Review

The Current Developments in Medicinal Plant Genomics Enabled the Diversification of Secondary Metabolites' Biosynthesis

Mohammad Murtaza Alami [ID], Zhen Ouyang, Yipeng Zhang, Shaohua Shu [ID], Guozheng Yang, Zhinan Mei and Xuekui Wang *

College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China
* Correspondence: wang-xuekui@mail.hzau.edu.cn

Abstract: Medicinal plants produce important substrates for their adaptation and defenses against environmental factors and, at the same time, are used for traditional medicine and industrial additives. Plants have relatively little in the way of secondary metabolites via biosynthesis. Recently, the whole-genome sequencing of medicinal plants and the identification of secondary metabolite production were revolutionized by the rapid development and cheap cost of sequencing technology. Advances in functional genomics, such as transcriptomics, proteomics, and metabolomics, pave the way for discoveries in secondary metabolites and related key genes. The multi-omics approaches can offer tremendous insight into the variety, distribution, and development of biosynthetic gene clusters (BGCs). Although many reviews have reported on the plant and medicinal plant genome, chemistry, and pharmacology, there is no review giving a comprehensive report about the medicinal plant genome and multi-omics approaches to study the biosynthesis pathway of secondary metabolites. Here, we introduce the medicinal plant genome and the application of multi-omics tools for identifying genes related to the biosynthesis pathway of secondary metabolites. Moreover, we explore comparative genomics and polyploidy for gene family analysis in medicinal plants. This study promotes medicinal plant genomics, which contributes to the biosynthesis and screening of plant substrates and plant-based drugs and prompts the research efficiency of traditional medicine.

Keywords: medicinal plants; genomics; multi-omics; secondary metabolites; biosynthetic pathways

1. Introduction

Medicinal plants produce thousands of secondary metabolites that have been used in traditional medicine for centuries and provide essential elements for industries, such as drugs, condiments, food additives, and essential oils. However, secondary metabolites synthesis in a low concentration in plants, the genetics of biosynthetic pathways, and the regulation of these pathways must be studied to overcome the low synthesis of various secondary metabolites in plant cells. The high-quality genome sequencing paired with metabolomics and proteomics increased our understanding of biosynthesis pathways toward expanding production of metabolites.

The progress of technology has played a crucial role in the development of plant genomics. Every technological innovation has dramatically promoted the advancement of genomics. Since the first-generation sequencing technology was invented in the 1970s, in just 50 years, sequencing technology has experienced three generations of innovation (Supplementary Figure S1). It is still booming and changing with each passing day [1]. The development of genome sequencing, assembly technologies, and supporting bioinformatics analysis tools has injected additional vitality into the comprehensive development of medicinal plant genomics and molecular biology, making it possible to obtain high-quality medicinal plant genomes. A high-quality genome sequence will help to carry out extensive and in-depth molecular genetics research and play a significant role in identifying genes related to the biosynthesis of secondary metabolites in medicinal plants and clarifying the

historical evolution status. The completion of whole-genome sequencing has laid a solid foundation for carrying out fundamental research on medicinal plants' molecular breeding and accelerating the cultivation process of varieties. In addition, by employing the genetic data and regulatory network of the species and the omics technologies in accordance, the genomics research of medicinal plants aims to expose their influence on the human body from the level of the genome and understand their molecular mechanism to prevent human illnesses. Recently, most newly published medicinal plant genomes have used third-generation sequencing technology combined with second-generation short reads and high-throughput chromosome conformation capture (Hi-C) sequencing, used, for example, with *Coptis chinensis* Franch [2], *Tripterygium wilfordii* [3], *Sarcophaga peregrina* [4], *Salvia miltiorrhiza* Bunge [5], *Aquilegia oxysepala* var. kansuensis [6], *Ophiorrhiza pumila* [7], tea [8], and many others, as shown in Supplementary Tables S1–S3.

The acquisition of genome sequence facilitates comparative genomics study and provides valuable resources for analyzing secondary plant metabolites and their synthesis; it has brought new insights into the study of medicinal plants' diversity and evolution. Previous reviews have reported on the sequencing of medicinal plants' genomes, such as Cheng et al. (2021) [9], who reviewed the plant genome situation, sequencing technology development, and medicinal plant genome application. Hamilton and Robin Buell (2012) [10] summarized the status of the sequenced plant genome, Chen et al. (2018) [11] the angiosperm plant genome, and Jiao and Schneeberger (2017) [12] the impact of long-read genome sequencing (third-generation sequencing) technology on plants' genomes. In recent years, the reduced cost and time enabled the sequence of the medicinal plant genome and found the biosynthesis pathways of many secondary metabolites. Yet, the sequencing of medicinal plants' genomes is far behind that of microorganisms, crops, and animals; the applications of high-quality genome assembly to find the biosynthesis pathways of secondary metabolites and related key genes still have not been fully elucidated, and the studies to uncover them need to be strengthened.

In this review, we conduct a systematic review of the medicinal plant genome and evolution. Moreover, we explore the development of multi-omics (genomics, transcriptomics, and metabolomics) approaches for discovering key genes related to secondary metabolite biosynthesis in medicinal plants. This study could be foundational information for biomedicine and gene identification and significantly accelerate the discovery of genes and metabolic engineering.

2. Comparative Genomics and Evolutionary Analysis Allowed Us to Identify the Genes in Secondary Metabolites' Biosynthesis

An important subfield of genomics is comparative genomics. Biological research contrasts the structural and genomic properties of the genomes to comprehend the genetic linkages, mechanisms of expression, and functions of various species [13]. Similarity and homology are the basic principles of comparative genomics [14]. Similarity only obtains the degree of similarities and differences between different species through comparing simple nucleotide or amino acid sequences and does not involve evolutionary origin, kinship, and the relationship between structure and function. Homology is the exact evolutionary origin of different species, which shows the identity of all or functionally important conserved domains of nucleotide or amino acid sequences. The essence of similarity is the quantitative relationship of similarity between genome sequences of different species, and homology is the evolutionary common-source relationship of genes among species inferred from additional biological data and statistical analysis. Homology can be further divided into orthology and paralogy [15]. Corresponding linear homologous genes refer to homologous genes distributed in the genomes of different species and derived from speciation events, originated from an ancestor in an evolutionary relationship, and inherited vertically. Although there may be variation among homologous genes during species differentiation, they are highly conservative in structure and function and maintain the similar functions of ancestral genes, and based on such characteristics, the phylogenetic tree constructed by

lineal homologous genes can reveal the evolutionary relationship between species. Collateral homologous genes are formed by lateral transmission due to the doubling of ancestral genes in the same species. They belong to the same gene. After continuous replication, the genes formed by replication enter similar but different evolutionary pathways due to mutation. Therefore, new gene functions are formed in the process of continuous variation and adaptation.

Based on the genetic link between the studied species, comparative genomics may be further subdivided into interspecific and intraspecific comparative genomics [16]. The typical species of comparative genomics among medicinal plants are *C. chinensis* and four additional species of the Ranunculales; Ranunculales had 42 gene families and 352 genes that were unique to the order. Functional annotation revealed that these gene family members were engaged in the metabolism of bisbenzylisoquinoline (BIA) alkaloids, suggesting that they could contribute to the Ranunculales' production of common precursors for the subsequent biosynthesis of other BIA alkaloids [2]. Additionally, comparative genomic analysis between *Scutellaria aicalensis* and *Scutellaria barbata* revealed that the current long terminal repeat (LTR) might result in chromosomal extension and rearrangement. Tandem duplication of paralogs after their speciation may have led to the varied development of flavonoid biosynthesis gene families, which provided a crucial starting point for research on the evolution and chemo-diversity of the Lamiaceae [17]. Numerous research has shown how comparative genomics plays a key role in exposing the traits of complex genomes (polyploidy, high heterozygosity, and high duplication) as well as the genesis and evolution of species [18–20]. With the launch of genome sequencing projects for more and more species, such as the 5000 insect genome (i5k) project [21], 10,000 vertebrate genome (g10kcos) project [22], 10,000 plant genome project (10 kp) [23], 10,000 bird genome (b10k) project [24], and earth biological genome project (EBP) [25], and the release of the high-quality genome, it is possible to study comparative genomics across more species and even the whole-order species. We can explore the evolutionary relationship of species through a more macro perspective, find the source of life explosion in the process of earth evolution, and connect the bridge between genes and phenotypes and products.

We can now fully use comparative and phylogenetic techniques for gene identification due to the increasing number of genomic sequences that are now accessible. The evolutionary links of critical metabolic enzymes were previously studied using single genes, and only a small quantity of sequencing data from a few species was employed. As a consequence of advancements in next generation sequencing (NGS) technology, more detailed examinations into the mechanisms by which metabolic pathways developed are now possible for primitive angiosperm taxa such as Amborella, gymnosperms, lycopods, and mosses as well as many species within a taxonomic family [26]. Owing to the availability of many genomes from the same or closely related taxonomic groups, it is now possible to identify orthologous across species and taxa in a biosynthetic pathway and forecast the appearance of a feature due to synteny (same gene order), which shows common evolution [27]. Due to a large number of genome sequences within particular clades and the variety of genome sequences throughout the plant kingdom, comparative analysis across and within species may be a potent tool for discovering the evolutionary origins of biochemical processes.

3. Genome Duplication Contributed to the Biosynthesis of Secondary Metabolites in Medicinal Plants

The genetic material is doubled after whole genome duplication (WGD) or polyploidization of the plant genome, and the duplicated genes are differentiated in distinct ways. This process might result in the plant genome producing novel gene functions, metabolic pathways, and phenotypic traits. The ability to identify the genes involved when combined with comparative genome and gene family analysis techniques may increase breeding options while preserving species variety. The differentiation of gene function brought on by polyploidy may aid in developing novel species-specific features. According to studies, the production of several essential elements is strongly tied to genomic poly-

ploidy events in the Cruciferae, Gramineae, Leguminosae, Compositae, Solanaceae, and angiosperms [28]. The most concrete proof is that many essential features are genetically regulated by a multi-copy gene regulatory network created by the genome polyploidy event. Close to the time of the family-specific genome polyploidy event is when a species or trait becomes unique. For instance, tea genome sequencing indicates that there have been two WGD events, the most recent of which took place around 30–40 million years ago (MYA), giving rise to catechins and caffeine, the key secondary metabolites responsible for the taste of tea (the primary bitter substance) [29].

Tandem duplication has traditionally been considered a required method by which plants increase their accumulation of secondary metabolites. For instance, the biosynthetic pathways for morphine [30] and caffeine [31] in poppy and tea plants underwent significant tandem gene duplication events that expanded their respective gene networks. Studies have also shown that WGD contributes to the manufacture of secondary metabolites [32], which is another vital element in the development of stress tolerance in plants [33,34]. Plant genomes are often complicated due to cyclical polyploidization and genome rearrangement, making it challenging to comprehend how they formed and evolved and discover novel gene functions [35].

Genes involved in biosynthesis had a significant increase in copy number. After conducting comparative genomic research with 10 sample plants, it was discovered that the tea genome considerably increased the copy number of genes synthesizing terpenes and other compounds associated with tea taste. It resulted in a more pronounced increase in tea's scent [36]. As a result, occurrences of genomic polyploidy are essential for plant diversification, differentiation, and the establishment of specific characteristics. For instance, the genomic rearrangement of the opium poppy (*Papaver somniferum*) [30] led to the emergence of BIAs alkaloid metabolism. The antitussive and anticancer drug noscapine, which is a member of the phthalideisoquinoline subclass of BIAs [37], is produced by a cluster of 10 genes that code for enzymes and a P450 oxidoreductase gene fusion [38–40] and is in charge of the crucial gateway reaction that shifts metabolites in favor of the morphinan branch and away from the noscapine branch. Wang et al. (2018) [32] reported that polyploidy contributed to the expansion of key functional genes, e.g., vitamin C biosynthesis genes in kiwifruit (*Actinidia chinensis*).

All angiosperms sequenced so far exhibit evidence of whole-genome duplication, which causes the increase of gene families and metabolic variety across evolutionary time. These complicate the studies on plant metabolism, particularly secondary metabolism. Segmental and tandem duplications and other partial genome duplications also contribute to the enormous gene family size seen in plant genomes. Genes that have been duplicated may undergo "neofunctionalization", "sub-functionalization", or even "pseudogenization", which is the division of functions into duplicates [41]. Compared to primary metabolism genes, secondary metabolism genes have the following characteristics: a propensity to be preserved after gene duplication, lineage specificity, and physical grouping within the genome [42].

4. Omics Tools for the Study of Secondary Metabolites

Research has focused greatly on how medicinal plants produce and manage their bioactive components. In the first two decades of the 21st century, we saw a revolution in genomic technologies, and these developments continue to impact all facets of plant biology, including plant metabolism. This is because access to the genome sequence and the large data sets associated with it, such as expression and metabolite profiles, can facilitate the identification of the biosynthetic and regulatory pathways for both distinct metabolites and broad classes of metabolites. The study of "omics" has opened new avenues for examining a specific medicinal plant's biosynthetic process (Figure 1). Creating de novo genome and transcriptome sequences, which have proven very informative across many species for understanding plant metabolism, is one apparent use of genome sequencing [43]. Many plant species now have genome sequences that may be utilized as a starting point

for secondary metabolite pathway identification due to recent advancements in assembly methods and annotation software as well as access to more powerful computing resources and genome technologies [44].
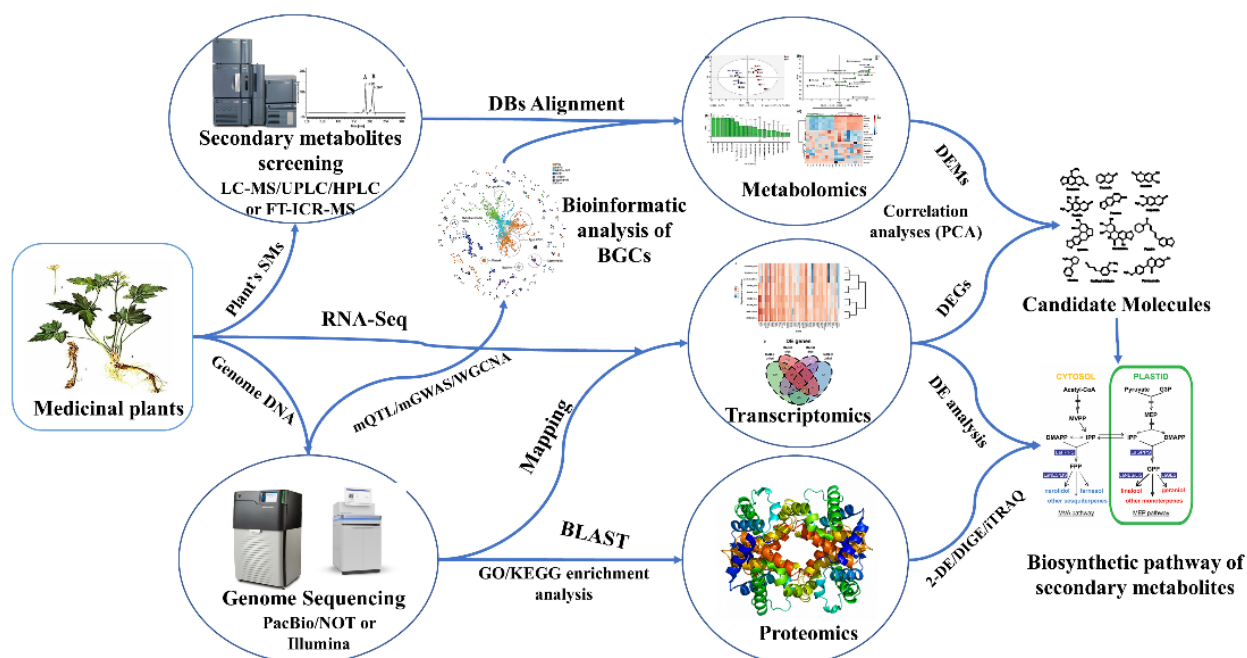

Figure 1. The multi-omics strategy and its applications. Numerous properties of a particular molecular type may be simultaneously measured in biological samples using various omics approaches. Data integration, including bioinformatic tools, pipelines, and databases, is crucial in identifying secondary metabolite biosynthesis pathways.

The secondary metabolites biosynthesis pathway analysis will be significantly accelerated by using bioinformatics and functional genomics approaches to screen and identify enzyme-coding genes on particular secondary biosynthesis pathways from many of the original medicinal plant species. For instance, Liu et al. (2021) [2] published high-quality chromosome-scale genome assembly, annotation, and metabolomics for *Coptis chinensis* Franch. by the integrated approach of omics tools; they discovered that the (S)-canadine synthase enzyme, which is involved in the berberine biosynthesis pathway, is encoded by the *CYP719* gene. The cytochrome P450 (*CYP728B70*) gene has been shown in similar studies to catalyze the oxidation of a methyl to the acid moiety of dehydroabietic acid in the triptolide biosynthesis [3]. According to research on the *Salvia miltiorrhiza* tanshinones biosynthetic route, *CYP71D373* and *CYP71D375* catalyze the hydroxylation at carbon-16 (C16) and the 14,16-ether (hetero)cyclization to create the D-ring, whereas *CYP71D411* catalyzes the upstream hydroxylation at C20 [45]. In addition, camptothecin biosynthesis was reported in *Camptotheca acuminata* by Kang et al. (2021) [46]. It is crucial to make the most of genetic data to determine the biosynthetic pathways through which medicinal plants create their active components. Candidate genes in these pathways may then be employed in synthetic biology for heterologous bioproduction. Table 1 displays the key related genes for various recently identified metabolites.

Table 1. Third-generation sequencing technology for studying secondary metabolites biosynthesis pathway.

Medicinal Plants	Sequencing Technologies	Identified Genes	Biosynthetic Pathways	Reference
Coptis chinensis Franch.	PacBio, O.N.T., and Hi-C	Cytochrome P450 (*CYP719A*)	BIAs	[2]
Tripterygiumwilfordii	PacBio, 10X genomics, and Hi-C	Cytochrome P450 (*CYP728B70*)	Triptolide	[3]
Artemisia annua	PacBio, Roche 454, and Illumina	*ADS*, *CYP71AV1*, and *FPS*	Artemisinin	[47]
Piper nigrum	PacBio, 10X Genomics, O.N.T., and HI-C	*GTF*, *CYP*, and *HCT*	Piperine	[48]
Allium sativum	PacBio, Illumina, and 10X Genomics	*AsGSH1b*, *AsGSH2*, *AsPCS1*, *AsFMO1*, and *AsGGT2*	Allicin	[49]
Magnolia biondii Pamp.	PacBio, 10X genomics, and Hi-C	*TPS* gene family	Terpenoids	[50]
Camptotheca acuminata	PacBio, Illumina, and Hi-C	*LAMT* and *SLAS*	Camptothecin	[46]
Platycodon grandiflorus	Illumina HiSeq 2500	*CYP716* and *bASs*	Platycoside	[51]
Ophiorrhiza pumila	ONT, Illumina, and Hi-C	*ASO/PAS*, *PNAE*, *PR*, *RH11H*, *SBE*, *SGD*, *T19AT*, and *THAS*	Camptothecin	[7]
Strobilanthes cusia	BioNano, Illumina, and Hi-C	*CYP*, *FMO*, *UGT*, and *BGL* gene family	Indigo	[52]
Scutellaria baicalensis	PacBio, Illumina, and Hi-C	*CHS-2*, *FNSII-2*, *F8H*, and *PFOMT5*	Wogonin	[53]
Scutellaria baicalensis and *Scutellaria barbata*	PacBio, ONT, and Hi-C	*PAL*, *4CL*, *CHS*, *F6H*, and *F8H*	Flavonoid	[17]
Chiococca alba	Illumina HiSeq 4000 and 10X Genomics	*TPSs* gene family	Unusual terpenoids	[54]
Acer truncatum	PacBio, Illumina, 10X Genomics, and Hi-C	*KCS*	Nervonic acid	[55]
Opium poppy	PacBio, Illumina, 10X Genomics, and ONT	*CYP 450* and *STORR*	Morphinan	[30]
Senna tora	PacBio, Illumina, and Hi-C	*CHS-L*	Anthraquinone	[56]
Hypericum perforatum	PacBio, Illumina, and 10X Genomics	*HpASMT1*, *HpASMT2*	Melatonin	[57]
Vernicia fordii Hemsl.	Illumina HiSeq 2000	*FAD2*, *FADX*	Triacylglycerol	[58]
Andrographis paniculata	PacBio, Illumina, and Hi-C	*diTPSs*, *CYP450*, *2OGDs*, and *UGT*	Diterpenoid neoan-drographolide	[59]
Eucommia ulmoides	PacBio, O.N.T., and Illumina	*FPSs*	Polyisoprene	[60]
Gardenia jasminoides	ONT and Hi-C	*GjCCD4a*, *ALDH*, and *UGT*	caffeine and crocin	[61]
Salvia bowleyana	PacBio, Illumina, and Hi-C	*SbPAL1*	Salvianolic acid B	[62]
Chimonanthus salicifolius	Illumina, PacBio, 10X Genomics, and Hi-C	*CHS* and *FLS*	Flavonoid	[63]
Strobilanthes cusia	PacBio, Illumina, and Hi-C	*UGT*, *IGPS*, *CYP450*, *EPSPS*, and *CS*	Indole alkaloids	[64]
Salvia miltiorrhiza	PacBio, Illumina, and Hi-C	*CYP71D373*, *CYP71D375*, and *CYP71D411*	Tanshinones	[45]
Rubus chingii Hu	ONT, Illumina, and Hi-C	*C.X.E.*, *U.G.T.*, and *SCPL*	Hydrolyzable tannin	[65]
Tripterygiumwilfordii	ONT, Illumina, and Hi-C	*TwCYP712K1* and *TwCYP712K2*	Celastrol	[66]
Papaver	ONT, Illumina, and Hi-C	*BIA gene*	Morphinan and noscapine	[66]

4.1. Transcriptomics

Transcriptome sequencing is currently a standard experimental technique for identifying new transcripts, measuring the level of gene expression, and discovering novel genes. Studies on plants' genetics and functional genomics have been transformed by transcriptome sequencing, especially for non-model species lacking sequenced genomes [67]. The transcriptome analysis of commercially significant plants has advanced significantly over the last ten years because of developments that overcame medicinal plant genome sequencing difficulties. The prospective method for illuminating biological processes, such

as the manufacture of bioactive substances, growth and development, and the genetic variety of medicinal plants, is transcriptome analysis [68–70]. Fast-track breeding of medicinal plants is made possible by identifying essential genes and genetic markers associated with manufacturing secondary metabolites [71]. Transcriptomics, the comprehensive study of transcription, genomics, and proteomics, has unquestionably aided in developing a systems biology approach, building on the substantial advancements in high-throughput technology (microarrays, automated sequencing, and mass spectrometry).

Nevertheless, data mining has been streamlined, speeding up the discovery process thanks to practical computational tools (intelligent data networks, query, retrieval, analysis, and visualization tools). For microarray and RNA-seq investigations, transcriptomics approaches are highly parallel and need extensive processing to obtain valuable data. Each spot on a microarray chip represents a distinct oligonucleotide probe, and fluorescence intensity directly measures the abundance of a particular sequence. Spots are sequenced one nucleotide at a time in a high-throughput sequencing flow cell, with the color at each round representing the next nucleotide in the sequence. Other iterations of similar methods use more or fewer color channels [72]. RNA-seq studies provide a significant amount of raw sequence reads, which must be processed to produce useable data. Depending on the experimental design and aims, various bioinformatics software packages are often needed for data analysis. The process has four steps: quality control, alignment, quantification, and differential expression [73]. A command-line interface is used to execute the most well-liked RNA-seq tools in a Unix environment or inside the R/Bioconductor statistical environment [74].

The express sequence tag (EST) technique, a standard transcriptome analysis method, has given a quick and affordable alternative path toward the identification of genes encoding particular enzymes and (or) regulatory factors involved in secondary biosynthesis pathways [75]. At present, the transcriptomes of a large number of medicinal plants, such as *C. chinensis* [2], opium poppy [30], *Artemisia annua* [47], *Camptotheca acuminata* [46], *Salvia miltiorrhiza* [45,76,77], and *Stevia rebaudiana* Bertoni [78], were published online. These transcriptome data identified genes involved in the manufacture of benzylisoquinoline alkaloids, glycyrhizin, artemisinin, camptothecin, saikosaponins, taxol, tanshinone, salvianolic acid, and steviol glycoside. Table 2 lists the genes for secondary metabolite production that were found after the genomes of several medicinal plants were assembled, and transcriptomics data were merged. Genes from *Trachyspermum ammi* that are involved in thymol oil synthesis [79] and eight putative diterpene synthases (*diTPS*) of *Tripterygium wilfordii* [80] that are involved in triptolide biosynthesis have been identified by comparative transcriptome analysis. Chang et al. (2022) [81] used RNA-seq-based data from five tissues to assess the expression of the 104 moringa genes identified in their study as putatively involved in glucosinolate biosynthesis. Of the 104 genes, 84 and 88 were expressed in seeds and leaves, respectively, where glucosinolate biosynthesis is more abundant.

Table 2. Multi-omics approaches to study the biosynthesis pathway of secondary metabolites.

Species	Omics Techniques	Metabolites	References
Salvia miltiorrhiza Bunge.	NMR-based metabolomics, transcriptomics	Salvianolic acids	[76]
Quercus ilex	Metabolomics, proteomics, transcriptomics	62 metabolites	[82]
Perilla frutescens	Metabolomics, transcriptomics	Terpenoids, flavonoids, and phenylpropanoid	[83]
Rosa roxburghii Tratt.	Metabolomics, transcriptomics	Amino acid, phenylpropanoid, and flavonoid	[84]

Table 2. *Cont.*

Species	Omics Techniques	Metabolites	References
Astragalus membranaceus Bge.	Metabolomics, transcriptomics	Phenylpropanoid, flavonoid, and isoflavonoid	[85]
Chenopodium quinoa Willd.	Transcriptomics, metabolomics	Flavonoid	[86]
Triticum aestivum L.	Transcriptomics, metabolomics	Phenolic	[87]
Gardenia jasminoides Ellis.	Transcriptomics, metabolomics	Iridoid and crocin	[88]
Flammulina velutipes	Transcriptomics, metabolomics	Ergosterol	[89]
Perilla frutescens L.	Transcriptomics, metabolomics	Flavonoid	[90]
Euphorbia lathyris L.	Transcriptomics, metabolomics	Ingenol	[91]
Triticum aestivum L.	Transcriptomics, metabolomics	Anthocyanin	[92]
Setaria italica L.	Transcriptomics, metabolomics	Phenylpropanoid, flavonoid, and lignin	[93]
Salvia miltiorrhiza	Transcriptomics, metabolomics	Tanshinone	[77]
Acer mandshuricum	Transcriptomics, metabolomics	Anthocyanin	[94]
Corydalis yanhusuo	Transcriptomics, metabolomics	Benzylisoquinoline alkaloid	[95]
Zanthoxylum schinifolium Sieb.	Transcriptomics, metabolomics	Phenylpropanoid, flavonoid, flavone, and flavonol	[96]
Leptobryum pyriforme	Transcriptomics, metabolomics	Flavonoid	[97]
Angelica sinensis	Transcriptomics, metabolomics	Phenylpropanoid	[98]
Dendrobium huoshanense	Transcriptomics, metabolomics	Flavonoid	[99]
Solanum lycopersicum L.	Transcriptomics, metabolomics	Phenolamide	[100]
Dendrobium sinense	Transcriptomics, metabolomics	Purine and phenylpropanoid	[101]
Nicotiana tabacum L.	Proteomic, metabolomic	Aroma precursors	[102]

Co-expression of genes is common among those involved in producing secondary metabolites. Based on the transcriptome data of many samples, co-expression analysis is used to screen genes engaged in a particular pathway and reduce the number of candidate genes. Transcriptome analysis of the tanshinone production in *S. miltiorrhiza* revealed the co-expression of six cytochrome P450 genes with *SmCPS* and *SmKSL* [103]. The comparative transcriptome studies have also been used to screen the potential genes for co-expression in the production of lignans and phenylethanoid glycosides in the three species of Forsythia, *F. suspensa*, *F. viridissima*, and *F. koreana* [104]. The chromosomal-scale genome assembly made it possible to examine medicinal plants' high contiguity and coverage transcriptome and greatly aided in identifying genes involved in manufacturing secondary compounds.

4.2. Metabolomics

The emergence of omics-based approaches such as genomics and metabolomics has highlighted the potential to identify new plant secondary metabolites and give extensive data on the regulatory networks of genes synthesizing the secondary metabolites. An effec-

tive platform for detecting and identifying known and unidentified secondary metabolites and enhancing secondary metabolite synthesis is provided by combining metabolomics with gene editing methods such as CRISPR-Cas9 [105]. Modern plant metabolomics, a more advanced plant metabolism, has sped up gene discovery and clarified several natural product biosynthesis pathways [106]. Metabolomics plays two roles in these studies: (1) identifying the target phytochemical's spatial and temporal distribution as influenced by plant development and environmental cues and (2) identifying related compounds that may be thought of as alternative products of promiscuous enzymes or as intermediates in the biosynthesis of the target phytochemical. Such information may open doors to discovering biosynthetic pathways when combined with the generalizable principles of organic synthetic chemistry. Therefore, integrating metabolomics with genome-based functional characterizations of gene products enables the fast identification of new biosynthetic routes to specialized metabolites.

Spectrophotometric techniques that were compared to a certain pure standard can fall short of the actual ones. A newly created LC-MS-based, extensively targeted metabolomics approach is used to ensure the differences in total contents and composition [107,108]. The ultra-high resolution capacity of FT-ICR-MS, which also permits the inference of the metabolite annotations following database searches, is used to identify the chemical composition of the metabolomic peaks found in the LC-MS study [109]. Metabolites have been subjected to hierarchical clustering analysis (HCA), principle component analysis (PCA), and partial least squares–discriminant analysis (PLS-DA) to investigate metabolite species-specific accumulation [110]. Extensive LC-MS metabolic profiling of different tissues revealed at least 54 flavonoid compounds in A. thaliana, including 35 flavonols, 11 anthocyanins, and 8 proanthocyanidins [111]. Furthermore, two P450s, i.e., *CYP88D6* [112] and *CYP72A15486*, catalyze the oxidation of β-amyrin to glycyrrhetinic acid in *Glycyrrhiza uralensis* (licorice), and were discovered by an elegant analysis integrating expressed sequence tag studies, gene expression analyses, and metabolic profiling. In addition, metabolomics has been used to identify the glycosyltransferases that catalyze the following "tailoring processes" of triterpene saponins. In *M. truncatula*, an assortment of co-expressed genes led to the discovery of the enzyme *UGT73F3*, which glucosylates hederagenin at the C-28 position [113]. Tryptophan decarboxylase and secologanin synthase, the genes encoding the enzymes that catalyze the first committed reactions in the biosynthetic pathways, are suppressed in RNA interference lines used in metabolomics research to identify candidates for intermediary metabolites in the biosynthetic pathway [114]. A wealth of knowledge on the biological conditions of medicinal plants is provided through the integration of omics, particularly metabolomic profiling and transcriptomics. The advancement of molecular biology technology may open the door for new omics methods. Transcriptomics, genomics, and proteomics may be utilized in conjunction with the integration of transcriptomics and metabolomics to aid in the analysis of molecular biology and the biosynthetic pathway in plant secondary metabolism studies.

4.3. Proteomics

There is a logical need to understand how this genetic information is fully expressed, which has given rise to the fields of transcriptomics, metabolomics, and proteomics. The development of advanced genomics capabilities is associated with the ability to comprehensively and quickly determine and assemble plant genomes. Identification of proteins involved in biological processes is made possible by proteomics, which is the systematic examination of (differentially) expressed proteins. Studies of the proteins involved in the metabolic pathways that result in secondary metabolites have previously been conducted using proteomics in plant sciences [115]. The time-consuming chemistry of the intermediates in the present method of protein identification, enzyme separation, and characterization makes it challenging to identify regulatory or transport proteins. Since the proteomic technique allows for the discovery of regulatory and transport proteins in addition to enzymes, it may be quicker and more thorough. One may identify highly

conserved proteins by comparing a protein's sequence to those of *Arabidopsis thaliana* and other plant species [116]. Dedicated EST databases have been developed and used to identify proteins [117]. Model organisms have a few SM, making sequence annotation a problematic task. Early on, following the publication of the grapevine [118,119] genome sequence, functionally annotating proteomic research has shown to be an effective technique. This method involves retrieving gene ontology (GO) words from the most closely related protein using a BLAST similarity search using the BLAST2GO tool [120]. A minority of annotations are carefully selected and based on experimental evidence, albeit caution is advised since many of them are excessively generic.

However, the structural and quantitative data obtained using various proteomic approaches remain crucial for locating the target metabolites and the acting metabolic enzymes co-localizing in the cell. It can also be used to characterize protein complexes and, to a lesser extent, to analyze protein–protein interactions. As a result of everything mentioned above, proteogenomics, a cutting-edge technique for genome annotation, was created [121]. Proteogenomics seeks to uncover variant or unidentified proteins in bottom-up proteomics by searching transcriptome or genome-derived bespoke protein databases. However, actual studies show that the peptide identifications produced by these massive proteogenomic databases are less sensitive. Several methods have been suggested, such as creating limited transcriptome-informed protein databases, which only include proteins whose transcripts have been found in the sample-matched transcriptome. It was discovered that they improved peptide identification sensitivity [122]. The microsomal fraction of *Nicotiana tabacum* trichomes underwent extensive fractionation, peptide identification by LC MALDI MS/MS, and sequence analysis for transmembrane span prediction, which resulted in the identification of 165 membrane proteins, of which 39 were putative transporters, and eight were of the ABC type [123]. A number of proteins, including alcohol dehydrogenases, terpene cyclases, cytochrome P450-dependent monooxygenases, and glycosyl- and methyl-transferases, has been proposed as candidates to be involved in other unidentified steps of this pathway in addition to being identified by analogy with other plant species of already characterized proteins that act in well-defined steps of the TIA biosynthesis pathway [117]. The capacity to effectively access the SM proteome has depended on the availability of species-specific nucleotide databases and the technological technique, highlighting the potency of combining deep mRNA and protein sequencing to access plant-specific SM. After a 2-DE-based process in opium poppy (*Papaver somniferum*) cell cultures, only one enzyme implicated in the production of BIAs was discovered [124]. In contrast, the identification of 1004 proteins, which comprised almost all of the pathway's enzymes, was made possible by the earlier construction of a particular EST database from deep transcriptome analysis, followed by an LC-MS/MS analysis of 1DE-fractionated complete protein extracts [125]. However, unlike the study of DNA, proteins provide some particular difficulties. For instance, proteins do not have a PCR equivalent, making it challenging to analyze low-quantity proteins. Additionally, natural protein conformations must be preserved to obtain significant findings from investigations of protein interactions.

Additionally, discovery investigations in plant SM have been carried out using differential proteomics techniques that outperform single staining 2-DE. These include differential in gel electrophoresis (DIGE) as a top-down protein-centric approach [126] and isobaric tags for relative and absolute quantitation (iTRAQ) [127] and label-free [128] as bottom-up, peptide-centric approaches. Among the 172 unregulated proteins discovered by DIGE, three TIA biosynthesis enzymes were found in *Catharanthus roseus* cell cultures along with five more enzymes that may be putatively implicated in the process [117]. When DIGE and iTRAQ data were compared, they revealed significant quantitative correlations between frequently recognized proteins although iTRAQ excels in proteome coverage. However, the slight overlap between the two proteome studies highlights the need to use several strategies better to comprehend the target process under investigation [127]. Even though protein analysis technology is developing quickly, it is still impossible to analyze proteins on par with nucleic acid research. Additionally, these methods must be error-tolerant of

accounting for sequencing mistakes, polymorphisms, and conservative substitutions to access unannotated DNA libraries from other species. Before protein analysis on a broad scale (such as mapping the proteome of medicinal plants) becomes a reality, new methods must be developed.

4.4. Multi-Omics Approaches

New metabolite discoveries and a new method of understanding biosynthetic pathways have been made possible by the advent of functional genomics, including transcriptomics, metabolomics, and proteomics [129]. The multi-omics study may significantly elucidate the diversity, location, and evolution of BGCs [84]. Genes synthesizing specific metabolites may often be identified by the association between their expression and the accumulation of related metabolites [130]. To demonstrate variations in metabolites and gene pathway enrichment concurrently, differentially expressed genes (DEGs) and differentially expressed metabolites (DEMs) are mapped to KEGG pathways. Each group's DEGs and DEMs undergo correlation analysis, and genes' and metabolites' Pearson correlation coefficients are computed. Canonical correlation analysis is utilized to represent the overall correlation between the transcriptome and metabolome [131]. The O2PLS model is built using all DEGs and DEMs, and the variables with the highest correlation and weight across datasets are chosen based on the load diagram, whereas significant variables that impact another set are filtered out [132]. For instance, the relative abundance of transcripts encoding the alkaloid biosynthetic enzymes correlates with the induction of BIAs accumulation in *Papaver somniferum* [133]. Zhou et al. (2021) [83] combined metabolomics and transcriptomics analyses to explore the bioactive constituent biosynthesis in the leaves, stem, and root of *Perilla frutescens* (L.) Britt. Transcriptome sequencing profiles revealed that terpenoids, flavonoids, and phenylpropanoids biosynthetic genes are regulated differently, with most of these genes highly expressed in leaves. Transcriptomics, proteomics, and metabolomics were used by Cristina et al. (2018) [82] to study the biosynthesis of secondary metabolites, carbohydrates, energy, amino acid, lipid, and nucleotide metabolism in the orphan species *Quercus ilex*.

Moreover, a comprehensive multi-omics analysis report by Liu et al. (2019) [76] employed nuclear magnetic resonance (NMR)-based metabolomics and transcriptomics techniques to study the biosynthesis pathway of salvianolic acids in *Salvia miltiorrhiza* Bunge. The recent studies on discovering secondary metabolite biosynthetic genes in medicinal plants by combining different omics data analyses are summarized in Table 2. A group of FAD2-related enzymes has been found by comparing the transcriptomes and metabolomes (in particular, fatty acids and lipids) of growing seeds that accumulate "unusual" fatty acids [134] that are responsible for the generation of hydroxy fatty acids, epoxy-fatty acids, conjugated fatty acids, and acetylenic fatty acids. A relatively complete biosynthetic pathway for the three *O*-methylated flavones, namely hispidulin, jaceosidin, and eupatilin, was conjected based on synergistic genome sequencing analysis transcriptomics and metabolomics [135]. Transcriptomics, phenolic metabolomics, and functional gene analyses were used to characterize an aggregation gene cluster associated with the biosynthesis of hydrolyzable tannins (HTs) [65]. Hence, comparing transcriptomics and metabolomics data enables the precise annotation of a wide variety of genes in specialized metabolism.

4.5. Other Tools and Platforms for Medicinal Plant Functional Genomics

4.5.1. Metabolite-Based, Genome-Wide Association Study (mGWAS)

To analyze plant metabolism's genetic and biochemical underpinnings, metabolite-based, genome-wide association studies (mGWAS) have become a potent alternative forward genetics technique [136]. mGWAS in plants gets extra benefits from both huge diversity [130] and high inheritability of most metabolites [137,138] (Figure 2). It is feasible to develop testable hypotheses that can be confirmed in subsequent experiments since putative associations' biological and functional justifications in mGWAS are corrob-

orated. Candidate genes can be discovered by searching for a protein or protein cluster that is biochemically related to the metabolic trait associated with it and encoded at the associated loci by performing cluster analysis of candidate genes concerning homologous genes with known functions and by cross-referencing with linkage mapping data. Rapid discovery of new genes and potential networks of plant metabolism is made possible by integrating GWAS findings with other types of genome-scale data, such as transcript profiling or proteomics datasets [137–139]. mGWAS was initially applied in the model species *A. thaliana* [140], then successfully performed and extended in some crops [137,141]. Within structured mapping populations, the interactions between genotypes, environment, and development have been shown to affect the accumulation of secondary metabolites. A considerable bias was found in mGWAS on the naturally occurring ring variation of GSL accumulation in *Arabidopsis* toward identifying several causative genes for the GSL phenotypes in the two distinct tissues and during various developmental phases [142]. It implies that the genetic regulation of the natural variation of GSLs is spatially and temporally regulated.
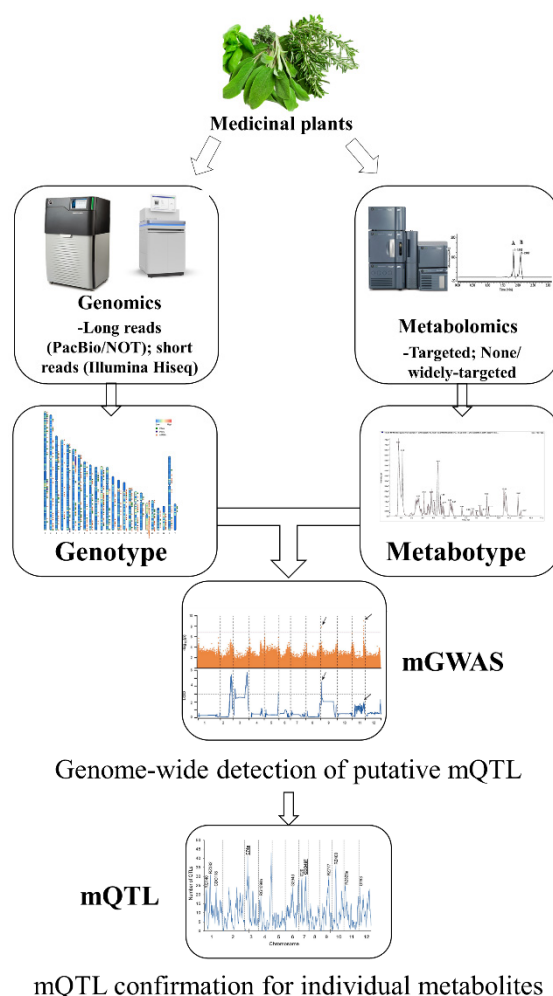

Figure 2. Flowchart of mGWAS. Combining the relevant data on the chemical structures and pathway architectures of associated/linked metabolites with the reported ortholog functions on metabolite analogs, candidate genes responsible for the mGWAS/mQTL assays loci are discovered.

Interestingly, different genetic control over metabolism was also imposed at the subspecies level [137]. Discovering the purposes of each gene in plant genomes is the goal of functional genomics. mGWAS may be used in functional genomics in addition to illuminating the genetic structure of plant metabolism. For example, mGWAS has identified 36 candidate genes encoding metabolites of physiological and dietary value using extensive

knowledge of several plant metabolic pathways [137]. Two spermidine hydroxycinnamoyl transferases were responsible for the normal variation of levels of spermidine conjugates in rice, according to a detailed analysis of spatiotemporal accumulation and natural variation of phenolamides in rice, followed by mGWAS [143]. However, regarding the use of mG-WAS in medicinal plants, there is no history. The reconstruction of biosynthetic pathways is made more accessible by understanding natural variation at the metabolic level using mGWAS, which is advantageous for biomedical research and the metabolic engineering of desired plant secondary metabolites [137,144,145].

4.5.2. Metabolic Quantitative Trait Loci (mQTL)

A practical method for identifying the genetic underpinnings of complex features in organisms such as medicinal plants is quantitative trait loci (QTL) mapping [146]. The self-incompatibility of medicinal plants, which results in limited populations and poor seed production, makes applying this strategy difficult. Despite this, more attention is being paid to the QTL mapping of numerous agronomic parameters of medical plants, notably yield, and secondary metabolites given the significance of medicinal plants in worldwide consumption [147–151]. Quantitative trait loci (QTL) analysis was employed in recent research to identify and pinpoint the locations of genomic areas that influence variance for quantitative traits since many aspects contributing to the phytochemical composition of medicinal plants are quantitative [152]. For example, Zhang et al. (2020) [153] have reported the application of QTL in ginsenoside biosynthesis in ginseng (*Panax ginseng* C.A. Meyer). They discovered that when expressed in the medicinal plant, most genes regulating quantitative traits were noticeably more likely to be spliced into numerous transcripts. The application of QTL mapping is also reported in several crops [127–129].

To understand the nature of inheritance of the amount and composition of these plant metabolites, metabolic quantitative trait loci (mQTL) investigations for targeted and non-targeted metabolites of therapeutic significance have also been carried out in numerous plant species [152]. In mQTL, MetaNetwork [154] software is used to map the metabolite variation. For example, 25 catechin-related QTLs were found using 2-year catechins data from 183 individuals crossbred from two tea plant kinds with different catechin compositions. Nine of these QTLs were confirmed throughout time and grouped in the chromosomal areas LG03 and LG11, pointing to a possible tandem duplication origin for the catechin genes in tea plants [149]. Similarly, 148 individuals from a pseudo-testcross population were used to determine theobromine and caffeine levels using 10 QTLs [148]. Recent research on QTL mapping discovered multiple new flavonoid-related QTLs due to the growth of the mapping populations [151]. To correctly identify QTLs in medicinal plants, more research will be required to expand the mapping populations and enhance the density of genetic linkage maps. Metabolites, on the other hand, are mechanistically more distant from the genome than mRNAs, leading to a crucial qualitative difference between metabolite QTL (mQTL) and eQTL studies. An advantageous feature of eQTL investigations is the ability to map mRNA to genes, which enables the search for a cis eQTL of each mRNA to be concentrated on a very narrow, gene-centered region. The connection between metabolite variation and genome-wide genetic variation is investigated in mQTL studies. Since there are many tests run, the effect sizes need to be much greater to be statistically significant. As a result, mQTLs are often more challenging to find than eQTLs with equal effect sizes in addition to perhaps being rarer.

4.5.3. Bulked-Segregant Analysis (BSA)

The recently developed BSA technique chooses and pools data from individuals with extreme phenotypes from biparental segregation populations, using the practical discovery of genes and alleles governing complex features via statistical genomic and phenotypic investigations. Bulked segregant analysis sequencing (BSA-seq) may be used to find the allele frequency (AF) of single-nucleotide polymorphisms by generating a DNA pool of offspring with extreme characteristics and utilizing molecular markers to analyze co-segregation

analysis between markers and traits in the two pools (SNPs). Therefore, BSA-seq allows the quick identification of essential genes linked with quantitative features or genes associated with key qualitative qualities regulated by a single gene [155]. The SNP-index technique was used to conduct SNP annotation, identify the effects (synonymous and non-synonymous mutations) of small indels [156] in the genome, and compare the variations in allele frequency across bulked pools [157,158]. The SNP index was formerly referred to as the number of short reads with SNPs that vary from the reference genome [159]. For instance, based on samples from the two tails of an F1 population that segregates with catechin concentration, the BSA technique, in conjunction with RNA sequencing, has aided the finding of a flavonoid three 5-hydroxylase (F3′5′H) genes in tea plants [160]. Studies conducted on the tomato and *Arabidopsis* have shown that the content of plant metabolites exhibits transgressive segregation, thus offering an opportunity to develop improved cultivars [161,162]. Traditionally, forward and reverse genetic screening, positional cloning, and quantitative trait loci mapping have been used to find causative genes within a metabolic pathway. High-throughput sequencing technologies may speed up and improve the efficiency of these traditional approaches. These sequencing technologies are not only capable of quickly detecting sequence variations (SNPs and insertions/deletions) across accessions; it can also be utilized in methods such as whole-genome resequencing in bulk segregant [163] analysis and *k-mer*-based [164] techniques to identify causative mutations. A single laboratory may currently employ sequencing techniques in combination with genetics methodologies for gene discovery because of the cheap cost, high coverage, accuracy, and accessible bioinformatics pipelines. For instance, having access to the genome sequences for several accessions enables association research to use linkage mapping or a genome-wide association technique to connect, for example, metabolite with genotype. With access to extensive transcriptome data sets, expression quantitative trait loci (eQTLs) associated to the production, control, or transport of metabolites may be found by tying transcript levels to sequence variations.

4.5.4. Weighted Gene Co-Expression Network Analysis (WGCNA)

The relationship between phenotype, including metabolites, patterns, and gene expression profiles, is a correlative method for annotating gene function. Access to genome sequencing enables the quick and affordable collection of expression abundance data. Coordinated expression studies (co-expression) may uncover genes within the same regulatory network from an expression atlas, in which various tissues and treatments are evaluated. WGCNA has effectively identified a subset of potential biosynthetic pathway genes for functional validation combined with examining the functional annotation of genes or transcripts [165]. WGCNA is a powerful approach to developing the regulatory network [166,167], and it is usually done in the R software WGCNA package [167,168]. Genes that were differently expressed in *Camptotheca acuminata* were used by Kang et al. (2021) [46] to create WGCNA. They discovered 30 clusters, including the seco-iridoid pathway genes geraniol synthase (*GS*), iridoid synthase (*IS*), 7-deoxyloganetic acid synthase (*7-DLS*), 7-deoxyloganetic acid O-glucosyltransferase (*7-DLGT*), and 7-deoxyloganic acid 7-hydroxylase (*7-DLH*), as well as other related genes, which were grouped. The WGCNA was also utilized by Tai et al. (2018) [169] to identify 35 co-expression modules in the tea plant, of which 20 modules were substantially linked to the production of catechins, theanine, and caffeine. The presence of 3 of the 20 co-expression modules, which included six genes involved in ascorbic acid biosynthesis and regeneration pathways, was shown to be positively linked with the ascorbic acid content of the guava (*Psidium guajava*), according to research on WGCNA [170]. According to WGCNA results in *indica-japonica* rice hybrids, ribosomal protein production and stress response were tightly related to the bisque4 and darkorange2 modules for types 1 and 3 of the core metabolites. Eight hub genes, namely *rpL32 8.1*, *novel.907*, *RPS15a*, and *RPL13a.4* in bisque4 as well as *OsSAP5*, *OsHSP20*, *PRX12*, and *DUF3741* in darkorange2, were discovered [171]. At various phases of the *C. panzhihuaensis* ovule and seed's development, 11 co-expression modules were

discovered by Liu et al. (2022) [172]. In summary, it enables the definition of modules (clusters), intra-modular hubs, and network nodes regarding their participation in modules as well as the analysis of the interactions between co-expression modules and the comparison of the network topologies of other networks (differential network analysis).

5. Conclusions and Prospects

Since the establishment of the human genome project (HGP), there has been a significant development regarding the high-throughput sequencing technologies, and accurate genome assembly has enabled the sequence of the medicinal plant genome at the chromosomal level in a short time and at low cost. Functional genomics (proteomics, transcriptomics, and metabolomics) paved many ways to identify key related genes and enzymes for studying secondary metabolite biosynthesis pathways. Furthermore, the combination of multi-omics provides a complete picture of medicinal plants, and their metabolism could identify novel genes and metabolites. Yet, there is not enough information regarding the application of mQTL, BSA, WGCNA, and mGWAS in medicinal plants, and our understanding of them needs to be strengthened.

However, high-throughput sequencing technologies and functional genomics reduced the cost and the complexity of DNA sequencing; many medicinal plants' genomes were sequenced and assembled, and shortly, this number will be increased; the multi-omics and bioinformatics approaches can be the best tools for the study of secondary metabolism. The promotion of medicinal plant genome projects contributes to the biosynthesis and screening of plant substrates and plant-based drugs and prompts the research efficiency of traditional medicine.

Abbreviations

BGCs	Biosynthetic gene clusters
NGS	Next-generation sequencing
BSA	Bulked-segregant analysis
WGCNA	Weighted gene co-expression network analysis
mGWAS	Metabolite-based, genome-wide association study
mQTL	Metabolic quantitative trait loci
HiSeq	*High-throughput sequencing*
WGD	Whole-genome duplication
BIAs	Benzylisoquinoline alkaloids
SNPs	Single-nucleotide polymorphisms
InDels	Insertions and deletions
SVs	Structure variations
NMR	*Nuclear magnetic resonance*

References

1. Sanger, F.; Nicklen, S.; Coulson, A.R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 5463–5467. [CrossRef] [PubMed]
2. Liu, Y.; Wang, B.; Shu, S.; Li, Z.; Song, C.; Liu, D.; Niu, Y.; Liu, J.; Zhang, J.; Liu, H.; et al. Analysis of the Coptis chinensis genome reveals the diversification of protoberberine-type alkaloids. *Nat. Commun.* **2021**, *12*, 3276. [CrossRef] [PubMed]
3. Tu, L.; Su, P.; Zhang, Z.; Gao, L.; Wang, J.; Hu, T.; Zhou, J.; Zhang, Y.; Zhao, Y.; Liu, Y.; et al. Genome of Tripterygium wilfordii and identification of cytochrome P450 involved in triptolide biosynthesis. *Nat. Commun.* **2020**, *11*, 971. [CrossRef] [PubMed]
4. Ren, L.; Shang, Y.; Yang, L.; Wang, S.; Wang, X.; Chen, S.; Bao, Z.; An, D.; Meng, F.; Cai, J.; et al. Chromosome-level de novo genome assembly of *Sarcophaga peregrina* provides insights into the evolutionary adaptation of flesh flies. *Mol. Ecol. Resour.* **2021**, *21*, 251–262. [CrossRef] [PubMed]
5. Song, Z.; Lin, C.; Xing, P.; Fen, Y.; Jin, H.; Zhou, C.; Gu, Y.Q.; Wang, J.; Li, X. A high-quality reference genome sequence of *Salvia miltiorrhiza* provides insights into tanshinone synthesis in its red rhizomes. *Plant Genome* **2020**, *13*, e20041. [CrossRef]
6. Xie, J.; Zhao, H.; Li, K.; Zhang, R.; Jiang, Y.; Wang, M.; Guo, X.; Yu, B.; Kong, H.; Jiao, Y.; et al. A chromosome-scale reference genome of *Aquilegia oxysepala* var. kansuensis. *Hortic. Res.* **2020**, *7*, 113. [CrossRef]
7. Rai, A.; Hirakawa, H.; Nakabayashi, R.; Kikuchi, S.; Hayashi, K.; Rai, M.; Tsugawa, H.; Nakaya, T.; Mori, T.; Nagasaki, H.; et al. Chromosome-level genome assembly of *Ophiorrhiza pumila* reveals the evolution of camptothecin biosynthesis. *Nat. Commun.* **2021**, *12*, 405. [CrossRef]
8. Chen, J.-D.; Zheng, C.; Ma, J.-Q.; Jiang, C.-K.; Ercisli, S.; Yao, M.-Z.; Chen, L. The chromosome-scale genome reveals the evolution and diversification after the recent tetraploidization event in tea plant. *Hortic. Res.* **2020**, *7*, 63. [CrossRef]
9. Cheng, Q.-Q.; Ouyang, Y.; Tang, Z.-Y.; Lao, C.-C.; Zhang, Y.-Y.; Cheng, C.-S.; Zhou, H. Review on the Development and Applications of Medicinal Plant Genomes. *Front. Plant Sci.* **2021**, *12*, 791219. [CrossRef]
10. Hamilton, J.P.; Robin Buell, C. Advances in plant genome sequencing. *Plant J.* **2012**, *70*, 177–190. [CrossRef]
11. Chen, F.; Dong, W.; Zhang, J.; Guo, X.; Chen, J.; Wang, Z.; Lin, Z.; Tang, H.; Zhang, L. The Sequenced Angiosperm Genomes and Genome Databases. *Front. Plant Sci.* **2018**, *9*, 418. [CrossRef]
12. Jiao, W.-B.; Schneeberger, K. The impact of third generation genomic technologies on plant genome assembly. *Curr. Opin. Plant Biol.* **2017**, *36*, 64–70. [CrossRef]
13. Caicedo, A.L.; Purugganan, M.D. Comparative Plant Genomics. Frontiers and Prospects. *Plant Physiol.* **2005**, *138*, 545–547. [CrossRef]
14. Koonin, E.V. Horizontal gene transfer: The path to maturity. *Mol. Microbiol.* **2003**, *50*, 725–727. [CrossRef]
15. Jensen, R.A. Orthologs and paralogs—We need to get it right. *Genome Biol.* **2001**, *2*, interactions1002.1. [CrossRef]
16. Murat, F.; Armero, A.; Pont, C.; Klopp, C.; Salse, J. Reconstructing the genome of the most recent common ancestor of flowering plants. *Nat. Genet.* **2017**, *49*, 490–496. [CrossRef]
17. Xu, Z.; Gao, R.; Pu, X.; Xu, R.; Wang, J.; Zheng, S.; Zeng, Y.; Chen, J.; He, C.; Song, J. Comparative Genome Analysis of *Scutellaria baicalensis* and *Scutellaria barbata* Reveals the Evolution of Active Flavonoid Biosynthesis. *Genom. Proteom. Bioinform.* **2020**, *18*, 230–240. [CrossRef]
18. Daccord, N.; Celton, J.-M.; Linsmith, G.; Becker, C.; Choisne, N.; Schijlen, E.; van de Geest, H.; Bianco, L.; Micheletti, D.; Velasco, R.; et al. High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. *Nat. Genet.* **2017**, *49*, 1099–1106. [CrossRef]
19. Zhao, Y.; Liu, X.; Guo, R.; Hu, K.; Cao, Y.; Dai, F. Comparative genomics and transcriptomics analysis reveals evolution patterns of selection in the *Salix phylogeny*. *BMC Genom.* **2019**, *20*, 253. [CrossRef]
20. Jarvis, E.D.; Mirarab, S.; Aberer, A.J.; Li, B.; Houde, P.; Li, C.; Ho, S.Y.W.; Faircloth, B.C.; Nabholz, B.; Howard, J.T.; et al. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* **2014**, *346*, 1320–1331. [CrossRef]

21. Robinson, G.E.; Hackett, K.J.; Purcell-Miramontes, M.; Brown, S.J.; Evans, J.D.; Goldsmith, M.R.; Lawson, D.; Okamuro, J.; Robertson, H.M.; Schneider, D.J. Creating a Buzz About Insect Genomes. *Science* **2011**, *331*, 1386. [CrossRef] [PubMed]

22. Genome 10K Community of Scientists. Genome 10K: A Proposal to Obtain Whole-Genome Sequence for 10 000 Vertebrate Species. *J. Hered.* **2009**, *100*, 659–674. [CrossRef] [PubMed]

23. Cheng, S.; Melkonian, M.; Smith, S.A.; Brockington, S.; Archibald, J.M.; Delaux, P.-M.; Li, F.-W.; Melkonian, B.; Mavrodiev, E.V.; Sun, W.; et al. 10KP: A phylodiverse genome sequencing plan. *Gigascience* **2018**, *7*, 1–9. [CrossRef] [PubMed]

24. OBrien, S.J.; Haussler, D.; Ryder, O. The birds of Genome10K. *Gigascience* **2014**, *3*, 32. [CrossRef] [PubMed]

25. Lewin, H.A.; Robinson, G.E.; Kress, W.J.; Baker, W.J.; Coddington, J.; Crandall, K.A.; Durbin, R.; Edwards, S.V.; Forest, F.; Gilbert, M.T.P.; et al. Earth BioGenome Project: Sequencing life for the future of life. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 4325–4333. [CrossRef]

26. Albert, V.A.; Barbazuk, W.B.; DePamphilis, C.W.; Der, J.P.; Leebens-Mack, J.; Ma, H.; Palmer, J.D.; Rounsley, S.; Sankoff, D.; Schuster, S.C.; et al. The Amborella Genome and the Evolution of Flowering Plants. *Science* **2013**, *342*, 1241089. [CrossRef]

27. Nelson, D.R.; Ming, R.; Alam, M.; Schuler, M.A. Comparison of Cytochrome P450 Genes from Six Plant Genomes. *Trop. Plant Biol.* **2008**, *1*, 216–235. [CrossRef]

28. Eric Schranz, M.; Mohammadin, S.; Edger, P.P. Ancient whole genome duplications, novelty and diversification: The WGD Radiation Lag-Time Model. *Curr. Opin. Plant Biol.* **2012**, *15*, 147–153. [CrossRef]

29. Xia, E.; Tong, W.; Hou, Y.; An, Y.; Chen, L.; Wu, Q.; Liu, Y.; Yu, J.; Li, F.; Li, R.; et al. The Reference Genome of Tea Plant and Resequencing of 81 Diverse Accessions Provide Insights into Its Genome Evolution and Adaptation. *Mol. Plant* **2020**, *13*, 1013–1026. [CrossRef]

30. Guo, L.; Winzer, T.; Yang, X.; Li, Y.; Ning, Z.; He, Z.; Teodor, R.; Lu, Y.; Bowser, T.A.; Graham, I.A.; et al. The opium poppy genome and morphinan production. *Science* **2018**, *362*, 343–347. [CrossRef]

31. Denoeud, F.; Carretero-Paulet, L.; Dereeper, A.; Droc, G.; Guyot, R.; Pietrella, M.; Zheng, C.; Alberti, A.; Anthony, F.; Aprea, G.; et al. The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science* **2014**, *345*, 1181–1184. [CrossRef]

32. Wang, J.-P.; Yu, J.-G.; Li, J.; Sun, P.-C.; Wang, L.; Yuan, J.-Q.; Meng, F.-B.; Sun, S.-R.; Li, Y.-X.; Lei, T.-Y.; et al. Two Likely Auto-Tetraploidization Events Shaped Kiwifruit Genome and Contributed to Establishment of the Actinidiaceae Family. *iScience* **2018**, *7*, 230–240. [CrossRef]

33. Van de Peer, Y.; Mizrachi, E.; Marchal, K. The evolutionary significance of polyploidy. *Nat. Rev. Genet.* **2017**, *18*, 411–424. [CrossRef]

34. Zhang, L.; Wu, S.; Chang, X.; Wang, X.; Zhao, Y.; Xia, Y.; Trigiano, R.N.; Jiao, Y.; Chen, F. The ancient wave of polyploidization events in flowering plants and their facilitated adaptation to environmental stress. *Plant Cell Environ.* **2020**, *43*, 2847–2856. [CrossRef]

35. Paterson, A.H.; Bowers, J.E.; Chapman, B.A. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9903–9908. [CrossRef]

36. Wei, C.; Yang, H.; Wang, S.; Zhao, J.; Liu, C.; Gao, L.; Xia, E.; Lu, Y.; Tai, Y.; She, G.; et al. Draft genome sequence of *Camellia sinensis* var. *sinensis* provides insights into the evolution of the tea genome and tea quality. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E4151–E4158. [CrossRef]

37. Winzer, T.; Gazda, V.; He, Z.; Kaminski, F.; Kern, M.; Larson, T.R.; Li, Y.; Meade, F.; Teodor, R.; Vaistij, F.E.; et al. A Papaver somniferum 10-Gene Cluster for Synthesis of the Anticancer Alkaloid Noscapine. *Science* **2012**, *336*, 1704–1708. [CrossRef]

38. Galanie, S.; Thodey, K.; Trenchard, I.J.; Interrante, M.F.; Smolke, C.D. Complete biosynthesis of opioids in yeast. *Science* **2015**, *349*, 1095–1100. [CrossRef]

39. Winzer, T.; Kern, M.; King, A.J.; Larson, T.R.; Teodor, R.I.; Donninger, S.L.; Li, Y.; Dowle, A.A.; Cartwright, J.; Bates, R.; et al. Morphinan biosynthesis in opium poppy requires a P450-oxidoreductase fusion protein. *Science* **2015**, *349*, 309–312. [CrossRef]

40. Farrow, S.C.; Hagel, J.M.; Beaudoin, G.A.W.; Burns, D.C.; Facchini, P.J. Stereochemical inversion of (S)-reticuline by a cytochrome P450 fusion in opium poppy. *Nat. Chem. Biol.* **2015**, *11*, 728–732. [CrossRef]

41. Conant, G.C.; Birchler, J.A.; Pires, J.C. Dosage, duplication, and diploidization: Clarifying the interplay of multiple models for duplicate gene evolution over time. *Curr. Opin. Plant Biol.* **2014**, *19*, 91–98. [CrossRef] [PubMed]

42. Chae, L.; Kim, T.; Nilo-Poyanco, R.; Rhee, S.Y. Genomic Signatures of Specialized Metabolism in Plants. *Science* **2014**, *344*, 510–513. [CrossRef] [PubMed]

43. Chen, S.; Xiang, L.; Guo, X.; Li, Q. An introduction to the medicinal plant genome project. *Front. Med.* **2011**, *5*, 178–184. [CrossRef] [PubMed]

44. Hirsch, C.N.; Robin Buell, C. Tapping the Promise of Genomics in Species with Complex, Nonmodel Genomes. *Annu. Rev. Plant Biol.* **2013**, *64*, 89–110. [CrossRef] [PubMed]

45. Ma, Y.; Cui, G.; Chen, T.; Ma, X.; Wang, R.; Jin, B.; Yang, J.; Kang, L.; Tang, J.; Lai, C.; et al. Expansion within the CYP71D subfamily drives the heterocyclization of tanshinones synthesis in Salvia miltiorrhiza. *Nat. Commun.* **2021**, *12*, 685. [CrossRef]

46. Kang, M.; Fu, R.; Zhang, P.; Lou, S.; Yang, X.; Chen, Y.; Ma, T.; Zhang, Y.; Xi, Z.; Liu, J. A chromosome-level Camptotheca acuminata genome assembly provides insights into the evolutionary origin of camptothecin biosynthesis. *Nat. Commun.* **2021**, *12*, 3531. [CrossRef]

47. Shen, Q.; Zhang, L.; Liao, Z.; Wang, S.; Yan, T.; Shi, P.; Liu, M.; Fu, X.; Pan, Q.; Wang, Y.; et al. The Genome of *Artemisia annua* Provides Insight into the Evolution of Asteraceae Family and Artemisinin Biosynthesis. *Mol. Plant* **2018**, *11*, 776–788. [CrossRef]

48. Hu, L.; Xu, Z.; Wang, M.; Fan, R.; Yuan, D.; Wu, B.; Wu, H.; Qin, X.; Yan, L.; Tan, L.; et al. The chromosome-scale reference genome of black pepper provides insight into piperine biosynthesis. *Nat. Commun.* **2019**, *10*, 4702. [CrossRef]

49. Sun, X.; Zhu, S.; Li, N.; Cheng, Y.; Zhao, J.; Qiao, X.; Lu, L.; Liu, S.; Wang, Y.; Liu, C.; et al. A Chromosome-Level Genome Assembly of Garlic (*Allium sativum*) Provides Insights into Genome Evolution and Allicin Biosynthesis. *Mol. Plant* **2020**, *13*, 1328–1339. [CrossRef]

50. Dong, S.; Liu, M.; Liu, Y.; Chen, F.; Yang, T.; Chen, L.; Zhang, X.; Guo, X.; Fang, D.; Li, L.; et al. The genome of *Magnolia biondii* Pamp. provides insights into the evolution of Magnoliales and biosynthesis of terpenoids. *Hortic. Res.* **2021**, *8*, 38. [CrossRef]

51. Kim, J.; Kang, S.-H.; Park, S.-G.; Yang, T.-J.; Lee, Y.; Kim, O.T.; Chung, O.; Lee, J.; Choi, J.-P.; Kwon, S.-J.; et al. Whole-genome, transcriptome, and methylome analyses provide insights into the evolution of platycoside biosynthesis in *Platycodon grandiflorus*, a medicinal plant. *Hortic. Res.* **2020**, *7*, 112. [CrossRef]

52. Xu, W.; Zhang, L.; Cunningham, A.B.; Li, S.; Zhuang, H.; Wang, Y.; Liu, A. Blue genome: Chromosome-scale genome reveals the evolutionary and molecular basis of indigo biosynthesis in *Strobilanthes cusia*. *Plant J.* **2020**, *104*, 864–879. [CrossRef]

53. Zhao, Q.; Yang, J.; Cui, M.-Y.; Liu, J.; Fang, Y.; Yan, M.; Qiu, W.; Shang, H.; Xu, Z.; Yidiresi, R.; et al. The Reference Genome Sequence of *Scutellaria baicalensis* Provides Insights into the Evolution of Wogonin Biosynthesis. *Mol. Plant* **2019**, *12*, 935–950. [CrossRef]

54. Lau, K.H.; Bhat, W.W.; Hamilton, J.P.; Wood, J.C.; Vaillancourt, B.; Wiegert-Rininger, K.; Newton, L.; Hamberger, B.; Holmes, D.; Hamberger, B.; et al. Genome assembly of Chiococca alba uncovers key enzymes involved in the biosynthesis of unusual terpenoids. *DNA Res.* **2020**, *27*, dsaa013. [CrossRef]

55. Ma, Q.; Sun, T.; Li, S.; Wen, J.; Zhu, L.; Yin, T.; Yan, K.; Xu, X.; Li, S.; Mao, J.; et al. The Acer truncatum genome provides insights into nervonic acid biosynthesis. *Plant J.* **2020**, *104*, 662–678. [CrossRef]

56. Kang, S.-H.; Pandey, R.P.; Lee, C.-M.; Sim, J.-S.; Jeong, J.-T.; Choi, B.-S.; Jung, M.; Ginzburg, D.; Zhao, K.; Won, S.Y.; et al. Genome-enabled discovery of anthraquinone biosynthesis in Senna tora. *Nat. Commun.* **2020**, *11*, 5875. [CrossRef]

57. Zhou, W.; Wang, Y.; Li, B.; Petijová, L.; Hu, S.; Zhang, Q.; Niu, J.; Wang, D.; Wang, S.; Dong, Y.; et al. Whole-genome sequence data of *Hypericum perforatum* and functional characterization of melatonin biosynthesis by N-acetylserotonin O-methyltransferase. *J. Pineal Res.* **2021**, *70*, e12709. [CrossRef]

58. Cui, P.; Lin, Q.; Fang, D.; Zhang, L.; Li, R.; Cheng, J.; Gao, F.; Shockey, J.; Hu, S.; Lü, S. Tung Tree (*Vernicia fordii*, Hemsl.) Genome and Transcriptome Sequencing Reveals Co-Ordinate Up-Regulation of Fatty Acid β-Oxidation and Triacylglycerol Biosynthesis Pathways During Eleostearic Acid Accumulation in Seeds. *Plant Cell Physiol.* **2018**, *59*, 1990–2003. [CrossRef]

59. Sun, W.; Leng, L.; Yin, Q.; Xu, M.; Huang, M.; Xu, Z.; Zhang, Y.; Yao, H.; Wang, C.; Xiong, C.; et al. The genome of the medicinal plant *Andrographis paniculata* provides insight into the biosynthesis of the bioactive diterpenoid neoandrographolide. *Plant J.* **2019**, *97*, 841–857. [CrossRef]

60. Wuyun, T.; Wang, L.; Liu, H.; Wang, X.; Zhang, L.; Bennetzen, J.L.; Li, T.; Yang, L.; Liu, P.; Du, L.; et al. The Hardy Rubber Tree Genome Provides Insights into the Evolution of Polyisoprene Biosynthesis. *Mol. Plant* **2018**, *11*, 429–442. [CrossRef]

61. Xu, Z.; Pu, X.; Gao, R.; Demurtas, O.C.; Fleck, S.J.; Richter, M.; He, C.; Ji, A.; Sun, W.; Kong, J.; et al. Tandem gene duplications drive divergent evolution of caffeine and crocin biosynthetic pathways in plants. *BMC Biol.* **2020**, *18*, 63. [CrossRef] [PubMed]

62. Zheng, X.; Chen, D.; Chen, B.; Liang, L.; Huang, Z.; Fan, W.; Chen, J.; He, W.; Chen, H.; Huang, L.; et al. Insights into salvianolic acid B biosynthesis from chromosome-scale assembly of the *Salvia bowleyana* genome. *J. Integr. Plant Biol.* **2021**, *63*, 1309–1323. [CrossRef] [PubMed]

63. Lv, Q.; Qiu, J.; Liu, J.; Li, Z.; Zhang, W.; Wang, Q.; Fang, J.; Pan, J.; Chen, Z.; Cheng, W.; et al. The Chimonanthus salicifolius genome provides insight into magnoliid evolution and flavonoid biosynthesis. *Plant J.* **2020**, *103*, 1910–1923. [CrossRef] [PubMed]

64. Hu, Y.; Ma, D.; Ning, S.; Ye, Q.; Zhao, X.; Ding, Q.; Liang, P.; Cai, G.; Ma, X.; Qin, X.; et al. High-Quality Genome of the Medicinal Plant Strobilanthes cusia Provides Insights Into the Biosynthesis of Indole Alkaloids. *Front. Plant Sci.* **2021**, *12*, 742420. [CrossRef] [PubMed]

65. Wang, L.; Lei, T.; Han, G.; Yue, J.; Zhang, X.; Yang, Q.; Ruan, H.; Gu, C.; Zhang, Q.; Qian, T.; et al. The chromosome-scale reference genome of Rubus chingii Hu provides insight into the biosynthetic pathway of hydrolyzable tannins. *Plant J.* **2021**, *107*, 1466–1477. [CrossRef]

66. Pei, T.; Yan, M.; Kong, Y.; Fan, H.; Liu, J.; Cui, M.; Fang, Y.; Ge, B.; Yang, J.; Zhao, Q. The genome of Tripterygium wilfordii and characterization of the celastrol biosynthesis pathway. *Gigabyte* **2021**, *2021*, 1–30. [CrossRef]

67. Ozsolak, F.; Milos, P.M. RNA sequencing: Advances, challenges and opportunities. *Nat. Rev. Genet.* **2011**, *12*, 87–98. [CrossRef]

68. Jennewein, S.; Wildung, M.R.; Chau, M.; Walker, K.; Croteau, R. Random sequencing of an induced Taxus cell cDNA library for identification of clones involved in *Taxol biosynthesis*. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9149–9154. [CrossRef]

69. Lin, X.; Zhang, J.; Li, Y.; Luo, H.; Wu, Q.; Sun, C.; Song, J.; Li, X.; Wei, J.; Lu, A.; et al. Functional genomics of a living fossil tree, Ginkgo, based on next-generation sequencing technology. *Physiol. Plant.* **2011**, *143*, 207–218. [CrossRef]

70. Murata, J.; Roepke, J.; Gordon, H.; De Luca, V. The Leaf Epidermome of Catharanthus roseus Reveals Its Biochemical Specialization. *Plant Cell* **2008**, *20*, 524–542. [CrossRef]

71. Graham, I.A.; Besser, K.; Blumer, S.; Branigan, C.A.; Czechowski, T.; Elias, L.; Guterman, I.; Harvey, D.; Isaac, P.G.; Khan, A.M.; et al. The Genetic Map of Artemisia annua L. Identifies Loci Affecting Yield of the Antimalarial Drug Artemisinin. *Science* **2010**, *327*, 328–331. [CrossRef]

72. Lowe, R.; Shirley, N.; Bleackley, M.; Dolan, S.; Shafee, T. Transcriptomics technologies. *PLoS Comput. Biol.* **2017**, *13*, e1005457. [CrossRef]

73. Van Verk, M.C.; Hickman, R.; Pieterse, C.M.J.; Van Wees, S.C.M. RNA-Seq: Revelation of the messengers. *Trends Plant Sci.* **2013**, *18*, 175–179. [CrossRef]

74. Huber, W.; Carey, V.J.; Gentleman, R.; Anders, S.; Carlson, M.; Carvalho, B.S.; Bravo, H.C.; Davis, S.; Gatto, L.; Girke, T.; et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* **2015**, *12*, 115–121. [CrossRef]

75. Ohlrogge, J.; Benning, C. Unraveling plant metabolism by EST analysis. *Curr. Opin. Plant Biol.* **2000**, *3*, 224–228. [CrossRef]

76. Liu, X.; Jin, M.; Zhang, M.; Li, T.; Sun, S.; Zhang, J.; Dai, J.; Wang, Y. The application of combined 1H NMR-based metabolomics and transcriptomics techniques to explore phenolic acid biosynthesis in *Salvia miltiorrhiza* Bunge. *J. Pharm. Biomed. Anal.* **2019**, *172*, 126–138. [CrossRef]

77. Gao, W.; Sun, H.-X.; Xiao, H.; Cui, G.; Hillwig, M.L.; Jackson, A.; Wang, X.; Shen, Y.; Zhao, N.; Zhang, L.; et al. Combining metabolomics and transcriptomics to characterize tanshinone biosynthesis in Salvia miltiorrhiza. *BMC Genom.* **2014**, *15*, 73. [CrossRef]

78. Xu, X.; Yuan, H.; Yu, X.; Huang, S.; Sun, Y.; Zhang, T.; Liu, Q.; Tong, H.; Zhang, Y.; Wang, Y.; et al. The chromosome-level Stevia genome provides insights into steviol glycoside biosynthesis. *Hortic. Res.* **2021**, *8*, 129. [CrossRef]

79. Soltani Howyzeh, M.; Sadat Noori, S.A.; Shariati, J.V.; Amiripour, M. Comparative transcriptome analysis to identify putative genes involved in thymol biosynthesis pathway in medicinal plant *Trachyspermum ammi* L. *Sci. Rep.* **2018**, *8*, 13405. [CrossRef]

80. Su, P.; Guan, H.; Zhao, Y.; Tong, Y.; Xu, M.; Zhang, Y.; Hu, T.; Yang, J.; Cheng, Q.; Gao, L.; et al. Identification and functional characterization of diterpene synthases for triptolide biosynthesis from *Tripterygium wilfordii*. *Plant J.* **2018**, *93*, 50–65. [CrossRef]

81. Chang, J.; Marczuk-Rojas, J.P.; Waterman, C.; Garcia-Llanos, A.; Chen, S.; Ma, X.; Hulse-Kemp, A.; Van Deynze, A.; Van de Peer, Y.; Carretero-Paulet, L. Chromosome-scale assembly of the *Moringa oleifera* Lam. genome uncovers polyploid history and evolution of secondary metabolism pathways through tandem duplication. *Plant Genome* **2022**, *15*, e20238. [CrossRef] [PubMed]

82. López-Hidalgo, C.; Guerrero-Sánchez, V.M.; Gómez-Gálvez, I.; Sánchez-Lucas, R.; Castillejo-Sánchez, M.A.; Maldonado-Alconada, A.M.; Valledor, L.; Jorrín-Novo, J.V. A Multi-Omics Analysis Pipeline for the Metabolic Pathway Reconstruction in the Orphan Species Quercus ilex. *Front. Plant Sci.* **2018**, *9*, 935. [CrossRef] [PubMed]

83. Zhou, P.; Yin, M.; Dai, S.; Bao, K.; Song, C.; Liu, C.; Wu, Q. Multi-omics analysis of the bioactive constituents biosynthesis of glandular trichome in Perilla frutescens. *BMC Plant Biol.* **2021**, *21*, 277. [CrossRef] [PubMed]

84. Li, N.; Jiang, L.; Liu, Y.; Zou, S.; Lu, M.; An, H. Metabolomics Combined with Transcriptomics Analysis Revealed the Amino Acids, Phenolic Acids, and Flavonol Derivatives Biosynthesis Network in Developing Rosa roxburghii Fruit. *Foods* **2022**, *11*, 1639. [CrossRef] [PubMed]

85. Wu, X.; Li, X.; Wang, W.; Shan, Y.; Wang, C.; Zhu, M.; La, Q.; Zhong, Y.; Xu, Y.; Nan, P.; et al. Integrated metabolomics and transcriptomics study of traditional herb Astragalus membranaceus Bge. var. mongolicus (Bge.) Hsiao reveals global metabolic profile and novel phytochemical ingredients. *BMC Genom.* **2020**, *21*, 697. [CrossRef]

86. Liu, Y.; Liu, J.; Kong, Z.; Huan, X.; Li, L.; Zhang, P.; Wang, Q.; Guo, Y.; Zhu, W.; Qin, P. Transcriptomics and metabolomics analyses of the mechanism of flavonoid synthesis in seeds of differently colored quinoa strains. *Genomics* **2022**, *114*, 138–148. [CrossRef]

87. Ma, D.; Xu, B.; Feng, J.; Hu, H.; Tang, J.; Yin, G.; Xie, Y.; Wang, C. Dynamic Metabolomics and Transcriptomics Analyses for Characterization of Phenolic Compounds and Their Biosynthetic Characteristics in Wheat Grain. *Front. Nutr.* **2022**, *9*, 844337. [CrossRef]

88. Pan, Y.; Zhao, X.; Wang, Y.; Tan, J.; Chen, D. Metabolomics integrated with transcriptomics reveals the distribution of iridoid and crocin metabolic flux in Gardenia jasminoides Ellis. *PLoS ONE* **2021**, *16*, e0256802. [CrossRef]

89. Wang, R.; Ma, P.; Li, C.; Xiao, L.; Liang, Z.; Dong, J. Combining transcriptomics and metabolomics to reveal the underlying molecular mechanism of ergosterol biosynthesis during the fruiting process of Flammulina velutipes. *BMC Genom.* **2019**, *20*, 999. [CrossRef]

90. Jiang, T.; Guo, K.; Liu, L.; Tian, W.; Xie, X.; Wen, S.; Wen, C. Integrated transcriptomic and metabolomic data reveal the flavonoid biosynthesis metabolic pathway in Perilla frutescens (L.) leaves. *Sci. Rep.* **2020**, *10*, 16207. [CrossRef]

91. Zhao, W.; Peng, J.; Wang, F.; Tian, M.; Li, P.; Feng, B.; Yin, M.; Xu, Y.; Xue, J.-Y.; Xue, J.; et al. Integrating metabolomics and transcriptomics to unveil the spatiotemporal distribution of macrocyclic diterpenoids and candidate genes involved in ingenol biosynthesis in the medicinal plant *Euphorbia lathyris* L. *Ind. Crops Prod.* **2022**, *184*, 115096. [CrossRef]

92. Wang, F.; Ji, G.; Xu, Z.; Feng, B.; Zhou, Q.; Fan, X.; Wang, T. Metabolomics and Transcriptomics Provide Insights into Anthocyanin Biosynthesis in the Developing Grains of Purple Wheat (*Triticum aestivum* L.). *J. Agric. Food Chem.* **2021**, *69*, 11171–11184. [CrossRef]

93. Pan, J.; Li, Z.; Dai, S.; Ding, H.; Wang, Q.; Li, X.; Ding, G.; Wang, P.; Guan, Y.; Liu, W. Integrative analyses of transcriptomics and metabolomics upon seed germination of foxtail millet in response to salinity. *Sci. Rep.* **2020**, *10*, 13660. [CrossRef]

94. Zhang, S.; Zhan, W.; Sun, A.; Xie, Y.; Han, Z.; Qu, X.; Wang, J.; Zhang, L.; Tian, M.; Pang, X.; et al. Combined transcriptome and metabolome integrated analysis of Acer mandshuricum to reveal candidate genes involved in anthocyanin accumulation. *Sci. Rep.* **2021**, *11*, 23148. [CrossRef]

95. Xu, D.; Lin, H.; Tang, Y.; Huang, L.; Xu, J.; Nian, S.; Zhao, Y. Integration of full-length transcriptomics and targeted metabolomics to identify benzylisoquinoline alkaloid biosynthetic genes in Corydalis yanhusuo. *Hortic. Res.* **2021**, *8*, 16. [CrossRef]

96. Huang, S.; Wang, L.; Wang, Z.; Yang, G.; Xiang, X.; An, Y.; Kan, J. Multiomics strategy reveals the accumulation and biosynthesis of bitter components in Zanthoxylum schinifolium Sieb. et Zucc. *Food Res. Int.* **2022**, *162*, 111964. [CrossRef]

97. Liu, S.; Fang, S.; Liu, C.; Zhao, L.; Cong, B.; Zhang, Z. Transcriptomics Integrated With Metabolomics Reveal the Effects of Ultraviolet-B Radiation on Flavonoid Biosynthesis in Antarctic Moss. *Front. Plant Sci.* **2021**, *12*, 788377. [CrossRef]

98. Yang, J.; Zhang, C.; Li, W.-H.; Zhang, T.-E.; Fan, G.-Z.; Guo, B.-F.; Yao, W.-Y.; Gong, D.-H.; Peng, Q.-L.; Ding, W.-J. Comprehensive Analysis of Transcriptomics and Metabolomics between the Heads and Tails of Angelica Sinensis: Genes Related to Phenylpropanoid Biosynthesis Pathway. *Comb. Chem. High Throughput Screen.* **2021**, *24*, 1417–1427. [CrossRef]

99. Yuan, Y.; Zuo, J.; Zhang, H.; Li, R.; Yu, M.; Liu, S. Integration of Transcriptome and Metabolome Provides New Insights to Flavonoids Biosynthesis in Dendrobium huoshanense. *Front. Plant Sci.* **2022**, *13*, 850090. [CrossRef]

100. Roumani, M.; Le Bot, J.; Boisbrun, M.; Magot, F.; Péré, A.; Robin, C.; Hilliou, F.; Larbat, R. Transcriptomics and Metabolomics Analyses Reveal High Induction of the Phenolamide Pathway in Tomato Plants Attacked by the Leafminer Tuta absoluta. *Metabolites* **2022**, *12*, 484. [CrossRef]

101. Zhang, C.; Chen, J.; Huang, W.; Song, X.; Niu, J. Transcriptomics and Metabolomics Reveal Purine and Phenylpropanoid Metabolism Response to Drought Stress in Dendrobium sinense, an Endemic Orchid Species in Hainan Island. *Front. Genet.* **2021**, *12*, 692702. [CrossRef] [PubMed]

102. Liu, A.; Yuan, K.; Xu, H.; Zhang, Y.; Tian, J.; Li, Q.; Zhu, W.; Ye, H. Proteomic and Metabolomic Revealed Differences in the Distribution and Synthesis Mechanism of Aroma Precursors in Yunyan 87 Tobacco Leaf, Stem, and Root at the Seedling Stage. *ACS Omega* **2022**, *7*, 33295–33306. [CrossRef] [PubMed]

103. Guo, J.; Zhou, Y.J.; Hillwig, M.L.; Shen, Y.; Yang, L.; Wang, Y.; Zhang, X.; Liu, W.; Peters, R.J.; Chen, X.; et al. CYP76AH1 catalyzes turnover of miltiradiene in tanshinones biosynthesis and enables heterologous production of ferruginol in yeasts. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 12108–12113. [CrossRef] [PubMed]

104. Sun, L.; Rai, A.; Rai, M.; Nakamura, M.; Kawano, N.; Yoshimatsu, K.; Suzuki, H.; Kawahara, N.; Saito, K.; Yamazaki, M. Comparative transcriptome analyses of three medicinal Forsythia species and prediction of candidate genes involved in secondary metabolisms. *J. Nat. Med.* **2018**, *72*, 867–881. [CrossRef] [PubMed]

105. Fukusaki, E.; Kobayashi, A. Plant metabolomics: Potential for practical operation. *J. Biosci. Bioeng.* **2005**, *100*, 347–354. [CrossRef]

106. Sumner, L.W.; Lei, Z.; Nikolau, B.J.; Saito, K. Modern plant metabolomics: Advanced natural product gene discoveries, improved technologies, and future prospects. *Nat. Prod. Rep.* **2015**, *32*, 212–229. [CrossRef]

107. Chen, W.; Gong, L.; Guo, Z.; Wang, W.; Zhang, H.; Liu, X.; Yu, S.; Xiong, L.; Luo, J. A Novel Integrated Method for Large-Scale Detection, Identification, and Quantification of Widely Targeted Metabolites: Application in the Study of Rice Metabolomics. *Mol. Plant* **2013**, *6*, 1769–1780. [CrossRef]

108. Dong, X.; Chen, W.; Wang, W.; Zhang, H.; Liu, X.; Luo, J. Comprehensive profiling and natural variation of flavonoids in rice. *J. Integr. Plant Biol.* **2014**, *56*, 876–886. [CrossRef]

109. Nakabayashi, R.; Saito, K. Metabolomics for unknown plant metabolites. *Anal. Bioanal. Chem.* **2013**, *405*, 5005–5011. [CrossRef]

110. Wang, S.; Tu, H.; Wan, J.; Chen, W.; Liu, X.; Luo, J.; Xu, J.; Zhang, H. Spatio-temporal distribution and natural variation of metabolites in citrus fruits. *Food Chem.* **2016**, *199*, 8–17. [CrossRef]

111. Saito, K.; Yonekura-Sakakibara, K.; Nakabayashi, R.; Higashi, Y.; Yamazaki, M.; Tohge, T.; Fernie, A.R. The flavonoid biosynthetic pathway in Arabidopsis: Structural and genetic diversity. *Plant Physiol. Biochem.* **2013**, *72*, 21–34. [CrossRef]

112. Seki, H.; Ohyama, K.; Sawai, S.; Mizutani, M.; Ohnishi, T.; Sudo, H.; Akashi, T.; Aoki, T.; Saito, K.; Muranaka, T. Licorice β-amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 14204–14209. [CrossRef]

113. Naoumkina, M.A.; Modolo, L.V.; Huhman, D.V.; Urbanczyk-Wochniak, E.; Tang, Y.; Sumner, L.W.; Dixon, R.A. Genomic and Coexpression Analyses Predict Multiple Genes Involved in Triterpene Saponin Biosynthesis in Medicago truncatula. *Plant Cell* **2010**, *22*, 850–866. [CrossRef]

114. Asano, T.; Kobayashi, K.; Kashihara, E.; Sudo, H.; Sasaki, R.; Iijima, Y.; Aoki, K.; Shibata, D.; Saito, K.; Yamazaki, M. Suppression of camptothecin biosynthetic genes results in metabolic modification of secondary products in hairy roots of Ophiorrhiza pumila. *Phytochemistry* **2013**, *91*, 128–139. [CrossRef]

115. Martínez-Esteso, M.J.; Martínez-Márquez, A.; Sellés-Marchart, S.; Morante-Carriel, J.A.; Bru-Martínez, R. The role of proteomics in progressing insights into plant secondary metabolism. *Front. Plant Sci.* **2015**, *6*, 504. [CrossRef]

116. Bhattacharyya, D.; Sinha, R.; Ghanta, S.; Chakraborty, A.; Hazra, S.; Chattopadhyay, S. Proteins differentially expressed in elicited cell suspension culture of Podophyllum hexandrum with enhanced podophyllotoxin content. *Proteome Sci.* **2012**, *10*, 34. [CrossRef]

117. Champagne, A.; Rischer, H.; Oksman-Caldentey, K.-M.; Boutry, M. In-depth proteome mining of cultured Catharanthus roseus cells identifies candidate proteins involved in the synthesis and transport of secondary metabolites. *Proteomics* **2012**, *12*, 3536–3547. [CrossRef]

118. Martínez-Esteso, M.J.; Casado-Vela, J.; Sellés-Marchart, S.; Elortza, F.; Pedreño, M.A.; Bru-Martínez, R. iTRAQ-based profiling of grape berry exocarp proteins during ripening using a parallel mass spectrometric method. *Mol. BioSyst.* **2011**, *7*, 749–765. [CrossRef]

119. Martínez-Esteso, M.J.; Sellés-Marchart, S.; Lijavetzky, D.; Pedreño, M.A.; Bru-Martínez, R. A DIGE-based quantitative proteomic analysis of grape berry flesh development and ripening reveals key events in sugar and organic acid metabolism. *J. Exp. Bot.* **2011**, *62*, 2521–2569. [CrossRef]

120. Conesa, A.; Gotz, S.; Garcia-Gomez, J.M.; Terol, J.; Talon, M.; Robles, M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **2005**, *21*, 3674–3676. [CrossRef]

121. Nesvizhskii, A.I. Proteogenomics: Concepts, applications and computational strategies. *Nat. Methods* **2014**, *11*, 1114–1125. [CrossRef] [PubMed]

122. Fancello, L.; Burger, T. An analysis of proteogenomics and how and when transcriptome-informed reduction of protein databases can enhance eukaryotic proteomics. *Genome Biol.* **2022**, *23*, 132. [CrossRef] [PubMed]

123. Van Cutsem, E.; Simonart, G.; Degand, H.; Faber, A.-M.; Morsomme, P.; Boutry, M. Gel-based and gel-free proteomic analysis of Nicotiana tabacum trichomes identifies proteins involved in secondary metabolism and in the (a)biotic stress response. *Proteomics* **2011**, *11*, 440–454. [CrossRef] [PubMed]

124. Zulak, K.G.; Khan, M.F.; Alcantara, J.; Schriemer, D.C.; Facchini, P.J. Plant Defense Responses in Opium Poppy Cell Cultures Revealed by Liquid Chromatography-Tandem Mass Spectrometry Proteomics. *Mol. Cell. Proteomics* **2009**, *8*, 86–98. [CrossRef] [PubMed]

125. Desgagné-Penix, I.; Khan, M.F.; Schriemer, D.C.; Cram, D.; Nowak, J.; Facchini, P.J. Integration of deep transcriptome and proteome analyses reveals the components of alkaloid metabolism in opium poppy cell cultures. *BMC Plant Biol.* **2010**, *10*, 252. [CrossRef]

126. Corchete, P.; Bru, R. Proteome alterations monitored by DIGE analysis in Silybum marianum cell cultures elicited with methyl jasmonate and methyl B cyclodextrin. *J. Proteom.* **2013**, *85*, 99–108. [CrossRef]

127. Robbins, M.L.; Roy, A.; Wang, P.-H.; Gaffoor, I.; Sekhon, R.S.; de O. Buanafina, M.M.; Rohila, J.S.; Chopra, S. Comparative proteomics analysis by DIGE and iTRAQ provides insight into the regulation of phenylpropanoids in maize. *J. Proteom.* **2013**, *93*, 254–275. [CrossRef]

128. Oldham, J.T.; Hincapie, M.; Rejtar, T.; Wall, P.K.; Carlson, J.E.; Lee-Parsons, C.W.T. Shotgun Proteomic Analysis of Yeast-Elicited California Poppy (Eschscholzia californica) Suspension Cultures Producing Enhanced Levels of Benzophenanthridine Alkaloids. *J. Proteome Res.* **2010**, *9*, 4337–4345. [CrossRef]

129. Yang, D.; Du, X.; Yang, Z.; Liang, Z.; Guo, Z.; Liu, Y. Transcriptomics, proteomics, and metabolomics to reveal mechanisms underlying plant secondary metabolism. *Eng. Life Sci.* **2014**, *14*, 456–466. [CrossRef]

130. Saito, K. Phytochemical genomics—A new trend. *Curr. Opin. Plant Biol.* **2013**, *16*, 373–380. [CrossRef]

131. Florea, L.; Song, L.; Salzberg, S.L. Thousands of exon skipping events differentiate among splicing patterns in sixteen human tissues. *F1000Research* **2013**, *2*, 188. [CrossRef]

132. el Bouhaddani, S.; Houwing-Duistermaat, J.; Salo, P.; Perola, M.; Jongbloed, G.; Uh, H.-W. Evaluation of O2PLS in Omics data integration. *BMC Bioinform.* **2016**, *17*, S11. [CrossRef]

133. Dang, T.-T.T.; Facchini, P.J. Characterization of Three O -Methyltransferases Involved in Noscapine Biosynthesis in Opium Poppy. *Plant Physiol.* **2012**, *159*, 618–631. [CrossRef]

134. Sperling, P.; Lee, M.; Girke, T.; Zähringer, U.; Stymne, S.; Heinz, E. A bifunctional Δ6-fatty acyl acetylenase/desaturase from the moss Ceratodon purpureus. *Eur. J. Biochem.* **2000**, *267*, 3801–3811. [CrossRef]

135. Miao, Y.; Luo, D.; Zhao, T.; Du, H.; Liu, Z.; Xu, Z.; Guo, L.; Chen, C.; Peng, S.; Li, J.X.; et al. Genome sequencing reveals chromosome fusion and extensive expansion of genes related to secondary metabolism in Artemisia argyi. *Plant Biotechnol. J.* **2022**, *20*, 1902–1915. [CrossRef]

136. Luo, J. Metabolite-based genome-wide association studies in plants. *Curr. Opin. Plant Biol.* **2015**, *24*, 31–38. [CrossRef]

137. Chen, W.; Gao, Y.; Xie, W.; Gong, L.; Lu, K.; Wang, W.; Li, Y.; Liu, X.; Zhang, H.; Dong, H.; et al. Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nat. Genet.* **2014**, *46*, 714–721. [CrossRef]

138. Wen, W.; Li, D.; Li, X.; Gao, Y.; Li, W.; Li, H.; Liu, J.; Liu, H.; Chen, W.; Luo, J.; et al. Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights. *Nat. Commun.* **2014**, *5*, 3438. [CrossRef]

139. Suhre, K.; Shin, S.-Y.; Petersen, A.-K.; Mohney, R.P.; Meredith, D.; Wägele, B.; Altmaier, E.; Deloukas, P.; Erdmann, J.; Grundberg, E.; et al. Human metabolic individuality in biomedical and pharmaceutical research. *Nature* **2011**, *477*, 54–60. [CrossRef]

140. Verslues, P.E.; Lasky, J.R.; Juenger, T.E.; Liu, T.-W.; Kumar, M.N. Genome-Wide Association Mapping Combined with Reverse Genetics Identifies New Effectors of Low Water Potential-Induced Proline Accumulation in Arabidopsis. *Plant Physiol.* **2014**, *164*, 144–159. [CrossRef]

141. Riedelsheimer, C.; Lisec, J.; Czedik-Eysenberg, A.; Sulpice, R.; Flis, A.; Grieder, C.; Altmann, T.; Stitt, M.; Willmitzer, L.; Melchinger, A.E. Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 8872–8877. [CrossRef] [PubMed]

142. Chan, E.K.F.; Rowe, H.C.; Corwin, J.A.; Joseph, B.; Kliebenstein, D.J. Combining Genome-Wide Association Mapping and Transcriptional Networks to Identify Novel Genes Controlling Glucosinolates in Arabidopsis thaliana. *PLoS Biol.* **2011**, *9*, e1001125. [CrossRef] [PubMed]

143. Dong, X.; Gao, Y.; Chen, W.; Wang, W.; Gong, L.; Liu, X.; Luo, J. Spatiotemporal Distribution of Phenolamides and the Genetics of Natural Variation of Hydroxycinnamoyl Spermidine in Rice. *Mol. Plant* **2015**, *8*, 111–121. [CrossRef] [PubMed]

144. Gong, L.; Chen, W.; Gao, Y.; Liu, X.; Zhang, H.; Xu, C.; Yu, S.; Zhang, Q.; Luo, J. Genetic analysis of the metabolome exemplified using a rice population. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 20320–20325. [CrossRef] [PubMed]

145. Chan, E.K.F.; Rowe, H.C.; Kliebenstein, D.J. Understanding the Evolution of Defense Metabolites in Arabidopsis thaliana Using Genome-wide Association Mapping. *Genetics* **2010**, *185*, 991–1007. [CrossRef]

146. Würschum, T. Mapping QTL for agronomic traits in breeding populations. *Theor. Appl. Genet.* **2012**, *125*, 201–210. [CrossRef]

147. Kamunya, S.M.; Wachira, F.N.; Pathak, R.S.; Korir, R.; Sharma, V.; Kumar, R.; Bhardwaj, P.; Chalo, R.; Ahuja, P.S.; Sharma, R.K. Genomic mapping and testing for quantitative trait loci in tea (Camellia sinensis (L.) O. Kuntze). *Tree Genet. Genomes* **2010**, *6*, 915–929. [CrossRef]

148. Ma, J.-Q.; Jin, J.-Q.; Yao, M.-Z.; Ma, C.-L.; Xu, Y.-X.; Hao, W.-J.; Chen, L. Quantitative Trait Loci Mapping for Theobromine and Caffeine Contents in Tea Plant (Camellia sinensis). *J. Agric. Food Chem.* **2018**, *66*, 13321–13327. [CrossRef]

149. Ma, J.-Q.; Yao, M.-Z.; Ma, C.-L.; Wang, X.-C.; Jin, J.-Q.; Wang, X.-M.; Chen, L. Construction of a SSR-Based Genetic Map and Identification of QTLs for Catechins Content in Tea Plant (Camellia sinensis). *PLoS ONE* **2014**, *9*, e93131. [CrossRef]

150. Tan, L.-Q.; Wang, L.-Y.; Xu, L.-Y.; Wu, L.-Y.; Peng, M.; Zhang, C.-C.; Wei, K.; Bai, P.-X.; Li, H.-L.; Cheng, H.; et al. SSR-based genetic mapping and QTL analysis for timing of spring bud flush, young shoot color, and mature leaf size in tea plant (Camellia sinensis). *Tree Genet. Genomes* **2016**, *12*, 52. [CrossRef]

151. Xu, L.-Y.; Wang, L.-Y.; Wei, K.; Tan, L.-Q.; Su, J.-J.; Cheng, H. High-density SNP linkage map construction and QTL mapping for flavonoid-related traits in a tea plant (Camellia sinensis) using 2b-RAD sequencing. *BMC Genom.* **2018**, *19*, 955. [CrossRef]

152. McCallum, J.; Clarke, A.; Pither-Joyce, M.; Shaw, M.; Butler, R.; Brash, D.; Scheffer, J.; Sims, I.; van Heusden, S.; Shigyo, M.; et al. Genetic mapping of a major gene affecting onion bulb fructan content. *Theor. Appl. Genet.* **2006**, *112*, 958–967. [CrossRef]

153. Zhang, M.; Liu, Y.-H.; Xu, W.; Smith, C.W.; Murray, S.C.; Zhang, H.-B. Analysis of the genes controlling three quantitative traits in three diverse plant species reveals the molecular basis of quantitative traits. *Sci. Rep.* **2020**, *10*, 10074. [CrossRef]

154. Fu, J.; Swertz, M.A.; Keurentjes, J.J.; Jansen, R.C. MetaNetwork: A computational protocol for the genetic study of metabolic networks. *Nat. Protoc.* **2007**, *2*, 685–694. [CrossRef]

155. Gao, Y.; Du, L.; Ma, Q.; Yuan, Y.; Liu, J.; Song, H.; Feng, B. Conjunctive Analyses of Bulk Segregant Analysis Sequencing and Bulk Segregant RNA Sequencing to Identify Candidate Genes Controlling Spikelet Sterility of Foxtail Millet. *Front. Plant Sci.* **2022**, *13*, 842336. [CrossRef]

156. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [CrossRef]

157. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* **2012**, *6*, 80–92. [CrossRef]

158. Fekih, R.; Takagi, H.; Tamiru, M.; Abe, A.; Natsume, S.; Yaegashi, H.; Sharma, S.; Sharma, S.; Kanzaki, H.; Matsumura, H.; et al. MutMap+: Genetic Mapping and Mutant Identification without Crossing in Rice. *PLoS ONE* **2013**, *8*, e68529. [CrossRef]

159. Abe, A.; Kosugi, S.; Yoshida, K.; Natsume, S.; Takagi, H.; Kanzaki, H.; Matsumura, H.; Yoshida, K.; Mitsuoka, C.; Tamiru, M.; et al. Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat. Biotechnol.* **2012**, *30*, 174–178. [CrossRef]

160. Jin, J.-Q.; Ma, J.-Q.; Yao, M.-Z.; Ma, C.-L.; Chen, L. Functional natural allelic variants of flavonoid 3′,5′-hydroxylase gene governing catechin traits in tea plant and its relatives. *Planta* **2017**, *245*, 523–538. [CrossRef]

161. Schauer, N.; Semel, Y.; Roessner, U.; Gur, A.; Balbo, I.; Carrari, F.; Pleban, T.; Perez-Melis, A.; Bruedigam, C.; Kopka, J.; et al. Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat. Biotechnol.* **2006**, *24*, 447–454. [CrossRef] [PubMed]

162. Keurentjes, J.J.B.; Fu, J.; de Vos, C.H.R.; Lommen, A.; Hall, R.D.; Bino, R.J.; van der Plas, L.H.W.; Jansen, R.C.; Vreugdenhil, D.; Koornneef, M. The genetics of plant metabolism. *Nat. Genet.* **2006**, *38*, 842–849. [CrossRef] [PubMed]

163. James, G.V.; Patel, V.; Nordström, K.J.; Klasen, J.R.; Salomé, P.A.; Weigel, D.; Schneeberger, K. User guide for mapping-by-sequencing in Arabidopsis. *Genome Biol.* **2013**, *14*, R61. [CrossRef] [PubMed]

164. Nordström, K.J.V.; Albani, M.C.; James, G.V.; Gutjahr, C.; Hartwig, B.; Turck, F.; Paszkowski, U.; Coupland, G.; Schneeberger, K. Mutation identification by direct comparison of whole-genome sequencing data from mutant and wild-type individuals using k-mers. *Nat. Biotechnol.* **2013**, *31*, 325–330. [CrossRef] [PubMed]

165. Weber, A.P.M. Discovering new biology through RNA-Seq. *Plant Physiol.* **2015**, *169*, 1524–1531. [CrossRef]

166. Han, Z.; Ahsan, M.; Adil, M.F.; Chen, X.; Nazir, M.M.; Shamsi, I.H.; Zeng, F.; Zhang, G. Identification of the gene network modules highly associated with the synthesis of phenolics compounds in barley by transcriptome and metabolome analysis. *Food Chem.* **2020**, *323*, 126862. [CrossRef]

167. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [CrossRef]

168. Langfelder, P.; Horvath, S. Fast R Functions for Robust Correlations and Hierarchical Clustering. *J. Stat. Softw.* **2012**, *46*, i11. [CrossRef]

169. Tai, Y.; Liu, C.; Yu, S.; Yang, H.; Sun, J.; Guo, C.; Huang, B.; Liu, Z.; Yuan, Y.; Xia, E.; et al. Gene co-expression network analysis reveals coordinated regulation of three characteristic secondary biosynthetic pathways in tea plant (*Camellia sinensis*). *BMC Genom.* **2018**, *19*, 616. [CrossRef]

170. Feng, C.; Feng, C.; Lin, X.; Liu, S.; Li, Y.; Kang, M. A chromosome-level genome assembly provides insights into ascorbic acid accumulation and fruit softening in guava (*Psidium guajava*). *Plant Biotechnol. J.* **2021**, *19*, 717–730. [CrossRef]

171. Zhang, X.; Zhang, G. Transcriptome and metabolome profiling reveal the regulatory mechanism of protein accumulation in inferior grains of indica-japonica rice hybrids. *Curr. Plant Biol.* **2021**, *28*, 100226. [CrossRef]

172. Liu, Y.; Wang, S.; Li, L.; Yang, T.; Dong, S.; Wei, T.; Wu, S.; Liu, Y.; Gong, Y.; Feng, X.; et al. The Cycas genome and the early evolution of seed plants. *Nat. Plants* **2022**, *8*, 389–401. [CrossRef]

173. Schmutz, J.; Cannon, S.B.; Schlueter, J.; Ma, J.; Mitros, T.; Nelson, W.; Hyten, D.L.; Song, Q.; Thelen, J.J.; Cheng, J.; et al. Genome sequence of the palaeopolyploid soybean. *Nature* **2010**, *463*, 178–183. [CrossRef]

174. The International Brachypodium Initiative. Genome sequencing and analysis of the model grass Brachypodium distachyon. *Nature* **2010**, *463*, 763–768. [CrossRef]

175. Chan, A.P.; Crabtree, J.; Zhao, Q.; Lorenzi, H.; Orvis, J.; Puiu, D.; Melake-Berhan, A.; Jones, K.M.; Redman, J.; Chen, G.; et al. Draft genome sequence of the oilseed species Ricinus communis. *Nat. Biotechnol.* **2010**, *28*, 951–956. [CrossRef]

176. Schnable, P.S.; Ware, D.; Fulton, R.S.; Stein, J.C.; Wei, F.; Pasternak, S.; Liang, C.; Zhang, J.; Fulton, L.; Graves, T.A.; et al. The B73 Maize Genome: Complexity, Diversity, and Dynamics. *Science* **2009**, *326*, 1112–1115. [CrossRef]

177. Paterson, A.H.; Bowers, J.E.; Bruggmann, R.; Dubchak, I.; Grimwood, J.; Gundlach, H.; Haberer, G.; Hellsten, U.; Mitros, T.; Poliakov, A.; et al. The Sorghum bicolor genome and the diversification of grasses. *Nature* **2009**, *457*, 551–556. [CrossRef]

178. Rensing, S.A.; Lang, D.; Zimmer, A.D.; Terry, A.; Salamov, A.; Shapiro, H.; Nishiyama, T.; Perroud, P.-F.; Lindquist, E.A.; Kamisugi, Y.; et al. The Physcomitrella Genome Reveals Evolutionary Insights into the Conquest of Land by Plants. *Science* **2008**, *319*, 64–69. [CrossRef]

179. Borrell, B. Papaya genome project bears fruit. *Nature* **2008**. [CrossRef]

180. The French–Italian Public Consortium for Grapevine Genome Characterization. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **2007**, *449*, 463–467. [CrossRef]

181. Tuskan, G.A.; DiFazio, S.; Jansson, S.; Bohlmann, J.; Grigoriev, I.; Hellsten, U.; Putnam, N.; Ralph, S.; Rombauts, S.; Salamov, A.; et al. The Genome of Black Cottonwood, Populus trichocarpa (Torr. & Gray). *Science* **2006**, *313*, 1596–1604. [CrossRef] [PubMed]

182. Yu, J.; Hu, S.; Wang, J.; Wong, G.K.-S.; Li, S.; Liu, B.; Deng, Y.; Dai, L.; Zhou, Y.; Zhang, X.; et al. A Draft Sequence of the Rice Genome (*Oryza sativa* L. ssp. *indica*). *Science* **2002**, *296*, 79–92. [CrossRef] [PubMed]

183. Goff, S.A.; Ricke, D.; Lan, T.-H.; Presting, G.; Wang, R.; Dunn, M.; Glazebrook, J.; Sessions, A.; Oeller, P.; Varma, H.; et al. A Draft Sequence of the Rice Genome (*Oryza sativa* L. ssp. *japonica*). *Science* **2002**, *296*, 92–100. [CrossRef] [PubMed]

184. Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* **2000**, *408*, 796–815. [CrossRef] [PubMed]

185. Weitemier, K.; Straub, S.C.K.; Fishbein, M.; Bailey, C.D.; Cronn, R.C.; Liston, A. A draft genome and transcriptome of common milkweed (*Asclepias syriaca*) as resources for evolutionary, ecological, and molecular studies in milkweeds and Apocynaceae. *PeerJ* **2019**, *7*, e7649. [CrossRef]

186. Kim, N.-H.; Jayakodi, M.; Lee, S.-C.; Choi, B.-S.; Jang, W.; Lee, J.; Kim, H.H.; Waminal, N.E.; Lakshmanan, M.; van Nguyen, B.; et al. Genome and evolution of the shade-requiring medicinal herb Panax ginseng. *Plant Biotechnol. J.* **2018**, *16*, 1904–1917. [CrossRef]

187. Hirakawa, H.; Sumitomo, K.; Hisamatsu, T.; Nagano, S.; Shirasawa, K.; Higuchi, Y.; Kusaba, M.; Koshioka, M.; Nakano, Y.; Yagi, M.; et al. De novo whole-genome assembly in Chrysanthemum seticuspe, a model species of Chrysanthemums, and its application to genetic and gene discovery analysis. *DNA Res.* **2019**, *26*, 195–203. [CrossRef]

188. Yang, Q.; Bi, H.; Yang, W.; Li, T.; Jiang, J.; Zhang, L.; Liu, J.; Hu, Q. The Genome Sequence of Alpine Megacarpaea delavayi Identifies Species-Specific Whole-Genome Duplication. *Front. Genet.* **2020**, *11*, 812. [CrossRef]

189. Kumari, P.; Singh, K.P.; Rai, P.K. Draft genome of multiple resistance donor plant Sinapis alba: An insight into SSRs, annotations and phylogenetics. *PLoS ONE* **2020**, *15*, e0231002. [CrossRef]

190. Patil, A.B.; Shinde, S.S.; Raghavendra, S.; Satish, B.N.; Kushalappa, C.G.; Vijay, N. The genome sequence of Mesua ferrea and comparative demographic histories of forest trees. *Gene* **2021**, *769*, 145214. [CrossRef]

191. Xie, D.; Xu, Y.; Wang, J.; Liu, W.; Zhou, Q.; Luo, S.; Huang, W.; He, X.; Li, Q.; Peng, Q.; et al. The wax gourd genomes offer insights into the genetic diversity and ancestral cucurbit karyotype. *Nat. Commun.* **2019**, *10*, 5158. [CrossRef] [PubMed]

192. Tamiru, M.; Natsume, S.; Takagi, H.; White, B.; Yaegashi, H.; Shimizu, M.; Yoshida, K.; Uemura, A.; Oikawa, K.; Abe, A.; et al. Genome sequencing of the staple food crop white Guinea yam enables the development of a molecular marker for sex determination. *BMC Biol.* **2017**, *15*, 86. [CrossRef]

193. Chang, Y.; Liu, H.; Liu, M.; Liao, X.; Sahu, S.K.; Fu, Y.; Song, B.; Cheng, S.; Kariba, R.; Muthemba, S.; et al. The draft genomes of five agriculturally important African orphan crops. *Gigascience* **2019**, *8*, giy152. [CrossRef]

194. Griesmann, M.; Chang, Y.; Liu, X.; Song, Y.; Haberer, G.; Crook, M.B.; Billault-Penneteau, B.; Lauressergues, D.; Keller, J.; Imanishi, L.; et al. Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. *Science* **2018**, *361*, eaat1743. [CrossRef]

195. Wan, T.; Liu, Z.-M.; Li, L.-F.; Leitch, A.R.; Leitch, I.J.; Lohaus, R.; Liu, Z.-J.; Xin, H.-P.; Gong, Y.-B.; Liu, Y.; et al. A genome for gnetophytes and early evolution of seed plants. *Nat. Plants* **2018**, *4*, 82–89. [CrossRef]

196. Nashima, K.; Shirasawa, K.; Ghelfi, A.; Hirakawa, H.; Isobe, S.; Suyama, T.; Wada, T.; Kurokura, T.; Uemachi, T.; Azuma, M.; et al. Genome sequence of Hydrangea macrophylla and its application in analysis of the double flower phenotype. *DNA Res.* **2021**, *28*, dsaa026. [CrossRef]

197. Stevens, K.A.; Woeste, K.; Chakraborty, S.; Crepeau, M.W.; Leslie, C.A.; Martínez-García, P.J.; Puiu, D.; Romero-Severson, J.; Coggeshall, M.; Dandekar, A.M.; et al. Genomic Variation Among and Within Six Juglans Species. *G3 Genes Genome Genet.* **2018**, *8*, 2153–2165. [CrossRef]

198. Malli, R.P.N.; Adal, A.M.; Sarker, L.S.; Liang, P.; Mahmoud, S.S. De novo sequencing of the Lavandula angustifolia genome reveals highly duplicated and optimized features for essential oil production. *Planta* **2019**, *249*, 251–256. [CrossRef]

199. Bornowski, N.; Hamilton, J.P.; Liao, P.; Wood, J.C.; Dudareva, N.; Buell, C.R. Genome sequencing of four culinary herbs reveals terpenoid genes underlying chemodiversity in the Nepetoideae. *DNA Res.* **2020**, *27*, dsaa016. [CrossRef]

200. He, Y.; Peng, F.; Deng, C.; Xiong, L.; Huang, Z.; Zhang, R.; Liu, M.; Peng, C. Building an octaploid genome and transcriptome of the medicinal plant Pogostemon cablin from Lamiales. *Sci. Data* **2018**, *5*, 180274. [CrossRef]

201. Gui, S.; Peng, J.; Wang, X.; Wu, Z.; Cao, R.; Salse, J.; Zhang, H.; Zhu, Z.; Xia, Q.; Quan, Z.; et al. Improving Nelumbo nucifera genome assemblies using high-resolution genetic maps and BioNano genome mapping reveals ancient chromosome rearrangements. *Plant J.* **2018**, *94*, 721–734. [CrossRef] [PubMed]

202. Yang, X.; Yue, Y.; Li, H.; Ding, W.; Chen, G.; Shi, T.; Chen, J.; Park, M.S.; Chen, F.; Wang, L. The chromosome-level quality genome provides insights into the evolution of the biosynthesis genes for aroma compounds of Osmanthus fragrans. *Hortic. Res.* **2018**, *5*, 72. [CrossRef] [PubMed]

203. Yuan, Y.; Jin, X.; Liu, J.; Zhao, X.; Zhou, J.; Wang, X.; Wang, D.; Lai, C.; Xu, W.; Huang, J.; et al. The Gastrodia elata genome provides insights into plant adaptation to heterotrophy. *Nat. Commun.* **2018**, *9*, 1615. [CrossRef] [PubMed]

204. Neller, K.C.M.; Diaz, C.A.; Platts, A.E.; Hudak, K.A. De novo Assembly of the Pokeweed Genome Provides Insight Into Pokeweed Antiviral Protein (PAP) Gene Expression. *Front. Plant Sci.* **2019**, *10*, 1002. [CrossRef] [PubMed]

205. Mosca, E.; Cruz, F.; Gómez-Garrido, J.; Bianco, L.; Rellstab, C.; Brodbeck, S.; Csilléry, K.; Fady, B.; Fladung, M.; Fussi, B.; et al. A Reference Genome Sequence for the European Silver Fir (*Abies alba* Mill.): A Community-Generated Genomic Resource. *G3 Genes Genomes Genet.* **2019**, *9*, 2039–2049. [CrossRef]

206. Soyturk, A.; Sen, F.; Uncu, A.T.; Celik, I.; Uncu, A.O. De novo assembly and characterization of the first draft genome of quince (Cydonia oblonga Mill.). *Sci. Rep.* **2021**, *11*, 3818. [CrossRef]

207. Wang, L.; He, F.; Huang, Y.; He, J.; Yang, S.; Zeng, J.; Deng, C.; Jiang, X.; Fang, Y.; Wen, S.; et al. Genome of Wild Mandarin and Domestication History of Mandarin. *Mol. Plant* **2018**, *11*, 1024–1037. [CrossRef]

208. Mahesh, H.B.; Subba, P.; Advani, J.; Shirke, M.D.; Loganathan, R.M.; Chandana, S.L.; Shilpa, S.; Chatterjee, O.; Pinto, S.M.; Prasad, T.S.K.; et al. Multi-Omics Driven Assembly and Annotation of the Sandalwood (*Santalum album*) Genome. *Plant Physiol.* **2018**, *176*, 2772–2788. [CrossRef]

209. Nong, W.; Law, S.T.S.; Wong, A.Y.P.; Baril, T.; Swale, T.; Chu, L.M.; Hayward, A.; Lau, D.T.W.; Hui, J.H.L. Chromosomal-level reference genome of the incense tree Aquilaria sinensis. *Mol. Ecol. Resour.* **2020**, *20*, 971–979. [CrossRef]

210. Luan, M.-B.; Jian, J.-B.; Chen, P.; Chen, J.-H.; Chen, J.-H.; Gao, Q.; Gao, G.; Zhou, J.-H.; Chen, K.-M.; Guang, X.-M.; et al. Draft genome sequence of ramie, *Boehmeria nivea* (L.) Gaudich. *Mol. Ecol. Resour.* **2018**, *18*, 639–645. [CrossRef]

211. Zhang, Y.; Zheng, L.; Zheng, Y.; Zhou, C.; Huang, P.; Xiao, X.; Zhao, Y.; Hao, X.; Hu, Z.; Chen, Q.; et al. Assembly and Annotation of a Draft Genome of the Medicinal Plant Polygonum cuspidatum. *Front. Plant Sci.* **2019**, *10*, 1274. [CrossRef]

212. Ding, X.; Mei, W.; Huang, S.; Wang, H.; Zhu, J.; Hu, W.; Ding, Z.; Tie, W.; Peng, S.; Dai, H. Genome survey sequencing for the characterization of genetic background of Dracaena cambodiana and its defense response during dragon's blood formation. *PLoS ONE* **2018**, *13*, e0209258. [CrossRef]

213. Hunt, S.P.; Jarvis, D.E.; Larsen, D.J.; Mosyakin, S.L.; Kolano, B.A.; Jackson, E.W.; Martin, S.L.; Jellen, E.N.; Maughan, P.J. A Chromosome-Scale Assembly of the Garden Orach (Atriplex hortensis L.) Genome Using Oxford Nanopore Sequencing. *Front. Plant Sci.* **2020**, *11*, 624. [CrossRef]

214. Wang, P.; Luo, Y.; Huang, J.; Gao, S.; Zhu, G.; Dang, Z.; Gai, J.; Yang, M.; Zhu, M.; Zhang, H.; et al. The genome evolution and domestication of tropical fruit mango. *Genome Biol.* **2020**, *21*, 60. [CrossRef]

215. Zeng, L.; Tu, X.-L.; Dai, H.; Han, F.-M.; Lu, B.-S.; Wang, M.-S.; Nanaei, H.A.; Tajabadipour, A.; Mansouri, M.; Li, X.-L.; et al. Whole genomes and transcriptomes reveal adaptation and domestication of pistachio. *Genome Biol.* **2019**, *20*, 79. [CrossRef]

216. Strijk, J.S.; Hinsinger, D.D.; Roeder, M.M.; Chatrou, L.W.; Couvreur, T.L.P.; Erkens, R.H.J.; Sauquet, H.; Pirie, M.D.; Thomas, D.C.; Cao, K. Chromosome-level reference genome of the soursop (*Annona muricata*): A new resource for Magnoliid research and tropical pomology. *Mol. Ecol. Resour.* **2021**, *21*, 1608–1619. [CrossRef]

217. Song, X.; Sun, P.; Yuan, J.; Gong, K.; Li, N.; Meng, F.; Zhang, Z.; Li, X.; Hu, J.; Wang, J.; et al. The celery genome sequence reveals sequential paleo-polyploidizations, karyotype evolution and resistance gene reduction in apiales. *Plant Biotechnol. J.* **2021**, *19*, 731–744. [CrossRef]

218. Song, X.; Wang, J.; Li, N.; Yu, J.; Meng, F.; Wei, C.; Liu, C.; Chen, W.; Nie, F.; Zhang, Z.; et al. Deciphering the high-quality genome sequence of coriander that causes controversial feelings. *Plant Biotechnol. J.* **2020**, *18*, 1444–1456. [CrossRef]

219. Li, S.-F.; Wang, J.; Dong, R.; Zhu, H.-W.; Lan, L.-N.; Zhang, Y.-L.; Li, N.; Deng, C.-L.; Gao, W.-J. Chromosome-level genome assembly, annotation and evolutionary analysis of the ornamental plant Asparagus setaceus. *Hortic. Res.* **2020**, *7*, 48. [CrossRef]

220. Jaiswal, S.K.; Mahajan, S.; Chakraborty, A.; Kumar, S.; Sharma, V.K. The genome sequence of Aloe vera reveals adaptive evolution of drought tolerance mechanisms. *iScience* **2021**, *24*, 102079. [CrossRef]

221. Laforest, M.; Martin, S.L.; Bisaillon, K.; Soufiane, B.; Meloche, S.; Page, E. A chromosome-scale draft sequence of the Canada fleabane genome. *Pest Manag. Sci.* **2020**, *76*, 2158–2169. [CrossRef] [PubMed]

222. He, S.; Dong, X.; Zhang, G.; Fan, W.; Duan, S.; Shi, H.; Li, D.; Li, R.; Chen, G.; Long, G.; et al. High quality genome of Erigeron breviscapus provides a reference for herbal plants in Asteraceae. *Mol. Ecol. Resour.* **2021**, 21, 153–169. [CrossRef] [PubMed]

223. Liu, B.; Yan, J.; Li, W.; Yin, L.; Li, P.; Yu, H.; Xing, L.; Cai, M.; Wang, H.; Zhao, M.; et al. Mikania micrantha genome provides insights into the molecular mechanism of rapid growth. *Nat. Commun.* **2020**, 11, 340. [CrossRef] [PubMed]

224. Chen, S.; Wang, Y.; Yu, L.; Zheng, T.; Wang, S.; Yue, Z.; Jiang, J.; Kumari, S.; Zheng, C.; Tang, H.; et al. Genome sequence and evolution of Betula platyphylla. *Hortic. Res.* **2021**, 8, 37. [CrossRef] [PubMed]

225. Li, Y.; Sun, P.; Lu, Z.; Chen, J.; Wang, Z.; Du, X.; Zheng, Z.; Wu, Y.; Hu, H.; Yang, J.; et al. The Corylus mandshurica genome provides insights into the evolution of Betulaceae genomes and hazelnut breeding. *Hortic. Res.* **2021**, 8, 54. [CrossRef]

226. Züst, T.; Strickler, S.R.; Powell, A.F.; Mabry, M.E.; An, H.; Mirzaei, M.; York, T.; Holland, C.K.; Kumar, P.; Erb, M.; et al. Independent evolution of ancestral and novel defenses in a genus of toxic plants (Erysimum, Brassicaceae). *eLife* **2020**, 9, e51712. [CrossRef]

227. Kang, M.; Wu, H.; Yang, Q.; Huang, L.; Hu, Q.; Ma, T.; Li, Z.; Liu, J. A chromosome-scale genome assembly of Isatis indigotica, an important medicinal plant used in traditional Chinese medicine. *Hortic. Res.* **2020**, 7, 18. [CrossRef]

228. Shang, J.; Tian, J.; Cheng, H.; Yan, Q.; Li, L.; Jamal, A.; Xu, Z.; Xiang, L.; Saski, C.A.; Jin, S.; et al. The chromosome-level wintersweet (Chimonanthus praecox) genome provides insights into floral scent biosynthesis and flowering in winter. *Genome Biol.* **2020**, 21, 200. [CrossRef]

229. Pu, X.; Li, Z.; Tian, Y.; Gao, R.; Hao, L.; Hu, Y.; He, C.; Sun, W.; Xu, M.; Peters, R.J.; et al. The honeysuckle genome provides insight into the molecular mechanism of carotenoid metabolism underlying dynamic flower coloration. *New Phytol.* **2020**, 227, 930–943. [CrossRef]

230. Ye, G.; Zhang, H.; Chen, B.; Nie, S.; Liu, H.; Gao, W.; Wang, H.; Gao, Y.; Gu, L. De novo genome assembly of the stress tolerant forest species Casuarina equisetifolia provides insight into secondary growth. *Plant J.* **2019**, 97, 779–794. [CrossRef]

231. Patterson, E.L.; Saski, C.A.; Sloan, D.B.; Tranel, P.J.; Westra, P.; Gaines, T.A. The Draft Genome of Kochia scoparia and the Mechanism of Glyphosate Resistance via Transposon-Mediated EPSPS Tandem Gene Duplication. *Genome Biol. Evol.* **2019**, 11, 2927–2940. [CrossRef]

232. Sun, G.; Xu, Y.; Liu, H.; Sun, T.; Zhang, J.; Hettenhausen, C.; Shen, G.; Qi, J.; Qin, Y.; Li, J.; et al. Large-scale gene losses underlie the genome evolution of parasitic plant Cuscuta australis. *Nat. Commun.* **2018**, 9, 2683. [CrossRef]

233. Vogel, A.; Schwacke, R.; Denton, A.K.; Usadel, B.; Hollmann, J.; Fischer, K.; Bolger, A.; Schmidt, M.H.-W.; Bolger, M.E.; Gundlach, H.; et al. Footprints of parasitism in the genome of the parasitic flowering plant Cuscuta campestris. *Nat. Commun.* **2018**, 9, 2515. [CrossRef]

234. Barrera-Redondo, J.; Ibarra-Laclette, E.; Vázquez-Lobo, A.; Gutiérrez-Guerrero, Y.T.; Sánchez de la Vega, G.; Piñero, D.; Montes-Hernández, S.; Lira-Saade, R.; Eguiarte, L.E. The Genome of Cucurbita argyrosperma (Silver-Seed Gourd) Reveals Faster Rates of Protein-Coding Gene and Long Noncoding RNA Turnover and Neofunctionalization within Cucurbita. *Mol. Plant* **2019**, 12, 506–520. [CrossRef]

235. Wu, H.; Zhao, G.; Gong, H.; Li, J.; Luo, C.; He, X.; Luo, S.; Zheng, X.; Liu, X.; Guo, J.; et al. A high-quality sponge gourd (*Luffa cylindrica*) genome. *Hortic. Res.* **2020**, 7, 128. [CrossRef]

236. Matsumura, H.; Hsiao, M.-C.; Lin, Y.-P.; Toyoda, A.; Taniai, N.; Tarora, K.; Urasaki, N.; Anand, S.S.; Dhillon, N.P.S.; Schafleitner, R.; et al. Long-read bitter gourd (*Momordica charantia*) genome and the genomic architecture of nonclassic domestication. *Proc. Natl. Acad. Sci. USA* **2020**, 117, 14543–14551. [CrossRef]

237. Fu, A.; Wang, Q.; Mu, J.; Ma, L.; Wen, C.; Zhao, X.; Gao, L.; Li, J.; Shi, K.; Wang, Y.; et al. Combined genomic, transcriptomic, and metabolomic analyses provide insights into chayote (*Sechium edule*) evolution and fruit development. *Hortic. Res.* **2021**, 8, 35. [CrossRef]

238. Xia, M.; Han, X.; He, H.; Yu, R.; Zhen, G.; Jia, X.; Cheng, B.; Deng, X.W. Improved de novo genome assembly and analysis of the Chinese cucurbit Siraitia grosvenorii, also known as monk fruit or luo-han-guo. *Gigascience* **2018**, 7, giy067. [CrossRef]

239. Ma, L.; Wang, Q.; Mu, J.; Fu, A.; Wen, C.; Zhao, X.; Gao, L.; Li, J.; Shi, K.; Wang, Y.; et al. The genome and transcriptome analysis of snake gourd provide insights into its evolution and fruit development and ripening. *Hortic. Res.* **2020**, 7, 199. [CrossRef]

240. Siadjeu, C.; Pucker, B.; Viehöver, P.; Albach, D.C.; Weisshaar, B. High Contiguity de novo Genome Sequence Assembly of Trifoliate Yam (*Dioscorea dumetorum*) Using Long Read Sequencing. *Genes* **2020**, 11, 274. [CrossRef]

241. Yang, F.-S.; Nie, S.; Liu, H.; Shi, T.-L.; Tian, X.-C.; Zhou, S.-S.; Bao, Y.-T.; Jia, K.-H.; Guo, J.-F.; Zhao, W.; et al. Chromosome-level genome assembly of a parent species of widely cultivated azaleas. *Nat. Commun.* **2020**, 11, 5269. [CrossRef] [PubMed]

242. Hovde, B.T.; Daligault, H.E.; Hanschen, E.R.; Kunde, Y.A.; Johnson, M.B.; Starkenburg, S.R.; Johnson, S.L. Detection of Abrin-Like and Prepropulchellin-Like Toxin Genes and Transcripts Using Whole Genome Sequencing and Full-Length Transcript Sequencing of Abrus precatorius. *Toxins* **2019**, 11, 691. [CrossRef] [PubMed]

243. Chen, X.; Lu, Q.; Liu, H.; Zhang, J.; Hong, Y.; Lan, H.; Li, H.; Wang, J.; Liu, H.; Li, S.; et al. Sequencing of Cultivated Peanut, Arachis hypogaea, Yields Insights into Genome Evolution and Oil Improvement. *Mol. Plant* **2019**, 12, 920–934. [CrossRef] [PubMed]

244. Hong, Z.; Li, J.; Liu, X.; Lian, J.; Zhang, N.; Yang, Z.; Niu, Y.; Cui, Z.; Xu, D. The chromosome-level draft genome of Dalbergia odorifera. *Gigascience* **2020**, 9, giaa084. [CrossRef] [PubMed]

245. Kreplak, J.; Madoui, M.-A.; Cápal, P.; Novák, P.; Labadie, K.; Aubert, G.; Bayer, P.E.; Gali, K.K.; Syme, R.A.; Main, D.; et al. A reference genome for pea provides insight into legume genome evolution. *Nat. Genet.* **2019**, 51, 1411–1422. [CrossRef]

246. Qin, S.; Wu, L.; Wei, K.; Liang, Y.; Song, Z.; Zhou, X.; Wang, S.; Li, M.; Wu, Q.; Zhang, K.; et al. A draft genome for Spatholobus suberectus. *Sci. Data* **2019**, *6*, 113. [CrossRef]

247. Xing, Y.; Liu, Y.; Zhang, Q.; Nie, X.; Sun, Y.; Zhang, Z.; Li, H.; Fang, K.; Wang, G.; Huang, H.; et al. Hybrid de novo genome assembly of Chinese chestnut (*Castanea mollissima*). *Gigascience* **2019**, *8*, giz112. [CrossRef]

248. Liu, Y.; Tang, Q.; Cheng, P.; Zhu, M.; Zhang, H.; Liu, J.; Zuo, M.; Huang, C.; Wu, C.; Sun, Z.; et al. Whole-genome sequencing and analysis of the Chinese herbal plant Gelsemium elegans. *Acta Pharm. Sin. B* **2020**, *10*, 374–382. [CrossRef]

249. Hamilton, J.P.; Godden, G.T.; Lanier, E.; Bhat, W.W.; Kinser, T.J.; Vaillancourt, B.; Wang, H.; Wood, J.C.; Jiang, J.; Soltis, P.S.; et al. Generation of a chromosome-scale genome assembly of the insect-repellent terpenoid-producing Lamiaceae species, *Callicarpa americana*. *Gigascience* **2020**, *9*, giaa093. [CrossRef]

250. Dong, A.-X.; Xin, H.-B.; Li, Z.-J.; Liu, H.; Sun, Y.-Q.; Nie, S.; Zhao, Z.-N.; Cui, R.-F.; Zhang, R.-G.; Yun, Q.-Z.; et al. High-quality assembly of the reference genome for scarlet sage, Salvia splendens, an economically important ornamental plant. *Gigascience* **2018**, *7*, giy068. [CrossRef]

251. Chaw, S.-M.; Liu, Y.-C.; Wu, Y.-W.; Wang, H.-Y.; Lin, C.-Y.I.; Wu, C.-S.; Ke, H.-M.; Chang, L.-Y.; Hsu, C.-Y.; Yang, H.-T.; et al. Stout camphor tree genome fills gaps in understanding of flowering plant genome evolution. *Nat. Plants* **2019**, *5*, 63–73. [CrossRef]

252. Chen, Y.-C.; Li, Z.; Zhao, Y.-X.; Gao, M.; Wang, J.-Y.; Liu, K.-W.; Wang, X.; Wu, L.-W.; Jiao, Y.-L.; Xu, Z.-L.; et al. The Litsea genome and the evolution of the laurel family. *Nat. Commun.* **2020**, *11*, 1675. [CrossRef]

253. Chen, J.; Hao, Z.; Guang, X.; Zhao, C.; Wang, P.; Xue, L.; Zhu, Q.; Yang, L.; Sheng, Y.; Zhou, Y.; et al. Liriodendron genome sheds light on angiosperm phylogeny and species–pair differentiation. *Nat. Plants* **2019**, *5*, 18–25. [CrossRef]

254. Gao, Y.; Wang, H.; Liu, C.; Chu, H.; Dai, D.; Song, S.; Yu, L.; Han, L.; Fu, Y.; Tian, B.; et al. De novo genome assembly of the red silk cotton tree (*Bombax ceiba*). *Gigascience* **2018**, *7*, giy051. [CrossRef]

255. Ji, Y.; Xiu, Z.; Chen, C.; Wang, Y.; Yang, J.; Sui, J.; Jiang, S.; Wang, P.; Yue, S.; Zhang, Q.; et al. Long read sequencing of Toona sinensis (A. Juss) Roem: A chromosome-level reference genome for the family Meliaceae. *Mol. Ecol. Resour.* **2021**, *21*, 1243–1255. [CrossRef]

256. Peng, X.; Liu, H.; Chen, P.; Tang, F.; Hu, Y.; Wang, F.; Pi, Z.; Zhao, M.; Chen, N.; Chen, H.; et al. A Chromosome-Scale Genome Assembly of Paper Mulberry (*Broussonetia papyrifera*) Provides New Insights into Its Forage and Papermaking Usage. *Mol. Plant* **2019**, *12*, 661–677. [CrossRef]

257. Usai, G.; Mascagni, F.; Giordani, T.; Vangelisti, A.; Bosi, E.; Zuccolo, A.; Ceccarelli, M.; King, R.; Hassani-Pak, K.; Zambrano, L.S.; et al. Epigenetic patterns within the haplotype phased fig (*Ficus carica* L.) genome. *Plant J.* **2020**, *102*, 600–614. [CrossRef]

258. Zhang, X.; Wang, G.; Zhang, S.; Chen, S.; Wang, Y.; Wen, P.; Ma, X.; Shi, Y.; Qi, R.; Yang, Y.; et al. Genomes of the Banyan Tree and Pollinator Wasp Provide Insights into Fig-Wasp Coevolution. *Cell* **2020**, *183*, 875–889.e17. [CrossRef]

259. Jiao, F.; Luo, R.; Dai, X.; Liu, H.; Yu, G.; Han, S.; Lu, X.; Su, C.; Chen, Q.; Song, Q.; et al. Chromosome-Level Reference Genome and Population Genomic Analysis Provide Insights into the Evolution and Improvement of Domesticated Mulberry (*Morus alba*). *Mol. Plant* **2020**, *13*, 1001–1012. [CrossRef]

260. Yang, Y.; Sun, P.; Lv, L.; Wang, D.; Ru, D.; Li, Y.; Ma, T.; Zhang, L.; Shen, X.; Meng, F.; et al. Prickly waterlily and rigid hornwort genomes shed light on early angiosperm evolution. *Nat. Plants* **2020**, *6*, 215–222. [CrossRef]

261. Li, L.-F.; Cushman, S.A.; He, Y.-X.; Li, Y. Genome sequencing and population genomics modeling provide insights into the local adaptation of weeping forsythia. *Hortic. Res.* **2020**, *7*, 130. [CrossRef] [PubMed]

262. Han, B.; Jing, Y.; Dai, J.; Zheng, T.; Gu, F.; Zhao, Q.; Zhu, F.; Song, X.; Deng, H.; Wei, P.; et al. A Chromosome-Level Genome Assembly of Dendrobium Huoshanense Using Long Reads and Hi-C Data. *Genome Biol. Evol.* **2020**, *12*, 2486–2490. [CrossRef] [PubMed]

263. Wu, S.; Sun, W.; Xu, Z.; Zhai, J.; Li, X.; Li, C.; Zhang, D.; Wu, X.; Shen, L.; Chen, J.; et al. The genome sequence of star fruit (Averrhoa carambola). *Hortic. Res.* **2020**, *7*, 95. [CrossRef] [PubMed]

264. Lv, S.; Cheng, S.; Wang, Z.; Li, S.; Jin, X.; Lan, L.; Yang, B.; Yu, K.; Ni, X.; Li, N.; et al. Draft genome of the famous ornamental plant Paeonia suffruticosa. *Ecol. Evol.* **2020**, *10*, 4518–4530. [CrossRef] [PubMed]

265. Xia, Z.; Huang, D.; Zhang, S.; Wang, W.; Ma, F.; Wu, B.; Xu, Y.; Xu, B.; Chen, D.; Zou, M.; et al. Chromosome-scale genome assembly provides insights into the evolution and flavor synthesis of passion fruit (Passiflora edulis Sims). *Hortic. Res.* **2021**, *8*, 14. [CrossRef]

266. Li, M.; Zhang, D.; Gao, Q.; Luo, Y.; Zhang, H.; Ma, B.; Chen, C.; Whibley, A.; Zhang, Y.; Cao, Y.; et al. Genome structure and evolution of *Antirrhinum majus* L. *Nat. Plants* **2019**, *5*, 174–183. [CrossRef]

267. Liu, H.; Shi, J.; Cai, Z.; Huang, Y.; Lv, M.; Du, H.; Gao, Q.; Zuo, Y.; Dong, Z.; Huang, W.; et al. Evolution and Domestication Footprints Uncovered from the Genomes of Coix. *Mol. Plant* **2020**, *13*, 295–308. [CrossRef]

268. Thielen, P.M.; Pendleton, A.L.; Player, R.A.; Bowden, K.V.; Lawton, T.J.; Wisecaver, J.H. Reference Genome for the Highly Transformable Setaria viridis ME034V. *G3 Genes Genomes Genet.* **2020**, *10*, 3467–3478. [CrossRef]

269. Ma, D.; Guo, Z.; Ding, Q.; Zhao, Z.; Shen, Z.; Wei, M.; Gao, C.; Zhang, L.; Li, H.; Zhang, S.; et al. Chromosome-level assembly of the mangrove plant Aegiceras corniculatum genome generated through Illumina, PacBio and Hi-C sequencing technologies. *Mol. Ecol. Resour.* **2021**, *21*, 1593–1607. [CrossRef]

270. Jiang, S.; An, H.; Xu, F.; Zhang, X. Chromosome-level genome assembly and annotation of the loquat (*Eriobotrya japonica*) genome. *Gigascience* **2020**, *9*, giaa015. [CrossRef]

271. Buti, M.; Moretto, M.; Barghini, E.; Mascagni, F.; Natali, L.; Brilli, M.; Lomsadze, A.; Sonego, P.; Giongo, L.; Alonge, M.; et al. The genome sequence and transcriptome of *Potentilla micrantha* and their comparison to *Fragaria vesca* (the woodland strawberry). *Gigascience* **2018**, *7*, giy010. [CrossRef]

272. Liu, C.; Feng, C.; Peng, W.; Hao, J.; Wang, J.; Pan, J.; He, Y. Chromosome-level draft genome of a diploid plum (*Prunus salicina*). *Gigascience* **2020**, *9*, giaa130. [CrossRef]

273. Raymond, O.; Gouzy, J.; Just, J.; Badouin, H.; Verdenaud, M.; Lemainque, A.; Vergne, P.; Moja, S.; Choisne, N.; Pont, C.; et al. The Rosa genome provides new insights into the domestication of modern roses. *Nat. Genet.* **2018**, *50*, 772–777. [CrossRef]

274. Peng, Z.; Bredeson, J.V.; Wu, G.A.; Shu, S.; Rawat, N.; Du, D.; Parajuli, S.; Yu, Q.; You, Q.; Rokhsar, D.S.; et al. A chromosome-scale reference genome of trifoliate orange (*Poncirus trifoliata*) provides insights into disease resistance, cold tolerance and genome evolution in Citrus. *Plant J.* **2020**, *104*, 1215–1232. [CrossRef]

275. Ma, J.; Wan, D.; Duan, B.; Bai, X.; Bai, Q.; Chen, N.; Ma, T. Genome sequence and genetic transformation of a widely distributed and cultivated poplar. *Plant Biotechnol. J.* **2019**, *17*, 451–460. [CrossRef]

276. Liang, Q.; Li, H.; Li, S.; Yuan, F.; Sun, J.; Duan, Q.; Li, Q.; Zhang, R.; Sang, Y.L.; Wang, N.; et al. The genome assembly and annotation of yellowhorn (*Xanthoceras sorbifolium* Bunge). *Gigascience* **2019**, *8*, giz071. [CrossRef]

277. Khayi, S.; Azza, N.E.; Gaboun, F.; Pirro, S.; Badad, O.; Claros, M.G.; Lightfoot, D.A.; Unver, T.; Chaouni, B.; Merrouch, R.; et al. First draft genome assembly of the Argane tree (*Argania spinosa*). *F1000Research* **2020**, *7*, 1310. [CrossRef]

278. Xu, Z.; Xin, T.; Bartels, D.; Li, Y.; Gu, W.; Yao, H.; Liu, S.; Yu, H.; Pu, X.; Zhou, J.; et al. Genome Analysis of the Ancient Tracheophyte Selaginella tamariscina Reveals Evolutionary Features Relevant to the Acquisition of Desiccation Tolerance. *Mol. Plant* **2018**, *11*, 983–994. [CrossRef]

279. De-la-Cruz, I.M.; Hallab, A.; Olivares-Pinto, U.; Tapia-López, R.; Velázquez-Márquez, S.; Piñero, D.; Oyama, K.; Usadel, B.; Núñez-Farfán, J. Genomic signatures of the evolution of defence against its natural enemies in the poisonous and medicinal plant *Datura stramonium* (Solanaceae). *Sci. Rep.* **2021**, *11*, 882. [CrossRef]

280. van Lieshout, N.; van der Burgt, A.; de Vries, M.E.; ter Maat, M.; Eickholt, D.; Esselink, D.; van Kaauwen, M.P.W.; Kodde, L.P.; Visser, R.G.F.; Lindhout, P.; et al. Solyntus, the New Highly Contiguous Reference Genome for Potato (*Solanum tuberosum*). *G3 Genes Genomes Genet.* **2020**, *10*, 3489–3495. [CrossRef]

281. Wang, Y.; Xin, H.; Fan, P.; Zhang, J.; Liu, Y.; Dong, Y.; Wang, Z.; Yang, Y.; Zhang, Q.; Ming, R.; et al. The genome of Shanputao (*Vitis amurensis*) provides a new insight into cold tolerance of grapevine. *Plant J.* **2021**, *105*, 1495–1506. [CrossRef] [PubMed]