



Editorial

Power Up Plant Genetic Research with Genomic Data

Man-Wah Li ^{1,2} , Sachiko Isobe ³ and Hon-Ming Lam ^{1,*}

¹ Centre for Soybean Research of the State Key Laboratory of Agrobiotechnology and School of Life Sciences, The Chinese University of Hong Kong, Hong Kong, China; manwah.li@yale.edu

² Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06510, USA

³ Laboratory of Plant Genetics and Genomics, Kazusa DNA Research Institute, Kisarazu 292-0828, Chiba, Japan; sisobe@kazusa.or.jp

* Correspondence: honming@cuhk.edu.hk; Tel.: +852-3943-6336

The official debut of the reference genome of *Arabidopsis thaliana* in 2000 [1] marked the genesis of the plant genome era. Yet, limited by the sequencing technology and the complexity of most plant genomes, reference-grade genomes have only been available for a few important plant species in the two decades since [2]. In the past few years, the rise of third-generation sequencing (or so-called long-read sequencing) platforms has accelerated the assembly of genomes at the chromosome level [2–4]. Up until now, it has been estimated that, out of the ~40,000 green plant species, less than 0.2% have had their genomes assembled at the chromosome level [2,3]. Despite this relatively small achievement, these published genomes have already been able to advance genome-guided research, including, but not limited to, population genomics, transcriptomics, proteomics, meta-genomics, meta-transcriptomics and epigenomics. In this Special Issue, we have collected eight such plant genomics studies from dedicated researchers.

The plant genome size ranges from several tens of megabases to hundreds of gigabases [5]. Nevertheless, the number of protein-coding genes is not strictly correlated with the genome size, as most plant genomes are predicted to have similar numbers of protein-coding genes, averaging around 30,000 [5]. For uncharacterized genomes, an easy way to identify functional genes is to search for known gene families based on sequence homology and previous research of model organisms. The sucrose nonfermenting 2 (Snf2) family of proteins have previously been identified to modulate reproductive organ development in the model plant *Arabidopsis*, making them potential candidates for affecting crop yield. Adopting the hidden Markov model (HMM)-based approach, Chen et al. identified 38 high-confidence Snf2 protein-encoding genes in the genome of barley (*Hordeum vulgare*) [6]. These Snf2 genes can be classified into 6 groups and 18 subfamilies phylogenetically. Some of these Snf2 genes were found to be differentially expressed in floral organs, supporting the notion that they are important for determining the yield of barley. Another study by Lin et al. successfully identified, in the pearl millet genome, 208 putative MYB-coding genes [7], a transcription factor family participating in diverse biological functions. Substitution rate analyses of orthologous MYB pairs suggested that they were under purifying selection during domestication. Expression analyses of 11 selected MYB genes revealed that their expressions were responsive to stress treatments, resembling the characterized MYB genes from other organisms. This catalog of MYB genes would be useful for supporting further research of this gene family in pearl millet.

Whole-genome resequencing, in theory, could detect any small genomic variations in the genome at low cost. Thus, it is a powerful tool for genotyping genetic populations for forward genetics studies. In most cases, genetic markers are no longer the limiting factor for genetic mapping; instead, the number of recombination events within the population is the limiting factor. Guan et al. resequenced 200 F2 eggplants from an interspecific cross (*Solanum melongena* × *S. incanum*) for the construction of a recombinant bin map [8]. Using the phenotypic data of 172 F2:3 lines and this recombinant bin map, a major quantitative



Citation: Li, M.-W.; Isobe, S.; Lam, H.-M. Power Up Plant Genetic Research with Genomic Data. *Int. J. Mol. Sci.* **2023**, *24*, 6876. <https://doi.org/10.3390/ijms24086876>

Received: 3 April 2023

Accepted: 5 April 2023

Published: 7 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

trait locus (QTL) was identified corresponding to the leaf vein color and fruit pericarp color. Using the single-nucleotide polymorphism (SNP) and insertion/deletion (INDEL) information from F2 resequencing, the same genomic region was also rediscovered. Similarly, Liang et al. made use of 267 maize double-haploid (DH) lines and 336 inbred lines to map the genomic loci associated with the ear shank length, a trait determining the grain yield and kernel dehydration rate [9]. Using previously available bin maps and SNP data, they successfully identified 16 QTLs and 23 SNPs associated with the ear shank length via QTL mapping and genome-wide association studies (GWAS), respectively. By overlapping the two sets of data, this group identified four loci common to all the study populations for further investigation. Through the analysis of transcriptional dynamics during maize shank elongation, the researchers narrowed down the candidate genes to a final list of 12 differentially expressed genes.

Transcriptome analyses have been adopted widely to study plant–pathogen interactions. By integrating histology, transcriptomics and metabolomics, Zhou et al. tapped into the changes in walnut (*Juglans regia* L.) upon infection by *Phomopsis capsica* [10]. They discovered that *P. capsica* can lead to severe destruction of the vasculature of the infected walnut branches. The association of differentially expressed genes with differentially expressed metabolites demonstrated that *P. capsica* infection disturbs the metabolism of carbohydrates, amino acids and some secondary metabolites.

Apart from simply examining the host response toward the pathogen, variants of transcriptomic analyses have been used to delineate bidirectional interactions. For example, the meta-transcriptome can be used to identify the pathogen's own pathogen. Fungal diseases are the major threats to rice harvest. Biological control of fungal diseases using mycoviruses can be a sustainable solution. To identify potential mycoviruses as suitable biological control agents for fungal diseases in rice, He et al. [11] sequenced the meta-transcriptome of 343 strains of fungal pathogens of rice that are potentially affected by viruses. In total, 68 mycoviruses were identified from 682 contigs, with 42 of them never described before. Associated with the stunted growth of the fungal pathogens, the mycoviruses identified in this study could potentially be new biological control agents. Another study made use of dual transcriptome analyses to study the interactions between *Colletotrichum higginsianum* and *Arabidopsis thaliana*, the former being a fungal pathogen of the latter, which is the model for cruciferous plants [12]. Autophagy has been demonstrated to be important for fungal pathogenicity. The host, *Arabidopsis*, was inoculated with either wildtype *C. higginsianum* (ChWT) or a mutant with impaired autophagy (*Chatg8Δ*). Through dual transcriptome analyses, it was shown that *Chatg8Δ* significantly affected the melanin pathway, in which a deletion in the most enriched gene encoding trihydroxynaphthalene reductase (*THR1*) could lead to the loss of pathogenicity.

A well-annotated reference genome is always an important foundation for other omics studies. Rehman et al. analyzed the salt-induced membrane proteomic changes in the roots and leaves of a salt-sensitive soybean cultivar using a label-free quantitation strategy, and identified 127 differentially expressed membrane proteins [13]. These proteins are involved in ion transport, protein transport, ATP hydrolysis, protein folding, receptor kinases, etc. Verified with previously published RNA-seq data, this study demonstrated the importance of solute transport and sensing in salt stress responses.

This Special Issue explored some new areas in omics research in plants, demonstrating the power of omics strategies and their potential applications in improving crop production. These studies collectively provide a large amount of valuable omics data and information for further investigation.

Author Contributions: Conceptualization, M.-W.L., S.I. and H.-M.L.; writing—original draft preparation, M.-W.L.; writing review and editing, M.-W.L., S.I. and H.-M.L.; supervision, H.-M.L.; project administration, H.-M.L.; funding acquisition, H.-M.L. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: We thank Jee Yan Chu for copy editing the manuscript. Any opinions, findings, conclusions or recommendations expressed in this publication do not reflect the views of the Government of the Hong Kong Special Administrative Region or the Innovation and Technology Commission.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kaul, S.; Koo, H.L.; Jenkins, J.; Rizzo, M.; Rooney, T.; Tallon, L.J.; Feldblyum, T.; Nierman, W.; Benito, M.I.; Lin, X.Y.; et al. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **2000**, *408*, 796–815. [[CrossRef](#)]
2. Marks, R.A.; Hotaling, S.; Frandsen, P.B.; VanBuren, R. Representation and participation across 20 years of plant genome sequencing. *Nat. Plants* **2021**, *7*, 1571–1578. [[CrossRef](#)] [[PubMed](#)]
3. Kress, W.J.; Soltis, D.E.; Kersey, P.J.; Wegrzyn, J.L.; Leebens-Mack, J.H.; Gostel, M.R.; Liu, X.; Soltis, P.S. Green plant genomes: What we know in an era of rapidly expanding opportunities. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2115640118. [[CrossRef](#)] [[PubMed](#)]
4. Murigneux, V.; Rai, S.K.; Furtado, A.; Bruxner, T.J.C.; Tian, W.; Harliwong, I.; Wei, H.M.; Yang, B.C.; Ye, Q.Y.; Anderson, E.; et al. Comparison of long-read methods for sequencing and assembly of a plant genome. *Gigascience* **2020**, *9*, 11. [[CrossRef](#)] [[PubMed](#)]
5. Michael, T.P. Plant genome size variation: Bloating and purging DNA. *Brief Funct. Genomics* **2014**, *13*, 308–317. [[CrossRef](#)] [[PubMed](#)]
6. Chen, G.; Mishina, K.; Zhu, H.J.; Kikuchi, S.; Sassa, H.; Oono, Y.; Komatsuda, T. Genome-Wide Analysis of *Snf2* Gene Family Reveals Potential Role in Regulation of Spike Development in Barley. *Int. J. Mol. Sci.* **2023**, *24*, 457. [[CrossRef](#)] [[PubMed](#)]
7. Lin, M.H.; Dong, Z.Y.; Zhou, H.K.; Wu, G.Y.; Xu, L.; Ying, S.; Chen, M. Genome-Wide Identification and Transcriptional Analysis of the MYB Gene Family in Pearl Millet (*Pennisetum glaucum*). *Int. J. Mol. Sci.* **2023**, *24*, 2484. [[CrossRef](#)] [[PubMed](#)]
8. Guan, W.X.; Ke, C.J.; Tang, W.Q.; Jiang, J.L.; Xia, J.; Xie, X.F.; Yang, M.; Duan, C.F.; Wu, W.R.; Zheng, Y. Construction of a High-Density Recombination Bin-Based Genetic Map Facilitates High-Resolution Mapping of a Major QTL Underlying Anthocyanin Pigmentation in Eggplant. *Int. J. Mol. Sci.* **2022**, *23*, 10258. [[CrossRef](#)] [[PubMed](#)]
9. Liang, Z.; Xi, N.; Liu, H.; Liu, P.; Xiang, C.; Zhang, C.; Zou, C.; Cheng, X.; Yu, H.; Zhang, M.; et al. An Integration of Linkage Mapping and GWAS Reveals the Key Genes for Ear Shank Length in Maize. *Int. J. Mol. Sci.* **2022**, *23*, 15073. [[CrossRef](#)] [[PubMed](#)]
10. Zhou, L.; Zhu, T.; Han, S.; Li, S.; Liu, Y.; Lin, T.; Qiao, T. Changes in the Histology of Walnut (*Juglans regia* L.) Infected with *Phomopsis capsici* and Transcriptome and Metabolome Analysis. *Int. J. Mol. Sci.* **2023**, *24*, 4879. [[CrossRef](#)] [[PubMed](#)]
11. He, Z.; Huang, X.; Fan, Y.; Yang, M.; Zhou, E. Metatranscriptomic Analysis Reveals Rich Mycoviral Diversity in Three Major Fungal Pathogens of Rice. *Int. J. Mol. Sci.* **2022**, *23*, 9192. [[CrossRef](#)] [[PubMed](#)]
12. Zhu, Y.; Duan, L.; Zhu, C.; Wang, L.; He, Z.; Yang, M.; Zhou, E. Dual Transcriptome Analysis Reveals That *ChATG8* Is Required for Fungal Development, Melanization and Pathogenicity during the Interaction between *Colletotrichum higginsianum* and *Arabidopsis thaliana*. *Int. J. Mol. Sci.* **2023**, *24*, 4376. [[CrossRef](#)] [[PubMed](#)]
13. Rehman, H.M.; Chen, S.; Zhang, S.; Khalid, M.; Uzair, M.; Wilmarth, P.A.; Ahmad, S.; Lam, H.-M. Membrane Proteomic Profiling of Soybean Leaf and Root Tissues Uncovers Salt-Stress-Responsive Membrane Proteins. *Int. J. Mol. Sci.* **2022**, *23*, 13270. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.