Discordance Benchmark - Eukaryota
BBH
Ensembl_Compara
Hieranoid2
OMA_GETHOGs_2.0
OMA_Groups_2.0
OMA_Pairs_2.0
OrthoFinder_2_BLAST__DendroBLAST_
OrthoFinder_2_BLAST__MSA
OrthoInspector_3.0
PANTHER_14.1_all
PANTHER_14.1_LDO_only
RSD
eggNOG
Broccoli_1.1
InParanoid
SonicParanoid_default
SonicParanoid_most
SonicParanoid_sensitive
Proteinortho_6.0.13_with-isoform
OrthoFinder_2_Default
OrthoFinder_2_DIAMOND__MSA
InParanoid-DIAMOND
Avg. Robinson-Foulds distance
0.07
0.06
0.05
0.04
0.03
8000
10000
12000
14000
Number of completed tree samples


G
Species Tree Discordance Benchmark - Bacteria
Avg. Robinson-Foulds distance
0.650
0.625
0.600
0.575
0.550
0.525
0.500
600
800
1000
1200
1400
Number of completed tree samples
BBH
Ensembl_Compara
Hieranoid2
OMA_GETHOGs_2.0
OMA_Groups_2.0
OMA_Pairs_2.0
OrthoFinder_2_BLAST__DendroBLAST_
OrthoFinder_2_BLAST__MSA
OrthoInspector_3.0
PANTHER_14.1_all
PANTHER_14.1_LDO_only
RSD
eggNOG
Broccoli_1.1
InParanoid
SonicParanoid_default
SonicParanoid_most
SonicParanoid_sensitive
Proteinortho_6.0.13_with-isoform
OrthoFinder_2_Default
OrthoFinder_2_DIAMOND__MSA
InParanoid-DIAMOND
H
Gene Ontology Conservation Benchmark
Avg. Schlicker similarity
0.52
0.50
0.48
0.46
0.44
0.42
0.40
0
200000
400000
600000
800000
1000000
Number of predicted pairwise ortholog relations
BBH
Ensembl_Compara
Hieranoid2
OMA_GETHOGs_2.0
OMA_Groups_2.0
OMA_Pairs_2.0
OrthoFinder_2_BLAST__DendroBLAST_
OrthoFinder_2_BLAST__MSA
OrthoInspector_3.0
PANTHER_14.1_all
PANTHER_14.1_LDO_only
RSD
eggNOG
Broccoli_1.1
InParanoid
SonicParanoid_default
SonicParanoid_most
SonicParanoid_sensitive
Proteinortho_6.0.13_with-isoform
OrthoFinder_2_Default
OrthoFinder_2_DIAMOND__MSA
InParanoid-DIAMOND
I
Enzyme Classification Conservation Benchmark
Avg. Schlicker similarity
1.00
0.95
0.90
0.85
0.80
0.75
100000
200000
300000
400000
500000
Number of predicted pairwise ortholog relations
BBH
Ensembl_Compara
Hieranoid2
OMA_GETHOGs_2.0
OMA_Groups_2.0
OMA_Pairs_2.0
OrthoFinder_2_BLAST__DendroBLAST_
OrthoFinder_2_BLAST__MSA
OrthoInspector_3.0
PANTHER_14.1_all
PANTHER_14.1_LDO_only
RSD
eggNOG
Broccoli_1.1
InParanoid
SonicParanoid_default
SonicParanoid_most
SonicParanoid_sensitive
Proteinortho_6.0.13_with-isoform
OrthoFinder_2_Default
OrthoFinder_2_DIAMOND__MSA
InParanoid-DIAMOND



Supplementary Figure 4. Results from the orthology benchmark service for the QFO18 reference proteomes.

References

Altenhoff,A.M. *et al.* (2016) Standardized benchmarking in the quest for orthologs. *Nat. Methods*, **13**, 425–430.

Altenhoff,A.M. *et al.* (2020) The Quest for Orthologs benchmark service and consensus calls in 2020. *Nucleic Acids Res.*, **48**, W538–W545.

Buchfink,B. *et al.* (2021) Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods*, **18**, 366–368.

Capella-Gutierrez,S. *et al.* Lessons Learned: Recommendations for Establishing Critical

Periodic Scientific Benchmarking.

Cosentino,S. and Iwasaki,W. (2019) SonicParanoid: fast, accurate and easy orthology inference. *Bioinformatics*, **35**, 149–151.

Emms,D.M. and Kelly,S. (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.*, **20**, 238.

Lechner,M. *et al.* (2011) Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinformatics*, **12**, 124.

Ostlund,G. *et al.* (2010) InParanoid 7: new algorithms and tools for eukaryotic orthology analysis. *Nucleic Acids Res.*, **38**, D196–203.

Remm,M. *et al.* (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J. Mol. Biol.*, **314**, 1041–1052.

Sonnhammer,E.L.L. and Östlund,G. (2015) InParanoid 8: orthology analysis between 273 proteomes, mostly eukaryotic. *Nucleic Acids Res.*, **43**, D234–9.