

EDITORIAL

Open Access



Medicinal plant genomics

Christian Siadjeu¹ and Boas Pucker^{2*}

Abstract

Recent developments in plant genomics have enabled a comprehensive analysis of the medicinal potential of plants based on their gene repertoire. Genes of biosynthesis pathways can be discovered through comparative genomics and through integration of transcriptomic data. Data-driven discovery of specialized metabolites could accelerate research.

The emergence and rapid development of long read sequencing technologies has been revolutionizing the field of genomics by substantially increasing assembly continuity [1]. An ongoing competition between Oxford Nanopore Technologies and Pacific Biosciences has led to rapid and continuous releases of innovations. Furthermore, the availability of portable and affordable sequencing instruments, as well as reduced sequencing costs, has democratized genomics and substantially increased the number of scientists who are applying genomic approaches and thus led to an increase in genome sequencing projects [2]. Now, studies are not only limited to the genomes of economically important crops, but also include the genomes of medicinal plants. Additional large omics data sets are often analyzed in an integrated way based on sophisticated bioinformatics tools. Indeed, a genomic sequence can advance projects that are aiming to unravel a certain biosynthesis pathway. While the generation of an almost complete genome sequence based on highly accurate long reads is becoming a routine task, the generation of the corresponding high-quality structural annotation is a growing challenge. RNA-seq and

direct RNA sequencing can provide external hints to support this structural annotation process but revealing the functions of genes has become a bottleneck on the route to fully exploiting the potential of plants.

Besides food production and ecological roles, plants are also amazing chemists that synthesize a plethora of specialized metabolites. As recently reviewed, the precise number of different metabolites produced by plants is unknown, but estimates go up to about one million [3]. Many of these metabolites are of high value for humans. For example, plant pigments like anthocyanins, betalains, and carotenoids can be used as food colorants [4]. In addition to a nice visual appearance, important nutritional roles or even medicinal properties have been reported for these compounds. Flavonoids, including anthocyanins, are well known for their antioxidant properties which can protect various cell components against reactive oxygen species under stress conditions. Consequently, an anthocyanin-rich diet might help to prevent the onset or development of certain diseases. As flavonoids are present across land plants, there is a huge diversity of different derivatives that occur in specific evolutionary lineages. Specific types of flavonoids have already been considered as drug candidates. For example, the biosynthesis of the α -amylase inhibitor montbretin A in *Crocoshmia x crocosmiiflora* was elucidated and has thus enabled its heterologous production [5].

Alkaloids, such as colchicine, have long been used to both prevent and treat gout attacks, a severe inflammation linked to arthritis. Despite having been utilized for more than a thousand years, the specific mechanism

*Correspondence:

Boas Pucker
b.pucker@tu-braunschweig.de

¹ Prinzessin Therese Von Bayern Chair of Systematics, Biodiversity and Evolution of Plants, Ludwig Maximilian University Munich, 80638 Munich, Germany

² Plant Biotechnology and Bioinformatics, Institute of Plant Biology & BRICS, TU Braunschweig, 38106 Braunschweig, Germany



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

of colchicine production is still not fully understood. Recently, a draft genome sequence of *Cremastra appendiculata* (D. Don) has facilitated the identification of candidate genes associated with colchicine biosynthesis, highlighting the predominant role of an O-methyltransferase [6].

Similarly, terpenes are bioactive compounds with significant medicinal properties. Artemisinin is an endoperoxide sesquiterpene lactone produced in the glandular trichomes of *Artemisia annua*. Genomic sequencing, combined with transcriptome analysis, has revealed novel genes involved in artemisinin biosynthesis, and demonstrated that overexpressing the transcription factor *AaMYB2* can significantly enhance the content of artemisinin and dihydroartemisinic acid in transgenic *A. annua* lines [7]. Particularly relevant in the context of specialized metabolites in plants are biosynthetic gene clusters that gained attention in recent years [8]. The products synthesized by the enzymes encoded by these biosynthetic gene clusters are often associated with responses against pathogens.

A high-quality genome sequence with an accurate structural annotation can serve as the basis for the identification of biosynthesis genes and information about the individual functions of genes can be transferred from other plant species based on the assumption that orthologs have the same functions. This enables a high-throughput annotation workflow. Researchers can thus access a wealth of information about individual genes through genome browsers which facilitates the search for promising plant enzymes. Additional hints for the identification of biosynthesis genes can also be derived from co-expression data. As biosynthesis genes in the same pathway are often expressed simultaneously, various types of co-expression analyses have emerged as a powerful approach towards the identification of candidate genes. Thus based on one known gene in a biosynthesis pathway, additional genes in the same pathway can be identified. These simple correlation analyses of gene expression values have been replaced by more sophisticated machine learning algorithms [9]. Indeed information about gene expression can be integrated with metabolomics and proteomics data in a multiomics approach. Gene expression can be linked to the presence of metabolites of interest to identify the genes responsible for the corresponding biosynthesis pathway. Once completely resolved, plant biosynthesis pathways will be useful resources to be harnessed for heterologous production in other systems.

With this collection we hope to attract submissions that cover a wide range of topics, including genome sequencing, transcriptomics, comparative genomics, and bioinformatics analyses of medicinal plants.

Acknowledgements

Many thanks to the Plant Biotechnology and Bioinformatics research group at TU Braunschweig for providing comments on the initial draft of this manuscript.

Authors' contributions

BP initialized the Collection and wrote the initial draft. CS and BP revised the manuscript. CS and BP read the final version and approved its publication in BMC Genomics.

Funding

No specific funding was received for this work.

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

CS and BP are guest editors of this Collection. BP is a member of the editorial board of BMC Genomics and head of the technology transfer center Plant Genomics and Applied Bioinformatics at iTUBS.

Received: 3 July 2023 Accepted: 28 July 2023

Published online: 01 August 2023

References

1. Marks RA, Hotaling S, Frandsen PB, VanBuren R. Representation and participation across 20 years of plant genome sequencing. *Nat Plants*. 2021;7:1571–8.
2. Pucker B, Irisarri I, de Vries J, Xu B. Plant genome sequence assembly in the era of long reads: Progress, challenges and future directions. *Quant Plant Biol*. 2022;3: e5.
3. Wang S, Alseekh S, Fernie AR, Luo J. The Structure and Function of Major Plant Metabolite Modifications. *Mol Plant*. 2019;12:899–919.
4. Tanaka Y, Sasaki N, Ohmiya A. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J*. 2008;54:733–49.
5. Irmisch S, Jancsik S, Man Saint Yuen M, Madilao LL, Bohlmann J. Complete Biosynthesis of the Anti-Diabetic Plant Metabolite Montbretin A. *Plant Physiol*. 2020;184:97–109.
6. Wang J, Xie J, Chen H, Qiu X, Cui H, Liu Y, et al. A draft genome of the medicinal plant *Cremastra appendiculata* (D. Don) provides insights into the colchicine biosynthetic pathway. *Commun Biol*. 2022;5:1–12.
7. Shen Q, Zhang L, Liao Z, Wang S, Yan T, Shi P, et al. The Genome of *Artemisia annua* Provides Insight into the Evolution of Asteraceae Family and Artemisinin Biosynthesis. *Mol Plant*. 2018;11:776–88.
8. Polturak G, Osbourn A. The emerging role of biosynthetic gene clusters in plant defense and plant interactions. *PLOS Pathog*. 2021;17: e1009698.
9. Delli-Ponti R, Shivhare D, Mutwil M. Using Gene Expression to Study Specialized Metabolism—A Practical Guide. *Front Plant Sci*. 2021;11:625035. <https://doi.org/10.3389/fpls.2020.625035>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.