Supplementary Materials 1

All databases used in testing the program can be download from following sources.

CAZy Database - Carbohydrate Active Enzymes database, http://www.cazy.org/
Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy) : an expert resource for Glycogenomics. **Nucleic Acids Res** 37:D233-238 [PMID : 18838391].

https://bcb.unl.edu/dbCAN2/download/CAZyDB.07142024.fa

TreeGenes - https://treegenesdb.org/

Wegrzyn J.L., Staton M.A., Street N. R., Main D., Grau E., Herndon N., Buehler S., Falk T., Zaman S., Ramnath R., Richter P., Sun L., Condon B., Almsaeed A., Chen M.,Mannapperuma C., Jung S., Ficklin S. Cyberinfrastructure to Improve Forest Health and Productivity: The Role of Tree Databases in Connecting Genomes, Phenomes, and the Environment, TreeGenes. Database, Volume 2019. doi:10.3389/fpls.2019.00813

Falk T., Herndon N., Grau E., Buehler S., Richter P., Zaman S., Baker E.M., Ramnath R., Ficklin S., Staton M., Feltus F.A., Jung S., Main D., Wegrzyn J.L. (2018). **Growing and cultivating the forest genomics database, TreeGenes.** Database, Volume 2018 doi:10.1093/database/bay084

Abies alba
https://treegenesdb.org/FTP/Genomes/Abal/v1.1/annotation/Abal.1_1.pep.fa.gz

Picea abies
https://treegenesdb.org/FTP/Genomes/Paab/v1.0b/annotation/Paab.1_0b.pep.fa.gz

Populus tremula x tremuloides

https://treegenesdb.org/FTP/Genomes/Potl/v1.0/annotation/Potl.1_0.pep.fa.gz

Ginkgo Biloba

https://treegenesdb.org/FTP/Genomes/Gibi/v1.0/annotation/Gibi.1_0.pep.fa.gz

OneKP - https://db.cngb.org/onekp/

J.H. Leebens-Mack, M.S. Barker, E.J. Carpenter, M.K. Deyholos, M.A. Gitzendanner, S.W. Graham, I. Grosse, Z. Li, M. Melkonian, S. Mirarab, ... G.K. Wong. Oct 2019. One thousand plant transcriptomes and the phylogenomics of green plants. Nature 574: 679-685. PMID: 31645766. DOI: 10.1038/s41586-019-1693-2
E.J. Carpenter, N. Matasci, S. Ayyampalayam, S. Wu, J. Sun, J. Yu, F.R. Jimenez Vieira, C. Bowler, R.G. Dorrell, M.A. Gitzendanner, ... G.K. Wong. Oct 2019. Access to RNA-sequencing data from 1173 plant species: The 1000 plant transcriptomes initiative (1KP). GigaScience 8: giz126. PMID: 31644802. DOI: 10.1093/gigascience/giz126

https://ftp.cngb.org/pub/SciRAID/onekp/assemblies/DZQM-Pinus_radiata/DZQM-translated-protein.fa.gz

TAIR database - The Arabidopsis Information Resource, www.arabidopsis.org

Leonore Reiser, Erica Bakker, Sabarinath Subramaniam, Xingguo Chen, Swapnil Sawant, Kartik Khosa, Trilok Prithvi, Tanya Z Berardini, The Arabidopsis Information Resource in 2024, *Genetics*, Volume 227, Issue 1, May 2024, iyae027, https://doi.org/10.1093/genetics/iyae027

https://www.arabidopsis.org/download/file?path=Proteins/Araport11_protein_lists/Araport11_pep_20250411.gz

Solgenomics - https://solgenomics.sgn.cornell.edu/

Noe Fernandez-Pozo, Naama Menda, Jeremy D. Edwards, Surya Saha, Isaak Y. Tecle, Susan R. Strickler, Aureliano Bombarely, Thomas Fisher-York, Anuradha Pujar, Hartmut Foerster, Aimin Yan, Lukas A. Mueller, The Sol Genomics Network (SGN)—from genotype to phenotype to breeding, *Nucleic Acids Research*, Volume 43, Issue D1, 28 January 2015, Pages D1036–D1041, https://doi.org/10.1093/nar/gku1195

https://solgenomics.net/ftp/genomes/Nicotiana_benthamiana/annotation/Niben101/Niben101_annotation.proteins.fasta.gz