Bibliography

"UniProt: the universal protein knowledgebase in 2021", *Nucleic acids research*, vol. 49, no. D1, D480–D489 2021

Abbas, U. L., Chen, J., and Shao, Q., "Assessing fairness of AlphaFold2 prediction of protein 3D structures", in "Proceedings of the 14th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics", pp. 1–10 2023

Abdel-Ghany, S. E., Ullah, F., Ben-Hur, A., and Reddy, A. S., "Transcriptome analysis of drought-resistant and drought-sensitive sorghum (Sorghum bicolor) genotypes in response to PEG-induced drought stress", *International Journal of Molecular Sciences*, vol. 21, no. 3, 772 2020

Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K., "Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling", *The Plant Cell*, vol. 15, no. 1, 63–78 2003

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J. *et al.*, "Accurate structure prediction of biomolecular interactions with AlphaFold 3", *Nature*, pp. 1–3 2024

Agarwal, G., Garg, V., Kudapa, H., Doddamani, D., Pazhamala, L. T., Khan, A. W., Thudi, M., Lee, S.-H., and Varshney, R. K., "Genome-wide dissection of AP2/ERF and HSP90 gene families in five legumes and expression profiles in chickpea and pigeonpea", *Plant biotechnology journal*, vol. 14, no. 7, 1563–1577 2016

Ahmad, M., "Genomics and transcriptomics to protect rice (Oryza sativa. L.) from abiotic stressors:-pathways to achieving zero hunger", *Frontiers in Plant Science*, vol. 13, 1002596 2022

Ahmad, Q. R., Nguyen, D. H., Wingerd, M. A., Church, G. M., and Steffen, M. A., "Molecular weight assessment of proteins in total proteome profiles using 1D-PAGE and LC/MS/MS", *Proteome science*, vol. 3, 1–7 2005

Akhtar, S., Rao, E., Uike, A., and Saatu, M., "Plant breeding strategies: traditional and modern approaches", *Genetic revolution in agriculture: unleashing the power of plant genetics. Elite Publishing House, New Delhi* 2023

Alam, S., Sundaram, K. T., Singh, U. M., Srinivas Prasad, M., Laha, G. S., Sinha, P., and Singh, V. K., "Superior haplotypes towards the development of blast and bacterial blight-resistant rice", *Frontiers in Plant Science*, vol. 15, 1272326 2024

Allen, M. D., Yamasaki, K., Ohme-Takagi, M., Tateno, M., and Suzuki, M., "A novel mode of DNA recognition by a β-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA", *The EMBO journal* 1998

Almagro Armenteros, J. J., Sønderby, C. K., Sønderby, S. K., Nielsen, H., and Winther, O., "DeepLoc: prediction of protein subcellular localization using deep learning", *Bioinformatics*, vol. 33, no. 21, 3387–3395 2017

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J., "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic acids research*, vol. 25, no. 17, 3389–3402 1997

Anand, A., Subramanian, M., and Kar, D., "Breeding techniques to dispense higher genetic gains", *Frontiers in Plant Science*, vol. 13, 1076094 2023

Andrews, S. *et al.*, "FastQC: a quality control tool for high throughput sequence data", 2010

Aziz, M. A. and Masmoudi, K., "Molecular Breakthroughs in Modern Plant Breeding Techniques", *Horticultural Plant Journal* 2024

Babu, R. C., Zhang, J., Blum, A., Ho, T.-H. D., Wu, R., and Nguyen, H., "HVA1, a LEA gene from barley

confers dehydration tolerance in transgenic rice (Oryza sativa L.) via cell membrane protection", *Plant Science*, vol. 166, no. 4, 855–862 2004

Baibout, M., Corcket, E., Kothari, S. L., and Fievet, V., "Ecosystem services provided by Prosopis cineraria (L.) Druce in the drylands of Southern and Western Asia", *Botany Letters*, vol. 169, no. 1, 30–42 2022

Bakade, R., Ingole, K. D., Deshpande, S., Pal, G., Patil, S. S., Bhattacharjee, S., Prasannakumar, M., and Ramu, V. S., "Comparative transcriptome analysis of rice resistant and susceptible genotypes to Xanthomonas oryzae pv. oryzae identifies novel genes to control bacterial leaf blight", *Molecular Biotechnology*, vol. 63, no. 8, 719–731 2021

Bargmann, B. O. and Munnik, T., "The role of phospholipase D in plant stress responses", *Current opinion in plant biology*, vol. 9, no. 5, 515–522 2006

Bari, B. S., Islam, M. N., Rashid, M., Hasan, M. J., Razman, M. A. M., Musa, R. M., Nasir, A. F. A., and Majeed, A. P., "A real-time approach of diagnosing rice leaf disease using deep learning-based faster R-CNN framework", *PeerJ Computer Science*, vol. 7 2021

Bashford, D. and Case, D. A., "Generalized born models of macromolecular solvation effects", *Annual review of physical chemistry*, vol. 51, no. 1, 129–152 2000

Bej, A., Sahoo, B. R., Swain, B., Basu, M., Jayasankar, P., and Samanta, M., "LRRsearch: an asynchronous server-based application for the prediction of leucine-rich repeat motifs and an integrative database of NOD-like receptors", *Computers in biology and medicine*, vol. 53, 164–170 2014

Beltran-Peña, A., Rosa, L., and D'Odorico, P., "Global food self-sufficiency in the 21st century under sustainable intensification of agriculture", *Environmental Research Letters*, vol. 15, no. 9, 095004 2020

Bertioli, D. J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D., Seijo, G., Leal-Bertioli, S. C., Ren, L., Farmer, A. D., Pandey, M. K. *et al.*, "The genome sequence of segmental allotetraploid peanut Arachis hypogaea", *Nature genetics*, vol. 51, no. 5, 877–884 2019

Bigman, L. S., Iwahara, J., and Levy, Y., "Negatively charged disordered regions are prevalent and functionally important across proteomes", *Journal of molecular biology*, vol. 434, no. 14, 167660 2022

Bihani, P., Char, B., and Bhargava, S., "Transgenic expression of sorghum DREB2 in rice improves tolerance and yield under water limitation", *The Journal of Agricultural Science*, vol. 149, no. 1, 95–101 2011

Blum, M., Chang, H.-Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., Nuka, G., Paysan-Lafosse, T., Qureshi, M., Raj, S. *et al.*, "The InterPro protein families and domains database: 20 years on", *Nucleic acids research*, vol. 49, no. D1, D344–D354 2021

Breseghello, F. and Coelho, A. S. G., "Traditional and modern plant breeding methods with examples in rice (Oryza sativa L.)", *Journal of agricultural and food chemistry*, vol. 61, no. 35, 8277–8286 2013

Calle Garcia, J., Guadagno, A., Paytuvi-Gallart, A., Saera-Vila, A., Amoroso, C. G., D'Esposito, D., Andolfo, G., Aiese Cigliano, R., Sanseverino, W., and Ercolano, M. R., "PRGdb 4.0: an updated database dedicated to genes involved in plant disease resistance process", *Nucleic Acids Research*, vol. 50, no. D1, D1483–D1490 2022

Chang, C.-C. and Lin, C.-J., "LIBSVM: a library for support vector machines", *ACM transactions on intelligent systems and technology (TIST)*, vol. 2, no. 3, 1–27 2011

Charfeddine, M., Saïdi, M. N., Charfeddine, S., Hammami, A., and Gargouri Bouzid, R., "Genome-wide analysis and expression profiling of the ERF transcription factor family in potato (Solanum tuberosum L.)", *Molecular biotechnology*, vol. 57, 348–358 2015

Chauhan, D., Walia, R., Singh, C., Deivakani, M., Kumbhkar, M., and Professor, A., "Detection of Maize Disease Using Random Forest Classification Algorithm", *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12 2021

Che, Y., Zhang, C., Xing, J., Xi, Q., Shao, Y., Zhao, L., Guo, S., and Zuo, Y., "Machine Learning-Based identification of resistance genes associated with sunflower broomrape", *Plant Methods*, vol. 21, no. 1, 1–15 2025

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., and Xia, R., "TBtools: an integrative toolkit developed for interactive analyses of big biological data", *Molecular plant*, vol. 13, no. 8, 1194–1202 2020

Chen, H., Hu, L., Wang, L., Wang, S., and Cheng, X., "Genome-wide identification and expression profiles

of AP2/ERF transcription factor family in mung bean (Vigna radiata L.)", *Journal of Applied Genetics*, vol. 63, no. 2, 223–236 2022

Cheng, W.-H., Endo, A., Zhou, L., Penney, J., Chen, H.-C., Arroyo, A., Leon, P., Nambara, E., Asami, T., Seo, M. *et al.*, "A unique short-chain dehydrogenase/reductase in Arabidopsis glucose signaling and abscisic acid biosynthesis and functions", *The Plant Cell*, vol. 14, no. 11, 2723–2743 2002

Choi, S. C., Lee, S., Kim, S.-R., Lee, Y.-S., Liu, C., Cao, X., and An, G., "Trithorax group protein Oryza sativa Trithorax1 controls flowering time in rice via interaction with early heading date3", *Plant physiology*, vol. 164, no. 3, 1326–1337 2014

Chukwu, S., Rafii, M., Ramlee, S., Ismail, S., Hasan, M., Oladosu, Y., Magaji, U., Akos, I., and Olalekan, K., "Bacterial leaf blight resistance in rice: a review of conventional breeding to molecular approach", *Molecular biology reports*, vol. 46, 1519–1532 2019

Close, T. J., Kortt, A. A., and Chandler, P. M., "A cDNA-based comparison of dehydration-induced proteins (dehydrins) in barley and corn", *Plant molecular biology*, vol. 13, 95–108 1989

Colovos, C. and Yeates, T. O., "Verification of protein structures: patterns of nonbonded atomic interactions", *Protein science*, vol. 2, no. 9, 1511–1519 1993

Cui, M., Haider, M. S., Chai, P., Guo, J., Du, P., Li, H., Dong, W., Huang, B., Zheng, Z., Shi, L. *et al.*, "Genome-wide identification and expression analysis of AP2/ERF transcription factor related to drought stress in cultivated peanut (Arachis hypogaea L.)", *Frontiers in Genetics*, vol. 12, 750761 2021

Das, S., Pattanayak, S., and Behera, P. R., "Application of machine learning: a recent advancement in plant diseases detection", *Journal of Plant Protection Research*, vol. 62, 122–135 2022, URL `https://www.plantprotection.pl/Application-of-machine-learning-a-recent-advancement-in-plant-diseases-detection,148094,0,2.html`

De Castro, E., Sigrist, C. J., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P. S., Gasteiger, E., Bairoch, A., and Hulo, N., "ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins", *Nucleic acids research*, vol. 34, no. suppl_2, W362–W365 2006

Dey, S., Pal, A., Guharoy, M., Sonavane, S., and Chakrabarti, P., "Characterization and prediction of the binding site in DNA-binding proteins: improvement of accuracy by combining residue composition, evolutionary conservation and structural parameters", *Nucleic acids research*, vol. 40, no. 15, 7150–7161 2012

Dhiman, V., Biswas, S., Shekhawat, R. S., Sadhukhan, A., and Yadav, P., "In silico characterization of five novel disease-resistance proteins in Oryza sativa sp. japonica against bacterial leaf blight and rice blast diseases", *3 Biotech*, vol. 14, no. 2, 48 2024

Diaz-Granados, A., Petrescu, A.-J., Goverse, A., and Smant, G., "SPRYSEC effectors: a versatile protein-binding platform to disrupt plant innate immunity", *Frontiers in plant science*, vol. 7, 1575 2016

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R., "STAR: ultrafast universal RNA-seq aligner", *Bioinformatics*, vol. 29, no. 1, 15–21 2013

Dossa, K., Wei, X., Li, D., Fonceka, D., Zhang, Y., Wang, L., Yu, J., Boshou, L., Diouf, D., Cissé, N. *et al.*, "Insight into the AP2/ERF transcription factor superfamily in sesame and expression profiling of DREB subfamily under drought stress", *BMC plant biology*, vol. 16, 1–16 2016

Eisenberg, D., Lüthy, R., and Bowie, J. U., "[20] VERIFY3D: assessment of protein models with three-dimensional profiles", in "Methods in enzymology", vol. 277, pp. 396–404, Elsevier 1997

Enany, S., "Structural and functional analysis of hypothetical and conserved proteins of Clostridium tetani", *Journal of infection and public health*, vol. 7, no. 4, 296–307 2014

FABIANPEDREGOSA, F. P., Michel, V., OLIVIERGRISEL, O. G., Blondel, M., Prettenhofer, P., Weiss, R., Vanderplas, J., Cournapeau, D., Pedregosa, F., Varoquaux, G., Gramfort, A., Thirion, B., Grisel, O., Dubourg, V., Passos, A., Brucher, M., andÉdouardand, M. P., andÉdouard Duchesnay, and EDOUARD-DUCHESNAY, F. D., "Scikit-learn: Machine Learning in Python Gaël Varoquaux Bertrand Thirion Vincent Dubourg Alexandre Passos PEDREGOSA, VAROQUAUX, GRAMFORT ET AL. Matthieu Perrot", *Journal of Machine Learning Research*, vol. 12, 2825–2830 2011, URL `http://scikit-learn.`

sourceforge.net.

Fan, W., Hai, M., Guo, Y., Ding, Z., Tie, W., Ding, X., Yan, Y., Wei, Y., Liu, Y., Wu, C. *et al.*, "The ERF transcription factor family in cassava: genome-wide characterization and expression analyses against drought stress", *Scientific reports*, vol. 6, no. 1, 37379 2016

Faraji, S., Filiz, E., Kazemitabar, S. K., Vannozzi, A., Palumbo, F., Barcaccia, G., and Heidari, P., "The AP2/ERF gene family in Triticum durum: genome-wide identification and expression analysis under drought and salinity stresses", *Genes*, vol. 11, no. 12, 1464 2020

Fariselli, P., Riccobelli, P., and Casadio, R., "Role of evolutionary information in predicting the disulfide-bonding state of cysteine in proteins", *Proteins: Structure, Function, and Bioinformatics*, vol. 36, no. 3, 340–346 1999

Freire-Rios, A., Tanaka, K., Crespo, I., Van der Wijk, E., Sizentsova, Y., Levitsky, V., Lindhoud, S., Fontana, M., Hohlbein, J., Boer, D. R. *et al.*, "Architecture of DNA elements mediating ARF transcription factor binding and auxin-responsive gene expression in Arabidopsis", *Proceedings of the National Academy of Sciences*, vol. 117, no. 39, 24557–24566 2020

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W., "CD-HIT: accelerated for clustering the next-generation sequencing data", *Bioinformatics*, vol. 28, no. 23, 3150–3152 2012

Fujimoto, S. Y., Ohta, M., Usui, A., Shinshi, H., and Ohme-Takagi, M., "Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box–mediated gene expression", *The Plant Cell*, vol. 12, no. 3, 393–404 2000

Garg, A. K., Kim, J.-K., Owens, T. G., Ranwala, A. P., Choi, Y. D., Kochian, L. V., and Wu, R. J., "Trehalose accumulation in rice plants confers high tolerance levels to different abiotic stresses", *Proceedings of the National Academy of Sciences*, vol. 99, no. 25, 15898–15903 2002

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A., "ExPASy: the proteomics server for in-depth protein knowledge and analysis", *Nucleic acids research*, vol. 31, no. 13, 3784–3788 2003

Geng, L., Ren, J., Ji, X., Yan, S., and Song, X. S., "Over-expression of DREB46 enhances drought tolerance in Populus trichocarpa", *Journal of Plant Physiology*, vol. 281, 153923 2023

Gigon, A., Matos, A.-R., Laffray, D., Zuily-Fodil, Y., and Pham-Thi, A.-T., "Effect of drought stress on lipid metabolism in the leaves of Arabidopsis thaliana (ecotype Columbia)", *Annals of botany*, vol. 94, no. 3, 345–351 2004

Gowda, M., Shirke, M. D., Mahesh, H., Chandarana, P., Rajamani, A., and Chattoo, B. B., "Genome analysis of rice-blast fungus Magnaporthe oryzae field isolates from southern India", *Genomics data*, vol. 5, 284–291 2015

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q. *et al.*, "Full-length transcriptome assembly from RNA-Seq data without a reference genome", *Nature biotechnology*, vol. 29, no. 7, 644–652 2011

Gu, X., Gao, S., Li, J., Song, P., Zhang, Q., Guo, J., Wang, X., Han, X., Wang, X., Zhu, Y. *et al.*, "The bHLH transcription factor regulated gene OsWIH2 is a positive regulator of drought tolerance in rice", *Plant Physiology and Biochemistry*, vol. 169, 269–279 2021

Guo, B., Wei, Y., Xu, R., Lin, S., Luan, H., Lv, C., Zhang, X., Song, X., and Xu, R., "Genome-wide analysis of APETALA2/ethylene-responsive factor (AP2/ERF) gene family in barley (Hordeum vulgare L.)", *PLoS One*, vol. 11, no. 9, e0161322 2016

Guo, Y., Qu, Y., Fan, R., Sun, F., Chen, Q., Shi, J., Zheng, K., Zhang, Y., Chen, Q., Wang, N. *et al.*, "Regulating drought tolerance in cotton by the expression of a specific allele of heat shock protein 70", *Industrial Crops and Products*, vol. 202, 116820 2023

Hand, S. C., Menze, M. A., Toner, M., Boswell, L., and Moore, D., "LEA proteins during water stress: not just for plants anymore", *Annual review of physiology*, vol. 73, no. 1, 115–134 2011

Hao, D., Ohme-Takagi, M., and Sarai, A., "Unique mode of GCC box recognition by the DNA-binding domain of ethylene-responsive element-binding factor (ERF domain) in plant", *Journal of Biological Chemistry*, vol. 273, no. 41, 26857–26861 1998

Harini, K., Kihara, D., and Gromiha, M. M., "PDA-Pred: Predicting the binding affinity of protein-DNA

complexes using machine learning techniques and structural features", *Methods*, vol. 213, 10–17 2023

Hassan, M. A., Dahu, N., Hongning, T., Qian, Z., Yueming, Y., Yiru, L., and Shimei, W., "Drought stress in rice: morpho-physiological and molecular responses and marker-assisted breeding", *Frontiers in Plant Science*, vol. 14, 1215371 2023

Hassan, S., Berk, K., and Aronsson, H., "Evolution and identification of DREB transcription factors in the wheat genome: modeling, docking and simulation of DREB proteins associated with salt stress", *Journal of Biomolecular Structure and Dynamics*, vol. 40, no. 16, 7191–7204 2022

He, R., Zhuang, Y., Cai, Y., Agüero, C. B., Liu, S., Wu, J., Deng, S., Walker, M. A., Lu, J., and Zhang, Y., "Overexpression of 9-cis-epoxycarotenoid dioxygenase cisgene in grapevine increases drought tolerance and results in pleiotropic effects", *Frontiers in plant science*, vol. 9, 970 2018

He, W., Luo, L., Xie, R., Chai, J., Wang, H., Wang, Y., Chen, Q., Wu, Z., Yang, S., Li, M. *et al.*, "Genome-wide identification and functional analysis of the AP2/ERF transcription factor family in citrus rootstock under waterlogging stress", *International Journal of Molecular Sciences*, vol. 24, no. 10, 8989 2023

He, Z., Xin, Y., Wang, C., Yang, H., Xu, Z., Cheng, J., Li, Z., Ye, C., Yin, H., Xie, Z. *et al.*, "Genomics-Assisted improvement of super high-yield hybrid rice variety "super 1000" for resistance to bacterial blight and blast diseases", *Frontiers in Plant Science*, vol. 13, 881244 2022

Hong-Bo, S., Zong-Suo, L., and Ming-An, S., "LEA proteins in higher plants: structure, function, gene expression and regulation", *Colloids and surfaces B: Biointerfaces*, vol. 45, no. 3-4, 131–135 2005

Hsieh, E.-J., Cheng, M.-C., and Lin, T.-P., "Functional characterization of an abiotic stress-inducible transcription factor AtERF53 in Arabidopsis thaliana", *Plant molecular biology*, vol. 82, 223–237 2013

Hsu, Y.-C., Chiu, C.-H., Yap, R., Tseng, Y.-C., and Wu, Y.-P., "Pyramiding bacterial blight resistance genes in Tainung82 for broad-spectrum resistance using marker-assisted selection", *International journal of molecular sciences*, vol. 21, no. 4, 1281 2020

Huang, X., Song, X., Chen, R., Zhang, B., Li, C., Liang, Y., Qiu, L., Fan, Y., Zhou, Z., Zhou, H. *et al.*, "Genome-wide analysis of the DREB subfamily in Saccharum spontaneum reveals their functional divergence during cold and drought stresses", *Frontiers in Genetics*, vol. 10, 1326 2020

Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., De Castro, E., Langendijk-Genevaux, P. S., Pagni, M., and Sigrist, C. J., "The PROSITE database", *Nucleic acids research*, vol. 34, no. suppl_1, D227–D230 2006

Hundertmark, M. and Hincha, D. K., "LEA (late embryogenesis abundant) proteins and their encoding genes in Arabidopsis thaliana", *BMC genomics*, vol. 9, 1–22 2008

Iqbal, Z., Iqbal, M. S., Hashem, A., Abd_Allah, E. F., and Ansari, M. I., "Plant defense responses to biotic stress and its interplay with fluctuating dark/light conditions", *Frontiers in Plant Science*, vol. 12, 631810 2021

Iuchi, S., Kobayashi, M., Taji, T., Naramoto, M., Seki, M., Kato, T., Tabata, S., Kakubari, Y., Yamaguchi-Shinozaki, K., and Shinozaki, K., "Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in Arabidopsis", *The Plant Journal*, vol. 27, no. 4, 325–333 2001

Jamaloddin, M., Mahender, A., Gokulan, C. G., Balachiranjeevi, C., Maliha, A., Patel, H. K., and Ali, J., "Molecular approaches for disease resistance in rice", in "Rice improvement: Physiological, molecular breeding and genetic perspectives", pp. 315–378, Springer International Publishing Cham 2021

Jarambasa, T., Regon, P., Jyoti, S. Y., Gupta, D., Panda, S. K., and Tanti, B., "Genome-wide identification and expression analysis of the Pisum sativum (L.) APETALA2/ethylene-responsive factor (AP2/ERF) gene family reveals functions in drought and cold stresses", *Genetica*, vol. 151, no. 3, 225–239 2023

Jayaraman, V., Toledo-Patiño, S., Noda-García, L., and Laurino, P., "Mechanisms of protein evolution", *Protein Science*, vol. 31, no. 7, e4362 2022

Jiang, F., Lu, Y., Chen, Y., Cai, D., and Li, G., "Image recognition of four rice leaf diseases based on deep learning and support vector machine", *Computers and Electronics in Agriculture*, vol. 179 2020

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., and Madden, T. L., "NCBI BLAST: a better web interface", *Nucleic Acids Research*, vol. 36, W5–W9 7 2008, URL `https://dx.doi.org/10.1093/nar/gkn201`

Jordan, M. I., Kearns, M. J., and Solla, S. A., *Advances in neural information processing systems 10: pro-*

ceedings of the 1997 conference, vol. 10, Mit Press 1998

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A. *et al.*, "Highly accurate protein structure prediction with AlphaFold", *nature*, vol. 596, no. 7873, 583–589 2021

Kalladan, R., Lasky, J. R., Sharma, S., Kumar, M. N., Juenger, T. E., Des Marais, D. L., and Verslues, P. E., "Natural variation in 9-cis-epoxycartenoid dioxygenase 3 and ABA accumulation", *Plant Physiology*, vol. 179, no. 4, 1620–1631 2019

Kanehisa, M., "The KEGG database", in "'In silico'simulation of biological processes: Novartis Foundation Symposium 247", vol. 247, pp. 91–103, Wiley Online Library 2002

Kartikeyan, P. and Shrivastava, G., "Review on Emerging Trends in Detection of Plant Diseases using Image Processing with Machine Learning", *International Journal of Computer Applications*, vol. 174 2021

Kaur, A., Vishwakarma, H., Maibam, A., and Padaria, J. C., "Cloning, characterization and in silico studies on abiotic stress responsive Hsp17. 9 from Prosopis cineraria", *Indian journal of plant physiology*, vol. 23, 731–740 2018

Keel, B. N., Lindholm-Perry, A. K., and Snelling, W. M., "Evolutionary and functional features of copy number variation in the cattle genome", *Frontiers in genetics*, vol. 7, 207 2016

Khalili, E., Kouchaki, S., Ramazi, S., and Ghanati, F., "Machine Learning Techniques for Soybean Charcoal Rot Disease Prediction", *Frontiers in Plant Science*, vol. 11 2020

Kim, S. L., Kim, N., Lee, H., Lee, E., Cheon, K.-S., Kim, M., Baek, J., Choi, I., Ji, H., Yoon, I. S. *et al.*, "High-throughput phenotyping platform for analyzing drought tolerance in rice", *Planta*, vol. 252, no. 3, 38 2020

Kirola, M., Joshi, K., Chaudhary, S., Singh, N., Anandaram, H., and Gupta, A., "Plants Diseases Prediction Framework: A Image-Based System Using Deep Learning", in "Proceedings - 2022 IEEE World Conference on Applied Intelligence and Computing, AIC 2022", 2022

Klasberg, S., Bitard-Feildel, T., and Mallet, L., "Computational identification of novel genes: current and future perspectives", *Bioinformatics and Biology insights*, vol. 10, BBI–S39950 2016

Kolberg, L., Raudvere, U., Kuzmin, I., Adler, P., Vilo, J., and Peterson, H., "g: Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update)", *Nucleic acids research*, vol. 51, no. W1, W207–W212 2023

Kumar, A., Kumar, R., Shukla, P., Singh, S., Alam, M., and Kumar Singh, D., "Update on cloning and molecular characterization of bacterial blight resistance genes in rice", in "Molecular Response and Genetic Engineering for Stress in Plants, Volume 2: Biotic stress", pp. 7–1, IOP Publishing Bristol, UK 2022

Kumar, I. S., Zaharin, N., and Nadarajah, K., "In silico Identification of Resistance and Defense Related Genes for Bacterial Leaf Blight (BLB) in Rice.", *Journal of Pure & Applied Microbiology*, vol. 12, no. 4 2018

Kumar, S., Meshram, S., and Sinha, A., "BACTERIAL DISEASES IN RICE AND THEIR ECO-FRIENDLY MANAGEMENT", 2017, URL `www.tjprc.org`

Kyte, J. and Doolittle, R. F., "A simple method for displaying the hydropathic character of a protein", *Journal of molecular biology*, vol. 157, no. 1, 105–132 1982

Lakhwani, D., Pandey, A., Dhar, Y. V., Bag, S. K., Trivedi, P. K., and Asif, M. H., "Genome-wide analysis of the AP2/ERF family in Musa species reveals divergence and neofunctionalisation during evolution", *Scientific reports*, vol. 6, no. 1, 18878 2016

Lamichhane, S. and Thapa, S., "Advances from conventional to modern plant breeding methodologies", *Plant breeding and biotechnology*, vol. 10, no. 1, 1–14 2022

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M., "PROCHECK: a program to check the stereochemical quality of protein structures", *Journal of applied crystallography*, vol. 26, no. 2, 283–291 1993

Lata, C. and Prasad, M., "Role of DREBs in regulation of abiotic stress responses in plants", *Journal of experimental botany*, vol. 62, no. 14, 4731–4748 2011

Lata, C., Bhutty, S., Bahadur, R. P., Majee, M., and Prasad, M., "Association of an SNP in a novel DREB2-

like gene SiDREB2 with stress tolerance in foxtail millet [Setaria italica (L.)]", *Journal of Experimental Botany*, vol. 62, no. 10, 3387–3401 2011

Le, T. T., Williams, B., and Mundree, S. G., "An osmotin from the resurrection plant Tripogon loliiformis (TlOsm) confers tolerance to multiple abiotic stresses in transgenic rice", *Physiologia plantarum*, vol. 162, no. 1, 13–34 2018

Lee, G. R., Won, J., Heo, L., and Seok, C., "GalaxyRefine2: simultaneous refinement of inaccurate local regions and overall protein structure", *Nucleic acids research*, vol. 47, no. W1, W451–W455 2019

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouzé, P., and Rombauts, S., "PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences", *Nucleic acids research*, vol. 30, no. 1, 325–327 2002

Letunic, I. and Bork, P., "Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation", *Nucleic acids research*, vol. 49, no. W1, W293–W296 2021

Letunic, I., Khedkar, S., and Bork, P., "SMART: recent updates, new developments and status in 2020", *Nucleic acids research*, vol. 49, no. D1, D458–D460 2021

Li, D., Wang, R., Xie, C., Liu, L., Zhang, J., Li, R., Wang, F., Zhou, M., and Liu, W., "A recognition method for rice plant diseases and pests video detection based on deep convolutional neural network", *Sensors*, vol. 20, no. 3, 578 2020

Li, J., Meng, L., Ren, S., Jia, C., Liu, R., Jiang, H., and Chen, J., "OsGSTU17, a tau class glutathione S-transferase gene, positively regulates drought stress tolerance in Oryza sativa", *Plants*, vol. 12, no. 17, 3166 2023

Li, W., Zhu, Z., Chern, M., Yin, J., Yang, C., Ran, L., Cheng, M., He, M., Wang, K., Wang, J. *et al.*, "A natural allele of a transcription factor in rice confers broad-spectrum blast resistance", *Cell*, vol. 170, no. 1, 114–126 2017

Li, X., Yang, Y., Yang, S., Sun, X., Yin, X., Zhao, Y., and Yang, Y., "Comparative proteomics analyses of intraspecific differences in the response of Stipa purpurea to drought", *Plant diversity*, vol. 38, no. 2, 101–117 2016

Li, X., Gao, B., Zhang, D., Liang, Y., Liu, X., Zhao, J., Zhang, J., and Wood, A. J., "Identification, classification, and functional analysis of AP2/ERF family genes in the desert moss Bryum argenteum", *International Journal of Molecular Sciences*, vol. 19, no. 11, 3637 2018

Li, Z., Wang, G., Liu, X., Wang, Z., Zhang, M., and Zhang, J., "Genome-wide identification and expression profiling of DREB genes in Saccharum spontaneum", *BMC genomics*, vol. 22, no. 1, 456 2021

Lin, F., Fan, J., and Rhee, S. Y., "QTG-Finder: a machine-learning based algorithm to prioritize causal genes of quantitative trait loci in Arabidopsis and rice", *G3: Genes, Genomes, Genetics*, vol. 9, no. 10, 3129–3138 2019a

Lin, F., Lazarus, E. Z., and Rhee, S. Y., "QTG-Finder2: a generalized machine-learning algorithm for prioritizing QTL causal genes in plants", *G3: Genes, Genomes, Genetics*, vol. 10, no. 7, 2411–2421 2020

Lin, Q., Yang, J., Wang, Q., Zhu, H., Chen, Z., Dao, Y., and Wang, K., "Overexpression of the trehalose-6-phosphate phosphatase family gene AtTPPF improves the drought tolerance of Arabidopsis thaliana", *BMC Plant Biology*, vol. 19, 1–15 2019b

Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K., and Shinozaki, K., "Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought-and low-temperature-responsive gene expression, respectively, in Arabidopsis", *The Plant Cell*, vol. 10, no. 8, 1391–1406 1998

Liu, Q., Zuo, S.-m., Peng, S., Zhang, H., Peng, Y., Li, W., Xiong, Y., Lin, R., Feng, Z., Li, H. *et al.*, "Development of Machine Learning Methods for Accurate Prediction of Plant Disease Resistance", *Engineering* 2024a

Liu, Q., Zuo, S.-m., Peng, S., Zhang, H., Peng, Y., Li, W., Xiong, Y., Lin, R., Feng, Z., Li, H. *et al.*, "Development of machine learning methods for accurate prediction of plant disease resistance", *Engineering*, vol. 40, 100–110 2024b

Liu, S., Wang, H., and Qin, F., "Genetic dissection of drought resistance for trait improvement in crops", *The Crop Journal*, vol. 11, no. 4, 975–985 2023

Liu, W. and Wang, G.-L., "Plant innate immunity in rice: a defense against pathogen infection", *National Science Review*, vol. 3, no. 3, 295–308 2016

Liu, Z., Zhu, Y., Shi, H., Qiu, J., Ding, X., and Kou, Y., "Recent progress in rice broad-spectrum disease resistance", *International journal of molecular sciences*, vol. 22, no. 21, 11658 2021

Liu, Z., Ren, Z., Yan, L., and Li, F., "DeepLRR: an online webserver for leucine-Rich-Repeat containing protein characterization based on deep learning", *Plants*, vol. 11, no. 1, 136 2022

Love, M. I., Huber, W., and Anders, S., "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2", *Genome biology*, vol. 15, 1–21 2014

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Marchler, G. H., Song, J. S. *et al.*, "CDD/SPARCLE: the conserved domain database in 2020", *Nucleic acids research*, vol. 48, no. D1, D265–D268 2020

Lu, Y., Zhong, Q., Xiao, S., Wang, B., Ke, X., Zhang, Y., Yin, F., Zhang, D., Jiang, C., Liu, L., Li, J., Yu, T., Wang, L., Cheng, Z., and Chen, L., "A new NLR disease resistance gene Xa47 confers durable and broad-spectrum resistance to bacterial blight in rice", *Frontiers in Plant Science*, vol. 13, 1037901 11 2022

Magdeldin, S., Yoshida, Y., Li, H., Maeda, Y., Yokoyama, M., Enany, S., Zhang, Y., Xu, B., Fujinaka, H., Yaoita, E. *et al.*, "Murine colon proteome and characterization of the protein pathways", *BioData mining*, vol. 5, 1–14 2012

Mahmood, U., Li, X., Fan, Y., Chang, W., Niu, Y., Li, J., Qu, C., and Lu, K., "Multi-omics revolution to promote plant breeding efficiency", *Frontiers in Plant Science*, vol. 13, 1062952 2022

Manavalan, R., "Automatic identification of diseases in grains crops through computational approaches: A review", 2020

Mansueto, L., Fuentes, R. R., Borja, F. N., Detras, J., Abrio-Santos, J. M., Chebotarov, D., Sanciangco, M., Palis, K., Copetti, D., Poliakov, A., Dubchak, I., Solovyev, V., Wing, R. A., Hamilton, R. S., Mauleon, R., McNally, K. L., and Alexandrov, N., "Rice SNP-seek database update: new SNPs, indels, and queries", *Nucleic Acids Research*, vol. 45, D1075 1 2017, URL `/pmc/articles/PMC5210592//pmc/articles/PMC5210592/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5210592/`

Martin, E. C., Sukarta, O. C., Spiridon, L., Grigore, L. G., Constantinescu, V., Tacutu, R., Goverse, A., and Petrescu, A.-J., "LRRpredictor—a new LRR motif detection method for irregular motifs of plant NLR proteins using an ensemble of classifiers", *Genes*, vol. 11, no. 3, 286 2020

McDonald, I. K. and Thornton, J. M., "Satisfying hydrogen bonding potential in proteins", *Journal of molecular biology*, vol. 238, no. 5, 777–793 1994

McGuffin, L. J., Bryson, K., and Jones, D. T., "The PSIPRED protein structure prediction server", *Bioinformatics*, vol. 16, no. 4, 404–405 2000

Mi, J., Yang, D., Chen, Y., Jiang, J., Mou, H., Huang, J., Ouyang, Y., and Mou, T., "Accelerated molecular breeding of a novel P/TGMS line with broad-spectrum resistance to rice blast and bacterial blight in two-line hybrid rice", *Rice*, vol. 11, 1–12 2018

Miller III, B. R., McGee Jr, T. D., Swails, J. M., Homeyer, N., Gohlke, H., and Roitberg, A. E., "MMPBSA. py: an efficient program for end-state free energy calculations", *Journal of chemical theory and computation*, vol. 8, no. 9, 3314–3321 2012

Moreno, P., Fexova, S., George, N., Manning, J. R., Miao, Z., Mohammed, S., Muñoz-Pomer, A., Fullgrabe, A., Bi, Y., Bush, N. *et al.*, "Expression Atlas update: gene and protein expression in multiple species", *Nucleic acids research*, vol. 50, no. D1, D129–D140 2022

Morgan, M., Blank, C., and Seetan, R., "Plant disease prediction using classification algorithms", *IAES International Journal of Artificial Intelligence*, vol. 10 2021a

Morgan, M., Blank, C., and Seetan, R., "Plant disease prediction using classification algorithms", *IAES International Journal of Artificial Intelligence*, vol. 10, no. 1, 257 2021b

Naing, A. H., Jeong, H. Y., Jung, S. K., and Kim, C. K., "Overexpression of 1-Aminocyclopropane-1-carboxylic acid Deaminase (acdS) gene in Petunia hybrida improves tolerance to abiotic stresses", *Frontiers in Plant Science*, vol. 12, 737490 2021

Najafabadi, M. Y., Hesami, M., and Eskandari, M., "Machine Learning-Assisted Approaches in Modern-

ized Plant Breeding Programs", *Genes*, vol. 14 4 2023, URL /pmc/articles/PMC10137951//pmc/articles/PMC10137951/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC10137951/

Nakano, T., Suzuki, K., Fujimura, T., and Shinshi, H., "Genome-wide analysis of the ERF gene family in Arabidopsis and rice", *Plant physiology*, vol. 140, no. 2, 411–432 2006

Narmadha, R., Sengottaiyan, N., and Kavitha, R., "Deep transfer learning based rice plant disease detection model.", *Intelligent Automation & Soft Computing*, vol. 31, no. 2 2022

Niu, X., Luo, T., Zhao, H., Su, Y., Ji, W., and Li, H., "Identification of wheat DREB genes and functional characterization of TaDREB3 in response to abiotic stresses", *Gene*, vol. 740, 144514 2020

Noleto-Dias, C., Picoli, E. A. d. T., Porzel, A., Wessjohann, L. A., Tavares, J. F., and Farag, M. A., "Metabolomics characterizes early metabolic changes and markers of tolerant Eucalyptus ssp. clones against drought stress", *Phytochemistry*, vol. 212, 113715 2023

Offord, V. and Werling, D., "LRRfinder2. 0: a webserver for the prediction of leucine-rich repeats", *Innate Immunity*, vol. 19, no. 4, 398–402 2013

Ogata, T., Ishizaki, T., Fujita, M., and Fujita, Y., "CRISPR/Cas9-targeted mutagenesis of OsERA1 confers enhanced responses to abscisic acid and drought stress and increased primary root growth under non-stressed conditions in rice", *PloS one*, vol. 15, no. 12, e0243376 2020

Oh, S.-J., Kwon, C.-W., Choi, D.-W., Song, S. I., and Kim, J.-K., "Expression of barley HvCBF4 enhances tolerance to abiotic stress in transgenic rice", *Plant Biotechnology Journal*, vol. 5, no. 5, 646–656 2007

Okonechnikov, K., Golosova, O., Fursov, M., and Team, U., "Unipro UGENE: a unified bioinformatics toolkit", *Bioinformatics*, vol. 28, no. 8, 1166–1167 2012

Oladosu, Y., Rafii, M. Y., Samuel, C., Fatai, A., Magaji, U., Kareem, I., Kamarudin, Z. S., Muhammad, I., and Kolapo, K., "Drought resistance in rice from conventional to molecular breeding: a review", *International journal of molecular sciences*, vol. 20, no. 14, 3519 2019

Onufriev, A., Bashford, D., and Case, D. A., "Exploring protein native states and large-scale conformational changes with a modified generalized born model", *Proteins: Structure, Function, and Bioinformatics*, vol. 55, no. 2, 383–394 2004

Osuna-Cruz, C. M., Paytuvi-Gallart, A., Di Donato, A., Sundesha, V., Andolfo, G., Aiese Cigliano, R., Sanseverino, W., and Ercolano, M. R., "PRGdb 3.0: a comprehensive platform for prediction and analysis of plant disease resistance genes", *Nucleic acids research*, vol. 46, no. D1, D1197–D1201 2018

Pagni, M., Ioannidis, V., Cerutti, L., Zahn-Zabal, M., Jongeneel, C. V., Hau, J., Martin, O., Kuznetsov, D., and Falquet, L., "MyHits: improvements to an interactive resource for analyzing protein sequences", *Nucleic acids research*, vol. 35, no. suppl_2, W433–W437 2007

Pal, A. and Levy, Y., "Balance between asymmetry and abundance in multi-domain DNA-binding proteins may regulate the kinetics of their binding to DNA", *PLoS Computational Biology*, vol. 16, no. 5, e1007867 2020

Pal, A., Chakrabarti, P., and Dey, S., "ProDFace: A web-tool for the dissection of protein-DNA interfaces", *Frontiers in Molecular Biosciences*, vol. 9, 978310 2022

Pal, T., Jaiswal, V., and Chauhan, R. S., "DRPPP: A machine learning based tool for prediction of disease resistance proteins in plants", *Computers in Biology and Medicine*, vol. 78, 42–48 11 2016

Pan, J., Li, Z., Wang, Q., Garrell, A. K., Liu, M., Guan, Y., Zhou, W., and Liu, W., "Comparative proteomic investigation of drought responses in foxtail millet", *BMC Plant Biology*, vol. 18, 1–19 2018

Pan, R., Buitrago, S., Feng, Z., Abou-Elwafa, S. F., Xu, L., Li, C., and Zhang, W., "HvbZIP21, a novel transcription factor from wild barley confers drought tolerance by modulating ROS scavenging", *Frontiers in Plant Science*, vol. 13, 878459 2022

Panda, D., Mishra, S. S., and Behera, P. K., "Drought tolerance in rice: focus on recent mechanisms and approaches", *Rice science*, vol. 28, no. 2, 119–132 2021

Pandey, G. K., Grant, J. J., Cheong, Y. H., Kim, B. G., Li, L., and Luan, S., "ABR1, an APETALA2-domain transcription factor that functions as a repressor of ABA response in Arabidopsis", *Plant Physiology*, vol. 139, no. 3, 1185–1193 2005

Park, C.-J. and Ronald, P. C., "Cleavage and nuclear localization of the rice XA21 immune receptor", *Nature*

communications, vol. 3, no. 1, 920 2012

Park, S., Shi, A., Meinhardt, L. W., and Mou, B., "Genome-wide characterization and evolutionary analysis of the AP2/ERF gene family in lettuce (Lactuca sativa)", *Scientific Reports*, vol. 13, no. 1, 21990 2023

Park, S.-I., Kwon, H. J., Cho, M. H., Song, J. S., Kim, B.-G., Baek, J., Kim, S. L., Ji, H., Kwon, T.-R., Kim, K.-H. *et al.*, "The OsERF115/AP2EREBP110 transcription factor is involved in the multiple stress tolerance to heat and drought in rice plants", *International journal of molecular sciences*, vol. 22, no. 13, 7181 2021

Perrakis, A. and Sixma, T. K., "AI revolutions in biology: The joys and perils of AlphaFold", *EMBO reports*, vol. 22, no. 11, e54046 2021

Ponder, J. W. and Case, D. A., "Force fields for protein simulations", *Advances in protein chemistry*, vol. 66, 27–85 2003

Prabhu, K. R., Kumar, A., Yumkhaibam, R. S., Janeja, H. S., Krishna, B., and Talekar, N., "A review on conventional and modern breeding approaches for developing climate resilient crop varieties: NA", *Journal of Applied and Natural Science*, vol. 15, no. 3, 987–997 2023

Pradhan, M., Bastia, D., Samal, K. C., Dash, M., and Sahoo, J. P., "Pyramiding resistance genes for bacterial leaf blight (Xanthomonas oryzae pv. Oryzae) into the popular rice variety, Pratikshya through marker assisted backcrossing", *Molecular Biology Reports*, vol. 50, no. 11, 9047–9060 2023

Puhakainen, T., Hess, M. W., Mäkelä, P., Svensson, J., Heino, P., and Palva, E. T., "Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in Arabidopsis", *Plant molecular biology*, vol. 54, 743–753 2004

Qi, Z., Du, Y., Yu, J., Zhang, R., Yu, M., Cao, H., Song, T., Pan, X., Liang, D., and Liu, Y., "Molecular detection and analysis of blast resistance genes in rice main varieties in Jiangsu province, China", *Agronomy*, vol. 13, no. 1, 157 2023

Qiu, W., Li, H., Song, Y., Ding, J., Chen, S., Ma, C., and Yu, B., "Genome-Wide Analysis of DREB Genes in Sugar Beet and Their Potential Functions in Response to Drought Stress", *Sugar Tech*, pp. 1–17 2024

Rahman, H., Ramanathan, V., Nallathambi, J., Duraialagaraja, S., and Muthurajan, R., "Over-expression of a NAC 67 transcription factor from finger millet (Eleusine coracana L.) confers tolerance against salinity and drought stress in rice", *BMC biotechnology*, vol. 16, 7–20 2016

Rai, M. K., Shekhawat, J. K., Kataria, V., Phulwaria, M., and Shekhawat, N., "Genomic and biotechnological interventions in Prosopis cineraria: current status, challenges and opportunities", *Trees*, pp. 1–13 2021

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z.-W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D. *et al.*, "Regulation of chromatin structure by site-specific histone H3 methyltransferases", *Nature*, vol. 406, no. 6796, 593–599 2000

Remmert, M., Biegert, A., Hauser, A., and Söding, J., "HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment", *Nature methods*, vol. 9, no. 2, 173–175 2012

Rohila, J. S., Jain, R. K., and Wu, R., "Genetic improvement of Basmati rice for salt and drought tolerance by regulated expression of a barley Hva1 cDNA", *Plant Science*, vol. 163, no. 3, 525–532 2002

Roth, J. A. and Galyon, J., "Food security: The ultimate one-health challenge", *One Health*, p. 100864 2024

Sadhukhan, A., Kobayashi, Y., Kobayashi, Y., Tokizawa, M., Yamamoto, Y. Y., Iuchi, S., Koyama, H., Panda, S. K., and Sahoo, L., "VuDREB2A, a novel DREB2-type transcription factor in the drought-tolerant legume cowpea, mediates DRE-dependent expression of stress-responsive genes and confers enhanced drought resistance in transgenic Arabidopsis", *Planta*, vol. 240, 645–664 2014

Saha, A., Das, S., Moin, M., Dutta, M., Bakshi, A., Madhav, M., and Kirti, P., "Genome-wide identification and comprehensive expression profiling of ribosomal protein small subunit (RPS) genes and their comparative analysis with the large subunit (RPL) genes in rice", *Frontiers in plant science*, vol. 8, 1553 2017

Sahu, M., Maurya, S., and Jha, Z., "In vitro selection for drought and salt stress tolerance in rice: An overview", *Plant Physiology Reports*, vol. 28, no. 1, 8–33 2023

Saijo, Y. and Loo, E. P.-i., "Plant immunity in signal integration between biotic and abiotic stress responses", *New Phytologist*, vol. 225, no. 1, 87–104 2020

Sakuma, Y., Liu, Q., Dubouzet, J. G., Abe, H., Shinozaki, K., and Yamaguchi-Shinozaki, K., "DNA-binding

specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression", *Biochemical and biophysical research communications*, vol. 290, no. 3, 998–1009 2002

Sakuma, Y., Maruyama, K., Osakabe, Y., Qin, F., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K., "Functional analysis of an Arabidopsis transcription factor, DREB2A, involved in drought-responsive gene expression", *The Plant Cell*, vol. 18, no. 5, 1292–1309 2006

Salazar-Retana, A. L., Maruri-Lopez, I., Hernández-Sánchez, I. E., Becerra-Flora, A., de la Luz Guerrero-Gonzalez, M., and Jiménez-Bremont, J. F., "PEST sequences from a cactus dehydrin regulate its proteolytic degradation", *PeerJ*, vol. 7, e6810 2019

Salmon, Y., Dietrich, L., Sevanto, S., Hölttä, T., Dannoura, M., and Epron, D., "Drought impacts on tree phloem: from cell-level responses to ecological significance", *Tree physiology*, vol. 39, no. 2, 173–191 2019

Sarkar, C., Gupta, D., Gupta, U., and Hazarika, B. B., "Leaf disease detection using machine learning and deep learning: Review and challenges", 2023

Schrödinger, L. and DeLano, W., "PyMOL", 2020, URL `http://www.pymol.org/pymol`

Shamsudin, N. A. A., Swamy, B. M., Ratnam, W., Sta. Cruz, M. T., Raman, A., and Kumar, A., "Marker assisted pyramiding of drought yield QTLs into a popular Malaysian rice cultivar, MR219", *BMC genetics*, vol. 17, 1–14 2016

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T., "Cytoscape: a software environment for integrated models of biomolecular interaction networks", *Genome research*, vol. 13, no. 11, 2498–2504 2003

Sharma, D., Verma, N., Pandey, C., Verma, D., Bhagat, P. K., Noryang, S., Singh, K., Tayyeba, S., Banerjee, G., and Sinha, A. K., "MAP kinase as regulators for stress responses in plants: An overview", *Protein Kinases and Stress Signaling in Plants: Functional Genomic Perspective*, pp. 369–392 2020

Sharma, M., Kumar, C. J., Talukdar, J., Singh, T. P., Dhiman, G., and Sharma, A., "Identification of rice leaf diseases and deficiency disorders using a novel DeepBatch technique", *Open Life Sciences*, vol. 18 1 2023, URL `https://www.degruyter.com/document/doi/10.1515/biol-2022-0689/html`

Sheik, S., Sundararajan, P., Hussain, A., and Sekar, K., "Ramachandran plot on the web", *Bioinformatics*, vol. 18, no. 11, 1548–1549 2002

Shendekar, S., Kute, N., Madhu, B., Gadpayale, D., Meshram, M., Basavaraj, P., and Totre, A., "Unlocking crop potential: speed breeding and its synergies with modern breeding techniques", in "Biol Forum Int J", vol. 15, pp. 89–100 2023

Shiraku, M. L., Magwanga, R. O., Cai, X., Kirungu, J. N., Xu, Y., Mehari, T. G., Hou, Y., Wang, Y., Wang, K., Peng, R. *et al.*, "Knockdown of 60S ribosomal protein L14-2 reveals their potential regulatory roles to enhance drought and salt tolerance in cotton", *Journal of Cotton Research*, vol. 4, 1–14 2021

Shumayla, Sharma, S., Kumar, R., Mendu, V., Singh, K., and Upadhyay, S. K., "Genomic dissection and expression profiling revealed functional divergence in Triticum aestivum leucine rich repeat receptor like kinases (TaLRRKs)", *Frontiers in Plant Science*, vol. 7, 1374 2016

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J. *et al.*, "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega", *Molecular systems biology*, vol. 7, no. 1, 539 2011

Silva, J. C. F., Ferreira, M. A., Carvalho, T. F., Silva, F. F., de A. Silveira, S., Brommonschenkel, S. H., and Fontes, E. P., "RLPredictiOme, a Machine Learning-Derived Method for High-Throughput Prediction of Plant Receptor-like Proteins, Reveals Novel Classes of Transmembrane Receptors", *International Journal of Molecular Sciences*, vol. 23, 12176 10 2022, URL `https://www.mdpi.com/1422-0067/23/20/12176/htmhttps://www.mdpi.com/1422-0067/23/20/12176`

Silva, R. J. S. and Micheli, F., "RRGPredictor, a set-theory-based tool for predicting pathogen-associated molecular pattern receptors (PRRs) and resistance (R) proteins from plants", *Genomics*, vol. 112, no. 3, 2666–2676 2020

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M., "BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs", *Bioinformatics*, vol. 31,

no. 19, 3210–3212 2015
Simón, D., Borsani, O., and Filippi, C. V., "RFPDR: a random forest approach for plant disease resistance protein prediction", *PeerJ*, vol. 10, e11683 2022
Singh, B. K. and Trivedi, P., "Microbiome and the future for food and nutrient security", *Microbial biotechnology*, vol. 10, no. 1, 50 2017
Singh, D., Dhiman, V. K., Pandey, H., Dhiman, V. K., and Pandey, D., "Crosstalk between salicylic acid and auxins, cytokinins and gibberellins under biotic stress", in "Auxins, cytokinins and gibberellins signaling in plants", pp. 249–262, Springer 2022
Singh, R. and Hymowitz, T., "The genomic relationship between Glycine max (L.) Merr. and G. soja Sieb. and Zucc. as revealed by pachytene chromosome analysis", *Theoretical and Applied Genetics*, vol. 76, no. 5, 705–711 1988
Snel, B., Lehmann, G., Bork, P., and Huynen, M. A., "STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene", *Nucleic acids research*, vol. 28, no. 18, 3442–3444 2000
Song, F. and Goodman, R. M., "Molecular biology of disease resistance in rice", *Physiological and Molecular Plant Pathology*, vol. 59, no. 1, 1–11 2001
Su, J., Song, S., Wang, Y., Zeng, Y., Dong, T., Ge, X., and Duan, H., "Genome-wide identification and expression analysis of DREB family genes in cotton", *BMC Plant Biology*, vol. 23, no. 1, 169 2023
Sudalaimuthuasari, N., Ali, R., Kottackal, M., Rafi, M., Al Nuaimi, M., Kundu, B., Al-Maskari, R. S., Wang, X., Mishra, A. K., Balan, J. *et al.*, "The genome of the mimosoid legume Prosopis cineraria, a desert tree", *International journal of molecular sciences*, vol. 23, no. 15, 8503 2022
Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P. *et al.*, "STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets", *Nucleic acids research*, vol. 47, no. D1, D607–D613 2019
Tenhaken, R., "Cell wall remodeling under abiotic stress", *Frontiers in plant science*, vol. 5, 771 2015
Thorvaldsdóttir, H., Robinson, J. T., and Mesirov, J. P., "Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration", *Briefings in bioinformatics*, vol. 14, no. 2, 178–192 2013
Törönen, P. and Holm, L., "PANNZER—a practical tool for protein function prediction", *Protein Science*, vol. 31, no. 1, 118–128 2022
Trejos, M., Aristizabal, Y., Aragón-Muriel, A., Oñate-Garzón, J., and Liscano, Y., "Characterization and classification in silico of peptides with dual activity (antimicrobial and wound healing)", *International Journal of Molecular Sciences*, vol. 24, no. 17, 13091 2023
Tusnady, G. E. and Simon, I., "The HMMTOP transmembrane topology prediction server", *Bioinformatics*, vol. 17, no. 9, 849–850 2001
Tyagi, A., Mir, Z. A., Almalki, M. A., Deshmukh, R., and Ali, S., "Genomics-Assisted Breeding: A Powerful Breeding Approach for Improving Plant Growth and Stress Resilience", *Agronomy*, vol. 14, no. 6, 1128 2024
Udage, A., "Introduction to plant mutation breeding: Different approaches and mutagenic agents.", *Journal of Agricultural Sciences (Sri Lanka)*, vol. 16, no. 3 2021
Udayashankar, A., Nayaka, S. C., Reddy, M., Srinivas, C. *et al.*, "Plant growth-promoting rhizobacteria mediate induced systemic resistance in rice against bacterial leaf blight caused by Xanthomonas oryzae pv. oryzae", *Biological Control*, vol. 59, no. 2, 114–122 2011
Umar, O. B., Ranti, L. A., Abdulbaki, A. S., Bola, A., Abdulhamid, A. K., Biola, M., and Victor, K., "Stresses in plants: Biotic and abiotic", *Current Trends in Wheat Research*, pp. 1–8 2021
Vaistij, F. E., Barros-Galvão, T., Cole, A. F., Gilday, A. D., He, Z., Li, Y., Harvey, D., Larson, T. R., and Graham, I. A., "MOTHER-OF-FT-AND-TFL1 represses seed germination under far-red light by modulating phytohormone responses in Arabidopsis thaliana", *Proceedings of the National Academy of Sciences*, vol. 115, no. 33, 8442–8447 2018
Viana, C. M., Freire, D., Abrantes, P., Rocha, J., and Pereira, P., "Agricultural land systems importance for supporting food security and sustainable development goals: A systematic review", *Science of the total*

environment, vol. 806, 150718 2022

Viana, V. E., Pegoraro, C., Busanello, C., and Costa de Oliveira, A., "Mutagenesis in rice: the basis for breeding a new super plant", *Frontiers in plant science*, vol. 10, 1326 2019

Vo, K. T. X., Rahman, M. M., Rahman, M. M., Trinh, K. T. T., Kim, S. T., and Jeon, J. S., "Proteomics and Metabolomics Studies on the Biotic Stress Responses of Rice: an Update", 2021

Wang, F., Wang, C., Liu, P., Lei, C., Hao, W., Gao, Y., Liu, Y.-G., and Zhao, K., "Enhanced rice blast resistance by CRISPR/Cas9-targeted mutagenesis of the ERF transcription factor gene OsERF922", *PloS one*, vol. 11, no. 4, e0154027 2016

Wang, H., Lu, S., Guan, X., Jiang, Y., Wang, B., Hua, J., and Zou, B., "Dehydration-responsive element binding protein 1C, 1E, and 1G promote stress tolerance to chilling, heat, drought, and salt in rice", *Frontiers in plant science*, vol. 13, 851731 2022a

Wang, L., Zhao, L., Zhang, X., Zhang, Q., Jia, Y., Wang, G., Li, S., Tian, D., Li, W.-H., and Yang, S., "Large-scale identification and functional analysis of NLR genes in blast resistance in the Tetep rice genome sequence", *Proceedings of the National Academy of Sciences*, vol. 116, no. 37, 18479–18487 2019

Wang, N., Gao, Z., Zhang, W., Qian, Y., Bai, D., Zhao, X., Bao, Y., Zheng, Z., Wang, X., Li, J. *et al.*, "Genome-Wide association analysis reveals the gene loci of yield traits under drought stress at the rice reproductive stage", *Agronomy*, vol. 13, no. 8, 2096 2023a

Wang, X., Bigman, L. S., Greenblatt, H. M., Yu, B., Levy, Y., and Iwahara, J., "Negatively charged, intrinsically disordered regions can accelerate target search by DNA-binding proteins", *Nucleic acids research*, vol. 51, no. 10, 4701–4712 2023b

Wang, Y., Xu, L., Zou, Q., and Lin, C., "prPred-DRLF: Plant R protein predictor using deep representation learning features", *Proteomics*, vol. 22 2022b

Wanga, M. A., Shimelis, H., Mashilo, J., and Laing, M. D., "Opportunities and challenges of speed breeding: A review", *Plant Breeding*, vol. 140, no. 2, 185–194 2021

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L. *et al.*, "SWISS-MODEL: homology modelling of protein structures and complexes", *Nucleic acids research*, vol. 46, no. W1, W296–W303 2018

Wickham, H. and Wickham, H., *Data analysis*, Springer 2016

Wilkinson, L., "ggplot2: elegant graphics for data analysis by WICKHAM, H.", 2011

Wilkinson, S. P. and Davy, S. K., "phylogram: an R package for phylogenetic analysis with nested lists", *Journal of Open Source Software*, vol. 3, no. 26, 790 2018, URL `https://doi.org/10.21105/joss.00790`

Xiao, N., Cao, D. S., Zhu, M. F., and Xu, Q. S., "protr/ProtrWeb: R package and web server for generating various numerical representation schemes of protein sequences", *Bioinformatics*, vol. 31, 1857–1859 6 2015, URL `https://dx.doi.org/10.1093/bioinformatics/btv042`

Xie, X., Chen, Z., Zhang, B., Guan, H., Zheng, Y., Lan, T., Zhang, J., Qin, M., and Wu, W., "Transcriptome analysis of xa5-mediated resistance to bacterial leaf streak in rice (Oryza sativa L.)", *Scientific Reports*, vol. 10, no. 1, 19439 2020

Xin, X., Su, T., Li, P., Wang, W., Zhao, X., Yu, Y., Zhang, D., Yu, S., and Zhang, F., "A histone H4 gene prevents drought-induced bolting in Chinese cabbage by attenuating the expression of flowering genes", *Journal of Experimental Botany*, vol. 72, no. 2, 623–635 2021

Xing, H., Jiang, Y., Zou, Y., Long, X., Wu, X., Ren, Y., Li, Y., and Li, H.-L., "Genome-wide investigation of the AP2/ERF gene family in ginger: Evolution and expression profiling during development and abiotic stresses", *BMC Plant Biology*, vol. 21, 1–21 2021

Xu, D. and Zhang, Y., "Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization", *Biophysical journal*, vol. 101, no. 10, 2525–2534 2011

Xu, L., Lan, Y., Lin, M., Zhou, H., Ying, S., and Chen, M., "Genome-Wide Identification and Transcriptional Analysis of AP2/ERF Gene Family in Pearl Millet (Pennisetum glaucum)", *International Journal of Molecular Sciences*, vol. 25, no. 5, 2470 2024

Xu, W., Tang, W., Wang, C., Ge, L., Sun, J., Qi, X., He, Z., Zhou, Y., Chen, J., Xu, Z. *et al.*, "SiMYB56

confers drought stress tolerance in transgenic rice by regulating lignin biosynthesis and ABA signaling pathway", *Frontiers in Plant Science*, vol. 11, 785 2020

Yadav, S., Sandhu, N., Dixit, S., Singh, V. K., Catolos, M., Mazumder, R. R., Rahman, M. A., and Kumar, A., "Genomics-assisted breeding for successful development of multiple-stress-tolerant, climate-smart rice for southern and southeastern Asia", *The Plant Genome*, vol. 14, no. 1, e20074 2021

Yan, W., Zhao, J.-M., Zhang, L.-X., Ping, W., Wang, S.-W., Hui, W., Wang, X.-X., Liu, Z.-H., and Zheng, W.-J., "Analysis of the diversity and function of the alleles of the rice blast resistance genes Piz-t, Pita and Pik in 24 rice cultivars", *Journal of Integrative Agriculture*, vol. 15, no. 7, 1423–1431 2016

Yan, Y., Tao, H., He, J., and Huang, S.-Y., "The HDOCK server for integrated protein–protein docking", *Nature protocols*, vol. 15, no. 5, 1829–1852 2020

Yoosefzadeh-Najafabadi, M., Earl, H. J., Tulpan, D., Sulik, J., and Eskandari, M., "Application of Machine Learning Algorithms in Plant Breeding: Predicting Yield From Hyperspectral Reflectance in Soybean", *Frontiers in Plant Science*, vol. 11 2021

Yu, F., Qian, L., Nibau, C., Duan, Q., Kita, D., Levasseur, K., Li, X., Lu, C., Li, H., Hou, C. *et al.*, "FERONIA receptor kinase pathway suppresses abscisic acid signaling in Arabidopsis by activating ABI2 phosphatase", *Proceedings of the National Academy of Sciences*, vol. 109, no. 36, 14693–14698 2012

Zafar, M. M., Rehman, A., Razzaq, A., Parvaiz, A., Mustafa, G., Sharif, F., Mo, H., Youlu, Y., Shakeel, A., and Ren, M., "Genome-wide characterization and expression analysis of Erf gene family in cotton", *BMC plant biology*, vol. 22, no. 1, 134 2022

Zaharias, S., Zhang, Z., Davis, K., Fargason, T., Cashman, D., Yu, T., and Zhang, J., "Intrinsically disordered electronegative clusters improve stability and binding specificity of RNA-binding proteins", *Journal of Biological Chemistry*, vol. 297, no. 2 2021

Zainuddin, F., Ismail, M. R., Hatta, M. A. M., and Ramlee, S. I., "Advancement in modern breeding and genomic approaches to accelerate rice improvement: speed breeding focus", *Euphytica*, vol. 220, no. 7, 109 2024

Zhai, K., Deng, Y., Liang, D., Tang, J., Liu, J., Yan, B., Yin, X., Lin, H., Chen, F., Yang, D. *et al.*, "RRM transcription factors interact with NLRs and regulate broad-spectrum blast resistance in rice", *Molecular cell*, vol. 74, no. 5, 996–1009 2019

Zhang, H., Pei, Y., He, Q., Zhu, W., Jahangir, M., Haq, S. u., Khan, A., and Chen, R., "Salicylic acid-related ribosomal protein CaSLP improves drought and Pst. DC3000 tolerance in pepper", *Molecular Horticulture*, vol. 3, no. 1, 6 2023a

Zhang, J., Chen, L., Fu, C., Wang, L., Liu, H., Cheng, Y., Li, S., Deng, Q., Wang, S., Zhu, J. *et al.*, "Comparative transcriptome analyses of gene expression changes triggered by Rhizoctonia solani AG1 IA infection in resistant and susceptible rice varieties", *Frontiers in plant science*, vol. 8, 1422 2017

Zhang, J., Yang, Y., Feng, X., Xu, H., Chen, J., and He, Y., "Identification of Bacterial Blight Resistant Rice Seeds Using Terahertz Imaging and Hyperspectral Imaging Combined With Convolutional Neural Network", *Frontiers in Plant Science*, vol. 11 2020

Zhang, X., Liu, S., and Takano, T., "Two cysteine proteinase inhibitors from Arabidopsis thaliana, AtCYSa and AtCYSb, increasing the salt, drought, oxidation and cold tolerance", *Plant Molecular Biology*, vol. 68, 131–143 2008

Zhang, Y., Zhu, J., Khan, M., Wang, Y., Xiao, W., Fang, T., Qu, J., Xiao, P., Li, C., and Liu, J.-H., "Transcription factors ABF4 and ABR1 synergistically regulate amylase-mediated starch catabolism in drought tolerance", *Plant physiology*, vol. 191, no. 1, 591–609 2023b

Zhang, Z.-L., Xie, Z., Zou, X., Casaretto, J., Ho, T.-h. D., and Shen, Q. J., "A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells", *Plant Physiology*, vol. 134, no. 4, 1500–1513 2004

Zhao, M., Haxim, Y., Liang, Y., Qiao, S., Gao, B., Zhang, D., and Li, X., "Genome-wide investigation of AP2/ERF gene family in the desert legume Eremosparton songoricum: Identification, classification, evolution, and expression profiling under drought stress", *Frontiers in Plant Science*, vol. 13, 885694 2022a

Zhao, Y., Zhu, X., Chen, X., and Zhou, J. M., "From plant immunity to crop disease resistance", *Journal of*

Genetics and Genomics, vol. 49, 693–703 8 2022b

Zhou, L., Feng, T., Xu, S., Gao, F., Lam, T. T., Wang, Q., Wu, T., Huang, H., Zhan, L., Li, L. *et al.*, "ggmsa: A visual exploration tool for multiple sequence alignment and associated data", *Briefings in Bioinformatics*, vol. 23, no. 4, bbac222 2022

Zhu, M., Zheng, L., Cao, S., Liu, Q., Wei, S., Zhou, Y., and Gao, F., "AnDREB5. 1, a A5 group DREB gene from desert shrub Ammopiptanthus nanus, confers osmotic and cold stress tolerances in transgenic tobacco", *Physiologia Plantarum*, vol. 176, no. 2, e14272 2024

Zhu, X., Wang, B., Liu, W., Wei, X., Wang, X., Du, X., and Liu, H., "Genome-wide analysis of AP2/ERF gene and functional analysis of CqERF24 gene in drought stress in quinoa", *International Journal of Biological Macromolecules*, vol. 253, 127582 2023

Zhuang, J., Cai, B., Peng, R.-H., Zhu, B., Jin, X.-F., Xue, Y., Gao, F., Fu, X.-Y., Tian, Y.-S., Zhao, W. *et al.*, "Genome-wide analysis of the AP2/ERF gene family in Populus trichocarpa", *Biochemical and biophysical research communications*, vol. 371, no. 3, 468–474 2008

Żmieńko, A., Samelak, A., Kozłowski, P., and Figlerowicz, M., "Copy number polymorphism in plant genomes", *Theoretical and applied genetics*, vol. 127, 1–18 2014

Appendix A

Supplementary information for Chapter 3

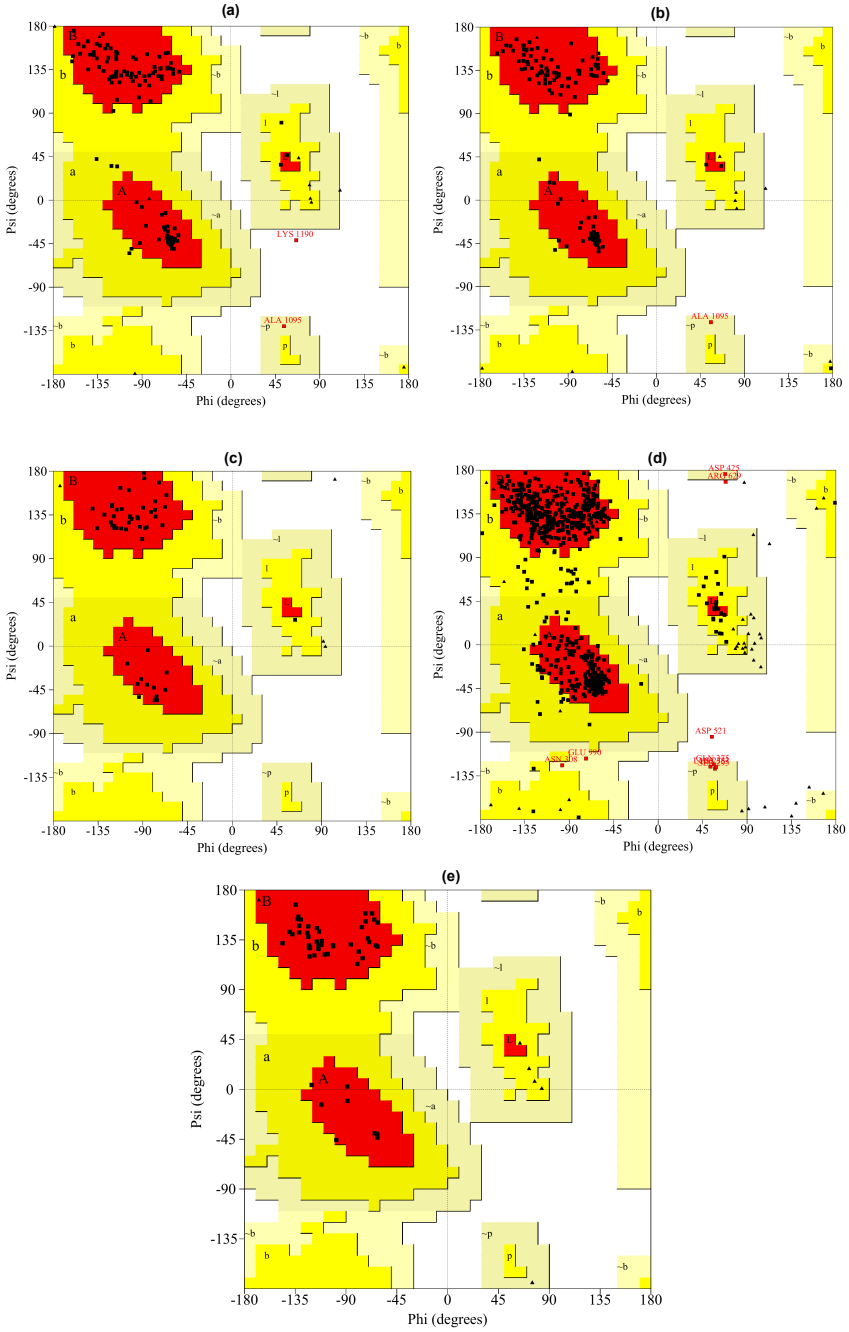


Figure A.1 : Analysis of the stereochemical properties of the 3D models of the disease-resistance proteins against bacterial blight and rice blast diseases in the rice plant. (a-b) Ramachandran plots of the 3D models of the disease-resistance proteins against bacterial blight disease. The stereochemical properties of the proteins BB.1 and BB.2 were identified using the Ramachandran plot server and PROCHECK tool. The Ramachandran plots describe the peptide backbone torsion angles, phi and psi (ϕ and ψ). **(c-e)** The stereochemical properties of the proteins RB.1, RB.2, and RB.3 were examined through the Ramachandran plot server and PROCHECK tool. In the plots, the most favored region, additionally allowed regions, generously allowed regions, and disallowed regions are indicated as red, yellow, light yellow, and white, respectively.

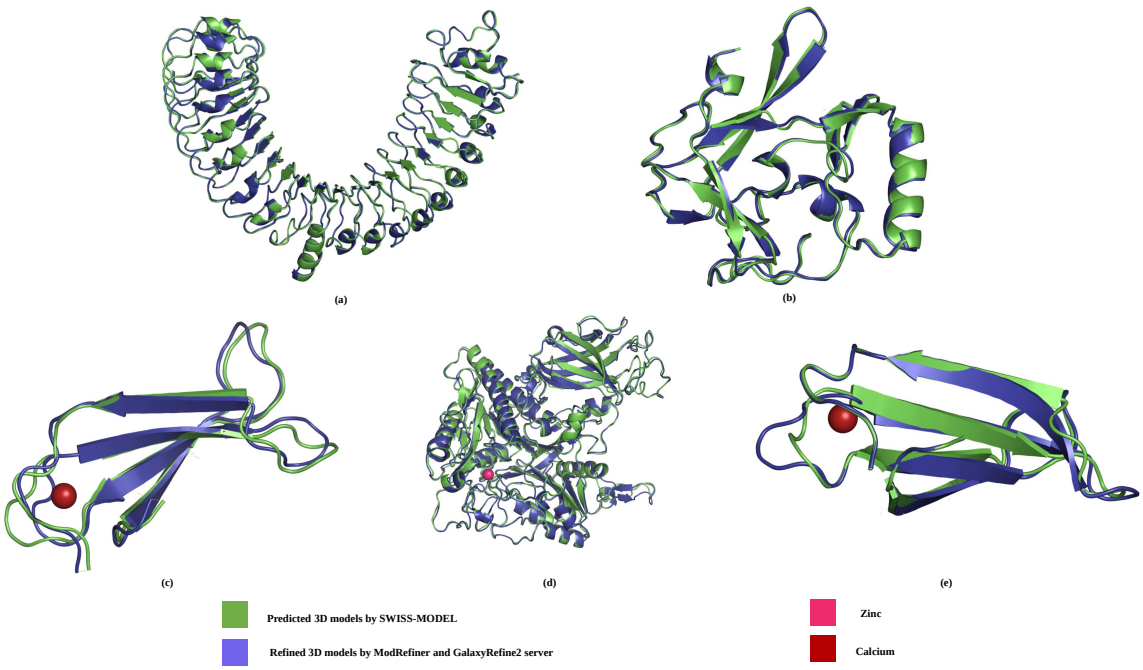


Figure A.2 : Structural alignment of the 3D models of the disease-resistance proteins against bacterial blight and rice blast diseases in the rice plant predicted by the SWISS-MODEL with the I-TASSER and AlphaFold. (a-d) Alignment of the 3D models predicted by the SWISS-MODEL with I-TASSER and Alphafold for the proteins BB.1 and BB.2 is shown. The alignment was performed by the 'align' function in the PyMOL, which superimposes the two protein structures on each other based on the homology between them. **(e-j)** Structural alignment of the 3D models of the proteins RB.1, RB.2, and RB.3 predicted by the SWISS-MODEL with I-TASSER and Alphafold is depicted. The alignment of disease-resistance proteins against bacterial blight and rice blast diseases in the rice plant indicates that SWISS-MODEL was the best-suited method compared to the I-TASSER and Alphafold.

Table A.1 : Overview of different databases and tools used in the study.

Categories	Description
Databases	
SNP-Seek database	Database providing information on genotype, phenotype, and variety for rice derived from the 3,000 Rice Genomes Project.
Expasy STRING (Search Tool for the Retrieval of Interacting Genes/Proteins)	Knowledgebase and software tool for protein-protein interactions
UniProt Knowledgebase (UniProtKB)	Richly curated protein database having protein sequence and functional information
ModelArchive	Archive for structural models which are not based on experimental data and are computationally derived

Conserved Domain Database (CDD)	Resource for the annotation of protein sequences based on the location of conserved domain footprints
DRAGO	Web-based tool in Plant Resistance Genes Database (PRGdb) database for providing a comprehensive overview of resistance genes (R-genes) in plants
STRING database	Database of known and predicted protein-protein interactions
EMBL-EBI Expression Atlas	open science resource providing gene and protein expression information
Tools	
Biomart tool	Open-source data management web tool for data refinement and retrieval available in the Ensembl plants database
Cytoscape	Open-source software platform for visualizing biomolecular interaction networks
PSI-BLAST Based Secondary Structure Prediction (PSIPRED)	Protein Analysis Workbench/tool for prediction of secondary structure
SWISS-MODEL	Web-based protein structure homology-modelling server
GalaxyRefine2 server	Web server for protein structure prediction, refinement, and related methods
ModRefiner	Algorithm for atomic-level, high-resolution protein structure refinement
Pymol	An open-source molecular visualization system for biomolecules, especially proteins
Ramachandran plot server	Web-based server for constructing Ramachandran plot
PROCHECK	Web-based tool for checking the stereochemical quality of a protein structure
Verify3D	Web-based tool for determining the compatibility of an atomic model (3D) with its own amino acid sequence (1D)
ERRAT	Web-based tool for evaluating protein models
ExPASy ProtParam	A web-based tool that computes various physical and chemical parameters for a given protein
CYSPRED	Neural-Network-based predictor for predicting the bonding state of cysteines in proteins

ModelArchive	Archive for structural models which are not based on experimental data and are computationally derived
CELLO2GO	Web server for protein subcellular localization prediction with functional gene ontology annotation
HMMTOP	Automatic server for predicting transmembrane helices and topology of proteins
InterProScan	Web-based tool for functional analysis of proteins by classifying them into families and predicting domains
ScanProsite	Web-based tool for detecting PROSITE signature matches in protein sequences.
Motif Scan	Finds all known motifs occurring in a protein sequence
ggmsa package (R software)	Tool for visualization and annotation of multiple sequence alignment of protein/DNA/RNA
phylogram package (R software)	Tool for developing phylogenetic trees
PANNZER (Protein Annotation with Z-score)	Webserver for functional annotation of proteins of unknown function
Assign KO tool (KEGG Database)	Interface to the BlastKOALA server for pathway analysis
Leucine Rich Repeats (LRR) predictor	Webserver for predicting leucine-rich repeats

Categories: tools and databases used in the study; Description: short detail of the tool and databases used.

Table A.2 : The secondary structure profile of the five unknown disease-resistant proteins.

Structure	**Number of amino acid residues**				
	BB.1	**BB.2**	**RB.1**	**RB.2**	**RB.3**
Helix	280	363	94	191	61
Coil	603	758	554	680	400

Structure: number of helices and coils present in the protein; BB.1 and BB.2: disease-resistant protein against bacterial blight disease; RB.1, RB.2 and RB.3: disease-resistant protein against rice blast disease.

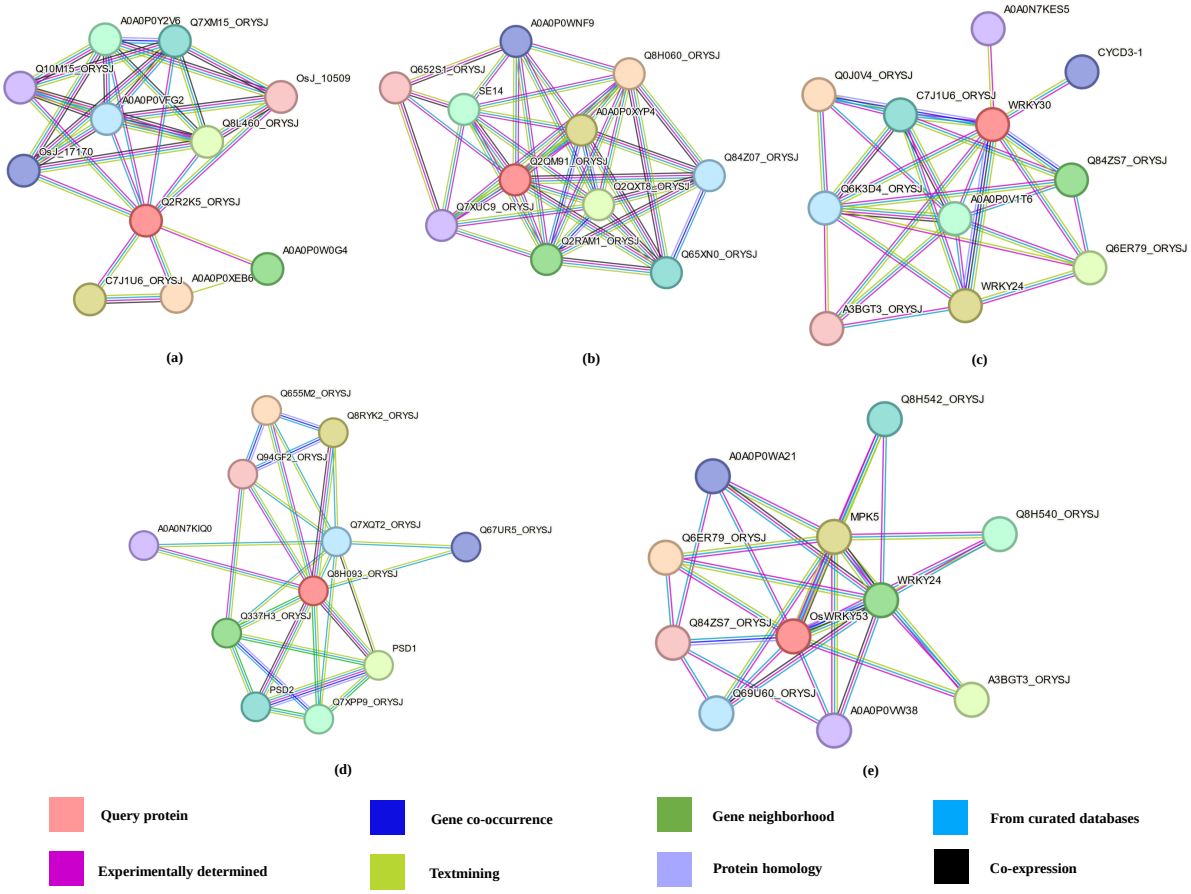


Figure A.3 : Structural alignment between the predicted 3D and refined models for the disease-resistance proteins against the bacterial blight and rice blast diseases. (a-b) Alignment of the proteins BB.1 and BB.2 against bacterial blight disease performed by the PyMOL is shown. The alignment between 3D and refined models of BB.1 and BB.2 was performed by the 'align' function in the PyMOL, which superimposes the two protein structures on each other based on the homology between them. **(c-e)** PyMOL was utilized to perform the structural alignment between the 3D and refined models of the proteins RB.1, RB.2, and RB.3. The alignment indicates the 3D models before and after the refinement executed by the Galaxy Refine2 server and ModRefiner.

Table A.3 : The stereochemical analysis of the five disease-resistant proteins using different tools.

Proteins	Ramachandran Plot	PROCHECK	Verify 3D	ERRAT
BB.1	0.70%	0.80%	96.58%	94.11
BB.2	0%	0%	100%	95.58
RB.1	0%	0%	80.70%	100
RB.2	0.10%	0.30%	91.11%	79.63
RB.3	0%	0%	83.05%	84.09

Protein: five disease-resistant proteins; PROCHECK: checks the stereochemical quality of protein structure by analyzing residue-by-residue geometry and overall structure geometry; Ramachandran Plot: refers to the percentage of amino acids in the disallowed region; Verify 3D: compatibility percentage of an atomic model (3D) with its own amino acid sequence (1D); ERRAT: depict the score (out of 100) of the non-bonded interactions between different atom types.

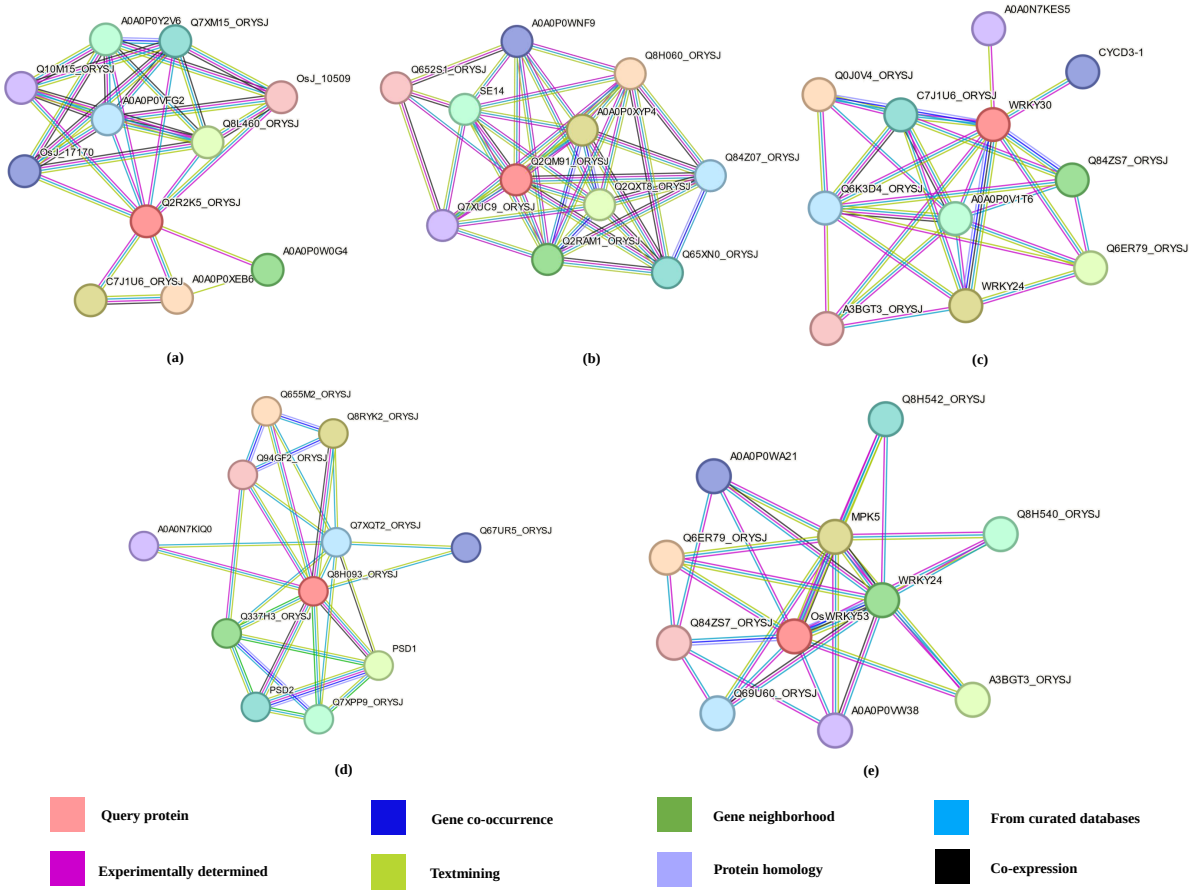


Figure A.4 : Protein-protein interactions between the identified disease-resistance and known disease-resistance proteins against bacterial blight and rice blast diseases in the rice plant. (a-b) Protein-protein interactions between proteins BB.1 and BB.2 and known disease-resistance proteins against bacterial blight disease. The interactions were constructed and visualized on the amino acid sequences of BB.1 and BB.2 in the Expasy STRING database. **(c-e)** A protein-protein network of the proteins RB.1, RB.2, and RB.3 with known disease-resistance proteins was constructed by utilizing the Expasy STRING database. The query proteins **(a-e)** were labelled with light pink, whereas other colors indicated the association of the query proteins with the known disease-resistance proteins. Purple, dark blue, green, dark green, light violet, light blue, and black colors indicate different types of associations between query and the known disease-resistance proteins, viz., experimentally-determined, gene co-occurrence, text mining, gene neighborhood, protein homology from curated databases, and co-expression.

Table A.4 : Comparison between SWISS-MODEL, I-TASSER and AlphaFold predicting five disease-resistant proteins.

Proteins	SWISS-MODEL	I-TASSER	AlphaFold
	Ramachandran outliers		
BB.1	0.8%	4.7%	0.6%
BB.2	0%	3.9%	17.3%
RB.1	0%	2.8%	31.9%
RB.2	0.3%	6.2%	2.8%
RB.3	0%	3.2%	20.2%

SWISS-MODEL: a fully automated protein homology modeling server; I-TASSER: a server with an integrated platform for protein structure and function prediction based on the sequence-to-structure-to-function paradigm; AlphaFold: an Artificial Intelligence (AI) based program to predict protein 3D structures; Ramachandran outliers or amino acid residues in the disallowed region: amino acid residues present in the disallowed region of the Ramachandran plot.

Table A.5 : Shows CYSPRED predicted cysteine residues playing important in disulphide bonding.

Proteins	Number of cysteine residues	Predictions	Reliability score
BB.1	21	Bonding state (3), non-bonding state (18)	9
BB.2	30	Non-bonding state	9
RB.1	8	Non-bonding state	9
RB.2	14	Non-bonding state	9
RB.3	8	Non-bonding state	9

Number of cysteine residues: total number of cysteine residues present in the proteins; Predictions: bonding and non-bonding state of the cysteine residues; Reliability score: score for the reliability of the predictions.

Table A.6 : Information on pathway analysis of the disease-resistance proteins.

Proteins	Pathways involved	Definition	Functional category
BB.1	Transcription (Basal transcription factors)	TFIIB, GTF2B, SUA7, tfb; transcription initiation factor TFIIB	Genetic Information Processing
BB.2	Transcription (Basal transcription factors)	TAF5; transcription initiation factor TFIID subunit 5	Genetic Information Processing
RB.1	Environmental adaptation (plant-pathogen interaction)	WRKY2; WRKY transcription factor 2	Organismal systems
RB.2	Metabolism (secondary metabolites biosynthesis, Glycerophospholipid, and ether lipid metabolism)	PLD1_2; phospholipase D1/2	Lipid metabolism
	Environmental Information Processing (Ras, phospholipase D, sphingolipid, cAMP signaling pathways)		
RB.3	Environmental Information Processing (MAPK signaling pathway), Environmental adaptation (plant-pathogen interaction)	WRKY33; WRKY transcription factor 33	Environmental Information Processing

Proteins: five disease-resistant proteins; Pathways involved: details of the identified pathways; Definition: short details of the disease-resistant proteins; Functional category: description of the biologically identified functional groups in the proteins.

Table A.7 : Comparison between the pIDDT and RMSD values of the SWISS-MODEL, AlphaFold and I-TASSER for the five disease-resistance proteins.

Proteins	**SWISS-MODEL**	**Alphafold**	**I-TASSER**
	pIDDT		**RMSD**
BB.1	404	150	13.1±4.1
BB.2	808	50	13.6±4.0
RB.1	600	100	10.0±4.6
RB.2	640	50	12.2±4.4
RB.3	575	380	11.0±4.6

Proteins: five disease-resistance proteins; SWISS-MODEL, AlphaFold and I-TASSER: different tools used to model the identified disease-resistance proteins; pIDDT: predicted local distance difference (pIDDT) test score; RMSD: Root Mean Square Deviation.

Appendix B

Supplementary information for Chapter 4

Table B.2 : Summary of the predicted disease-resistance proteins by the DL-based MLP model on the external dataset.

Protein IDs	Description	Functional role
NP_001384967.1	Cinnamoyl-CoA reductase-like SNL6	Resistance against the bacterial pathogen, *Xanthomonas oryzae* pv. *oryzae* (*Xoo*)
P0C0M2.1	Auxin-responsive GH3-like protein 2	Broad-spectrum resistance in *Oryza sativa*
Q52QH4.1	NAC domain-containing protein 68	Resistance against *Xoo* and fungal pathogen *Magnaporthe oryzae* (*M. oryzae*)
Q5N712.1	Lysine-specific demethylase JMJ705	Resistance against *Xoo*
XP_025878158.1	3-phosphoinositide-dependent protein kinase 1 isoform X2	Resistance against *Xoo* and fungal pathogen (*M. oryzae*)
P49027.1	Small ribosomal subunit protein RACK1z	Innate immunity
XP_015646387.1	Putative disease resistance RPP13-like protein 1	Innate immunity
Q7F2L3.1	NAC domain-containing protein 48	Disease resistance
Q6EPW7.2	Cysteine and histidine-rich domain-containing protein RAR1	Innate immunity
Q6ZHA3.1	Rac-like GTP-binding protein 6	Innate immunity
Q6ZH94.1	S-(+)-linalool synthase, chloroplastic	Resistance against *Xoo*
Q6Z437.1	Mitogen-activated protein kinase 3	Resistance against *Xoo*
XP_015631594.1	Serine/threonine-protein kinase EDR1	Disease resistance
XP_015630864.1	Leucine-rich repeat receptor-like serine/threonine-protein kinase BAM1	Disease resistance
XP_015627771.1	Omega-3 fatty acid desaturase, endoplasmic reticulum	Resistance against *Magnaporthe grisea* (*M. grisea*)
Q84T94.1	Protein phosphatase 2C 35	Innate immunity
Q7FAZ3.1	G-type lectin S-receptor-like serine/threonine-protein kinase LECRK1	Innate immunity
XP_015633586.1	Serine/threonine-protein kinase OXI1	Disease resistance
NP_001406724.1	Calcium-dependent protein kinase 12	Innate immunity
A3AZ88.1	Metallothionein-like protein 2C	Disease resistance
Q65XV2.1	E3 ubiquitin-protein ligase XB3	Resistance against *Xoo*
XP_015637594.1	Probable WRKY transcription factor 70	Disease resistance
XP_015640814.1	1-aminocyclopropane-1-carboxylate oxidase	Disease resistance

Q6I581.1	Jasmonoyl-L-amino acid synthetase GH3.5; Auxin-responsive GH3-like protein 5; Indole-3-aceti acid-amido synthetase GH3.5; Jasmonic acid-amido synthetase JAR1; Protein JASMONATE RESISTANT 1	Disease resistance against *M. oryzae*
XP_015643825.1	Serine-threonine kinase receptor-associated protein	Disease resistance
XP_015641578.1	WRKY transcription factor WRKY28-like	Disease resistance against *M. oryzae*
BAS97287.1	Os06g0286700	Disease resistance against *M. oryzae* and *M. grisea*
BAS97557.1	Os06g0330100	Disease resistance against *M. oryzae*
XP_015643661.1	G-type lectin S-receptor-like serine/threonine-protein kinase SD2-5	Disease resistance
Q0D4Z6.1	Indole-3-acetic acid-amido synthetase GH3.8; Auxin-responsive GH3-like protein 8	Disease resistance
Q7X993.1	Transcription factor TGA2.1; bZIP transcription factor 63	Disease resistance
Q69LD8.1	Gibberellin 20 oxidase 3; GA 20-oxidase 3; Gibberellin C-20 oxidase 3	Disease resistance
Q0D715.1	Calcium-dependent protein kinase 18	Disease resistance against *M. oryzae*
XP_015646465.1	Omega-3 fatty acid desaturase, chloroplastic	Disease resistance
XP_025883131.1	Probable WRKY transcription factor 2	Disease resistance
NP_001390599.1	Bidirectional sugar transporter SWEET11	Disease resistance against *Rhizoctonia solani*
XP_015648496.1	Stromal cell-derived factor 2-like protein	Disease resistance
Q69P94.1	Protein TIFY 10c; Jasmonate ZIM domain-containing protein 8	Disease resistance
Q67UE8.1	LysM domain-containing GPI-anchored protein LYP4; LysM domain-containing protein 4	Disease resistance
XP_015611599.1	Probable serine/threonine-protein kinase PBL19	Disease resistance
Q6ES51.1	Protein TIFY 6b; Jasmonate ZIM domain-containing protein 4	Disease resistance
XP_015616838.1	Receptor kinase-like protein Xa21	Resistance against *Xoo*
NP_001391837.1	Putative cysteine-rich receptor-like protein kinase 12	Disease resistance
XP_015617551.1	Lipase-like PAD4	Disease resistance
Q0IMG9.2	E3 ubiquitin-protein ligase SPL11; Cell death-related protein SPL11; Protein spotted leaf 11; RING-type E3 ubiquitin transferase SPL11	Disease resistance
XP_015620642.1	Disease resistance protein RGA5-like isoform X1	Disease resistance against *M. oryzae*

Protein IDs: identifiers of disease-resistance proteins predicted by the DL-based MLP model; Description: details of the predicted proteins by the DL-based MLP model.

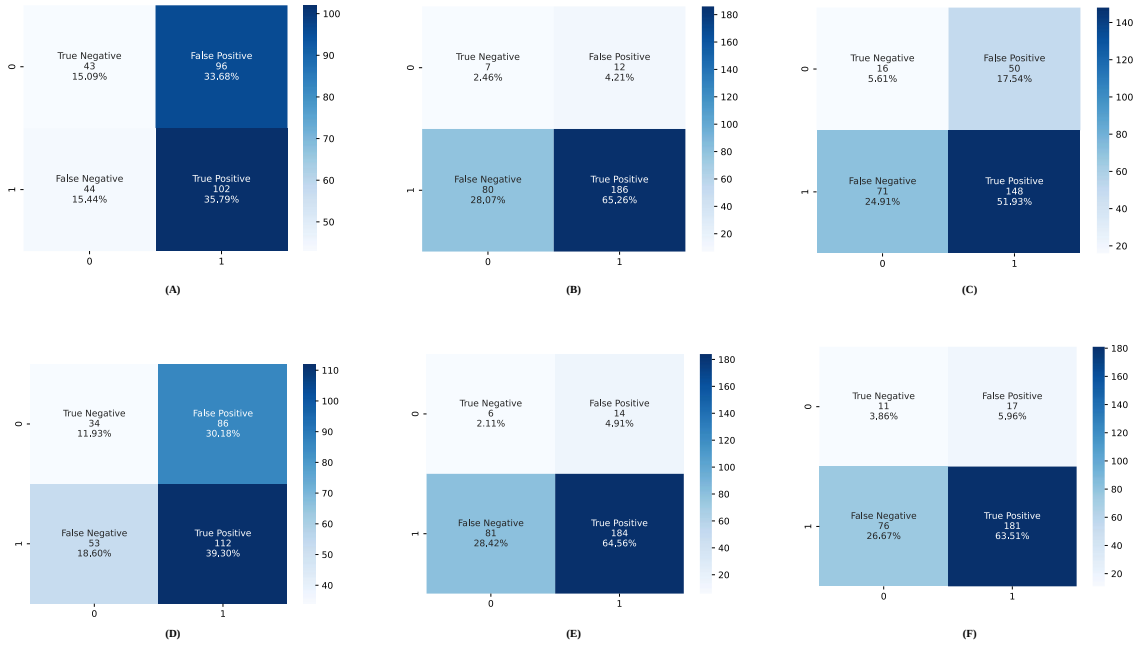


Figure B.1 : Confusion matrix of each model on the internal dataset. (A) to (F) Confusion Matrix of ML and DL models, including LR, RF, SVM, NB, KNN and MLP, respectively. The matrix was generated through the TP, TN, FP, and FN values of each model. It shows the highest TP values in the DL-based MLP model compared to other ML models (see "Results" section).

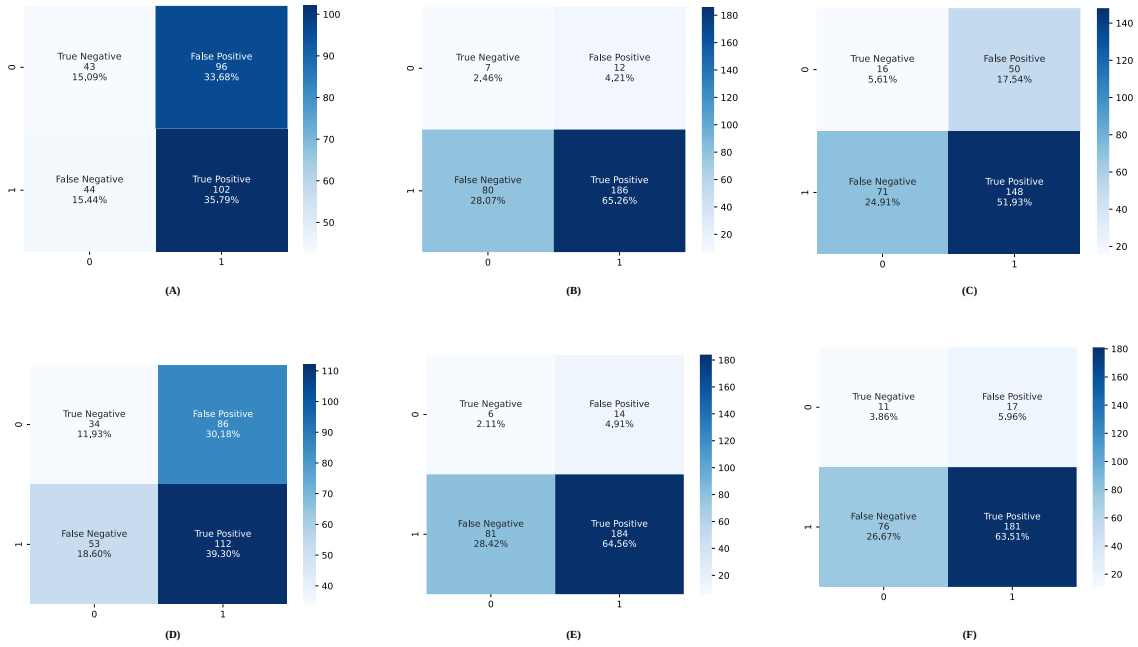


Figure B.2 : Confusion matrix of the DL-based MLP model and ML models for the external dataset. **(A)** to **(F)** A comparison between TP, TN, FP, and FN values in the confusion matrix for each of the LR, RF, SVM, NB, KNN, and MLP models. The confusion matrix was constructed for the external dataset provided to the models for evaluating their performance.

Table B.1 : List of various descriptors computed by protr package.

Descriptor groups	Descriptors
Amino acid composition	Amino acid composition
	Dipeptide composition
	Tripeptide composition
	Normalized Moreau-Broto
Autocorrelation	Moran
	Geary
	Composition
Composition, Transition, Distribution	Transition
	Distribution
Conjoint Triad	Conjoint Triad
Quasi-sequence-order	Sequence-order-coupling number
	Quasi-sequence-order descriptors
Pseudo-amino acid composition	Type I
	Type II
Proteochemometric descriptors	Principal components analysis (amino acid properties based)
	Principal components analysis (2D and 3D molecular descriptors based)
	Factor analysis (amino acid properties based)
	Factor analysis (2D and 3D molecular descriptors based)
	Multidimensional scaling (amino acid properties based)
	Multidimensional scaling (2D and 3D molecular descriptors based)
	BLOSUM and PAM matrix-derived descriptors
PSSM	PSSM-profile

Descriptor groups: Eight different types of descriptor groups for proteins computed by the protr package; Descriptors: Description of the protein features present in the protr package such as dipeptide and tripeptide composition, profile-based protein features based on the position-specific scoring matrix (PSSM), Proteochemometric modeling including BLOcks SUbstitution Matrices (BLOSUM), Point Accepted Mutation (PAM) and much more.

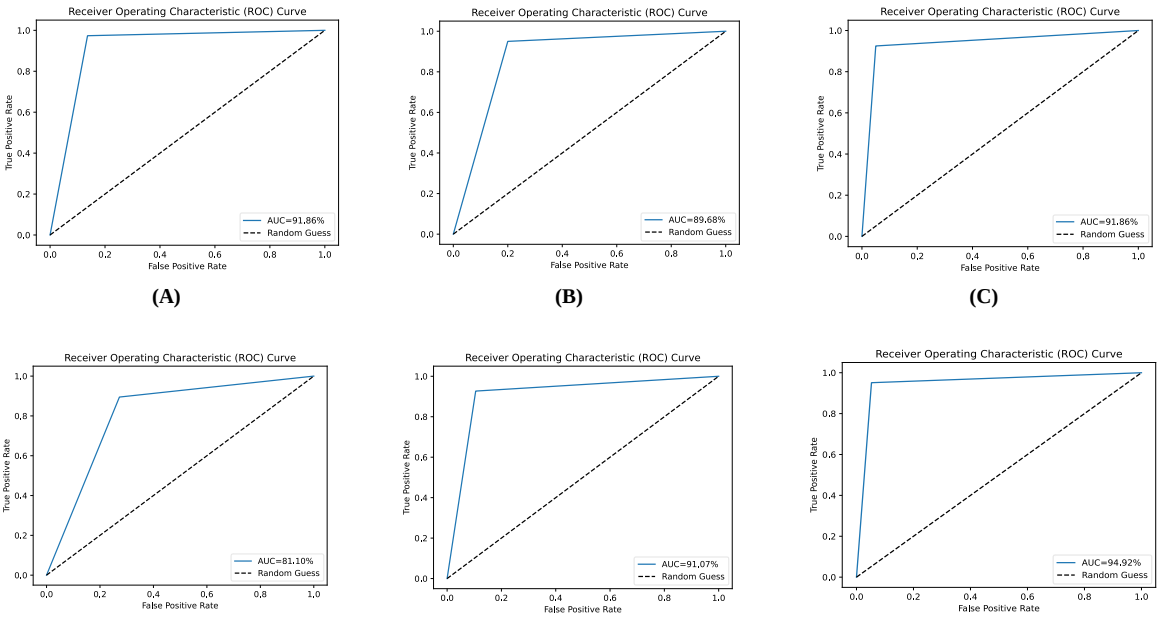


Figure B.3 : AUC curves for the external dataset for all the ML and DL models. **(A)** to **(F)** A comparison between AUC scores for all the models shows that the panel **(F)** i.e., the DL-based MLP model has the best performance. The x-axis shows the False Positive Rate (FPR) and True Positive Rate (TPR) which can be found highest in the DL-based MLP model.

Appendix C

Supplementary information for Chapter 5

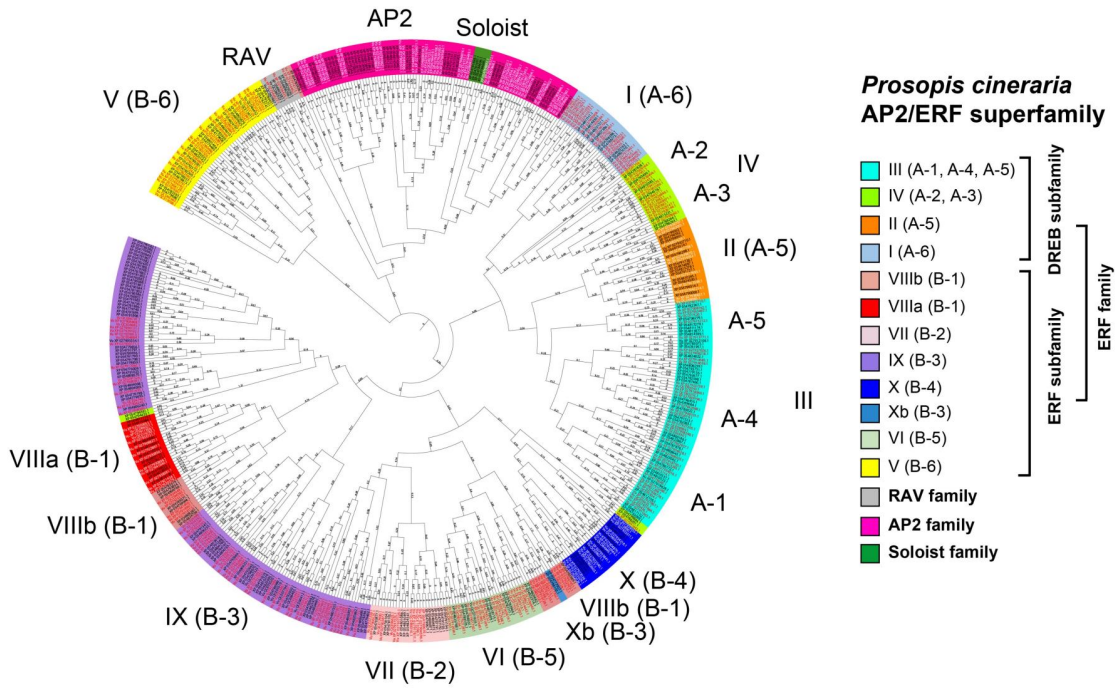


Figure C.1 : Phylogenetic tree of AP2/ERFs from *Vigna unguiculata* and *Prosopis cineraria*. The amino acid sequences of 60 differentially expressed genes of the AP2/ERF superfamily of *P. cineraria* under drought stress (see Table S1) were mined from the NCBI database. NCBI PSI-BLAST and HHblits searches were performed exhaustively on these 60 proteins, and unique sequences with AP2 domain filtered. Multiple Sequence Alignment (MSA) was performed using Clustal Omega on the 255 unique hits obtained from the NCBI BLASTP search, together with 290 *V. unguiculata* AP2/ERF superfamily proteins. The phylogenetic tree was constructed using the maximum likelihood method with the Jones-Taylor-Thornton model with 1000 bootstrap replicates in the MEGA11 application. The groups I-X were divided based on the classification of Nakano et al. (2006), whereas groups A1-A6 and B1-B6 followed the classification of Sakuma et al. (2002). *V. unguiculata* proteins are shown in red or white font.

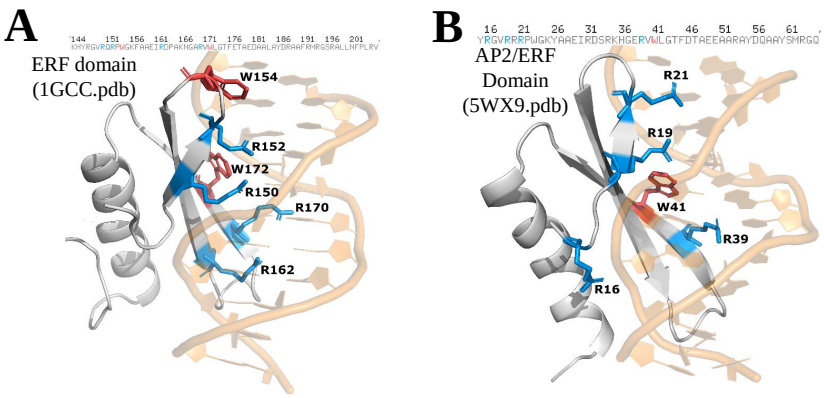


Figure C.2 : Structures of DNA-binding domain of ERF from *Arabidopsis thaliana* and *Prosopis cineraria*. (A, B) Cartoon representation of the DNA binding ERF (1GCC, NMR structure, in grey) and AP2/ERF domain (5WX9, x-ray structure, in grey) from Arabidopsis in complex with its target DNA fragment (in orange). Residues involved in specific interactions with DNA are shown in blue (positively charged) and red (aromatic) sticks.

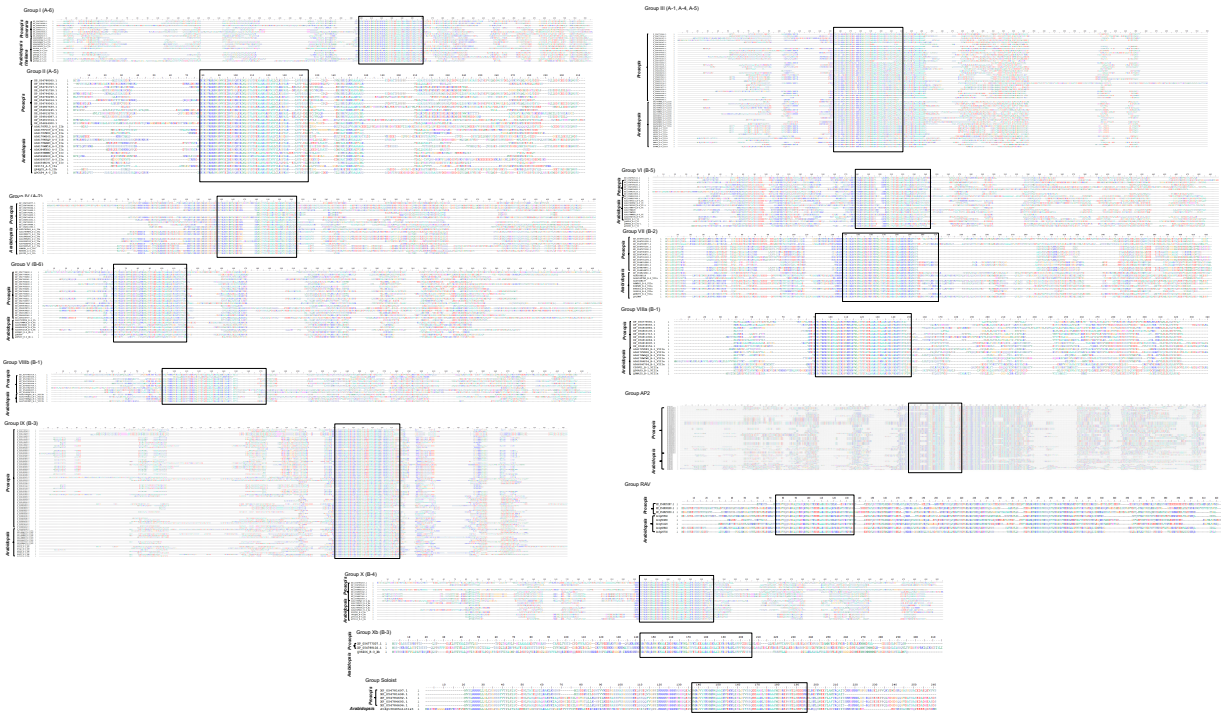


Figure C.3 : Multiple sequence alignments of AP2/ERF superfamily proteins of *Prosopis cineraria* and *Arabidopsis thaliana*. The alignments were conducted by CLUSTALW in BioEdit software, separately for each group of the phylogenetic tree (5.2). The black boxes indicate AP2 domains.

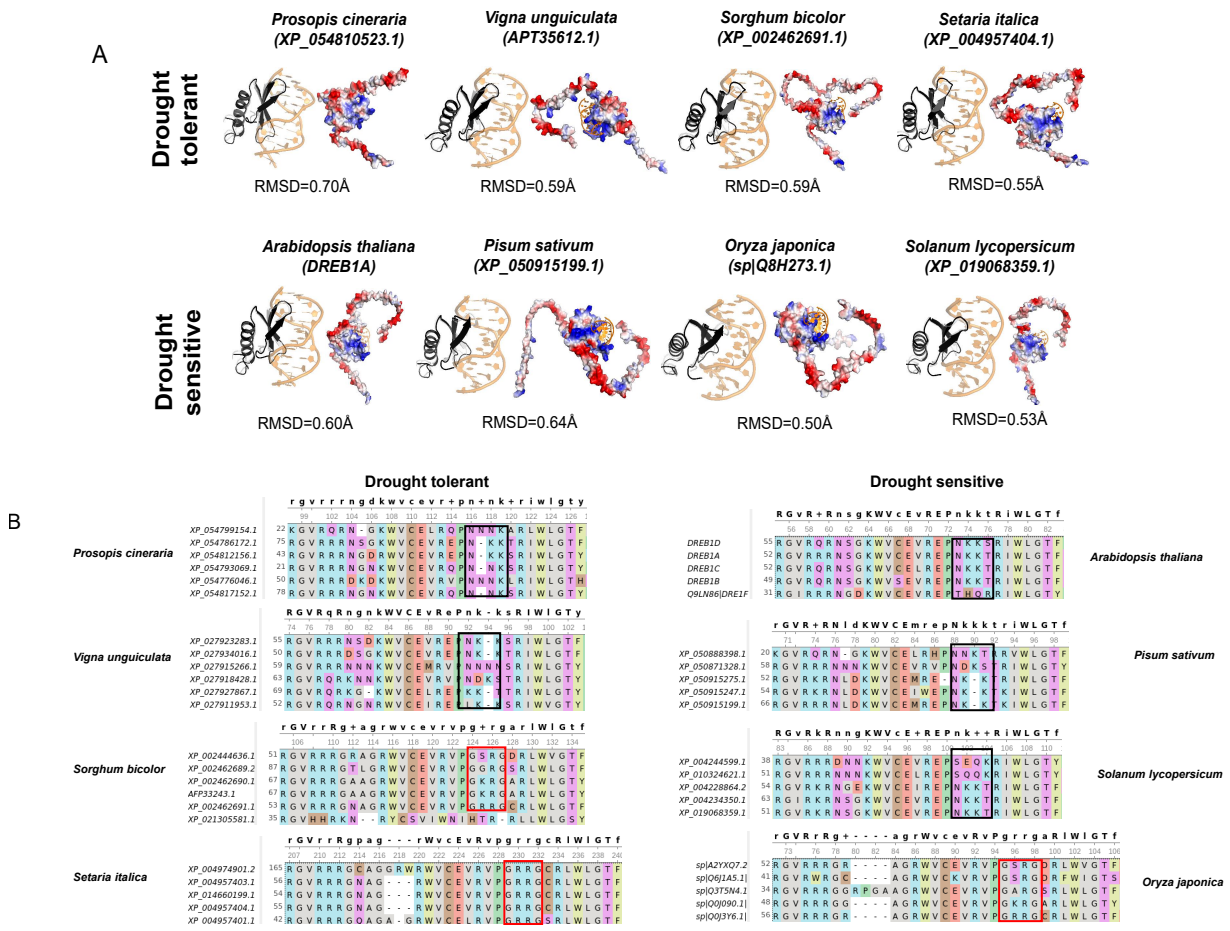


Figure C.4 : Structural comparison of DREB1 proteins from drought tolerant and sensitive species. Modeled full-length structure of dehydration-responsive element binding protein 1 (DREB1) from drought tolerant (top panel) and drought sensitive (bottom panel) species using AlphaFold2. The same representation schemes as that of Figure 1 are used. For each species, the left pane shows only the DNA-binding domain (in black cartoon) superposed on 1GCC.pdb (grey cartoon) whereas the right panel shows the electrostatic surface of the full-length protein in which blue and red correspond to the positive and negatively charged surfaces, respectively. DNA is shown in an orange cartoon. (B) Sequence alignments of DREB1 proteins from four drought-sensitive and tolerant species are shown.

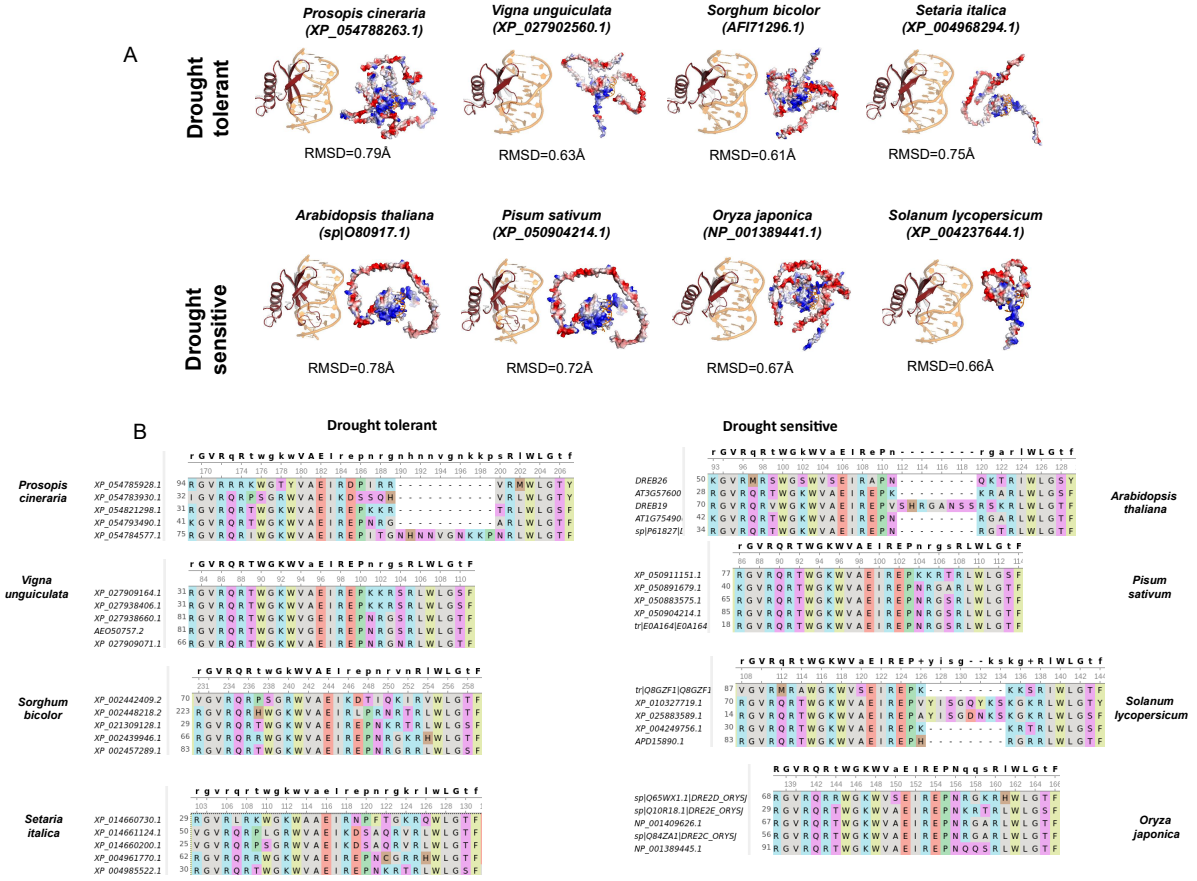


Figure C.5 : Structural comparison of DREB2 proteins from drought tolerant and sensitive species. (A) Modeled full-length structure of dehydration-responsive element binding protein 2 (DREB2) from drought tolerant (top panel) and drought sensitive (bottom panel) species using AlphaFold2. The same representation schemes as that of Figure 1 are used. For each species, the left pane shows only the DNA-binding domain (in black cartoon) superposed on 1GCC.pdb (grey cartoon) whereas the right panel shows the electrostatic surface of the full-length protein in which blue and red correspond to the positive and negatively charged surfaces, respectively. DNA is shown in an orange cartoon. (B) Sequence alignments of DREB2 proteins from four drought-sensitive and tolerant species are shown.

List of Publications

Published Journal Articles:

J1. Dhiman, V., Biswas, S., Shekhawat, R. S., Sadhukhan, A., & Yadav, P. (2024). *In silico* characterization of five novel disease-resistance proteins in *Oryza sativa* sp. *japonica* against bacterial leaf blight and rice blast diseases. 3 Biotech, 14(2), 48. https://doi.org/10.1007/s13205-023-03893-5

J2. Dhiman, V., Biswas, S., Swain, A. K., Sadhukhan, A., & Yadav, P. (2023). Deep learning-based method to identify disease-resistance proteins in *Oryza sativa* and relative species. bioRxiv, 2023-11. https://doi.org/10.1101/2023.11.25.568625

J3. Dhiman, V., Marik, D., Shekhawat, R. S., Swain, A. K., Dey, A., Yadav, P., Pal, A., Dey, S. & Sadhukhan, A. (2024). AP2/ERF Transcription Factor rthologs of the Desert Tree *Prosopis cineraria* Show Higher Copy Number and DNA-Binding Affinity than Drought-Sensitive Species. Journal of Plant Growth Regulation, 1-25. https://doi.org/10.1007/s00344-024-11532-3

Published Book Chapter:

J4. Shekhawat, R. S., Kashyap, P., Swain, A. K., Dhiman, V., Sharma, J., Kumar, A., & Yadav, P. (2024). Bioinformatics for Multi-Omics https://doi.org/10.1016/B978-0-323-95502-7.00159-7