SCIENCE CHINA
Life Sciences

•REVIEW•

July 2019 Vol.62 No.7: 913–920
https://doi.org/10.1007/s11427-018-9472-y

Herbgenomics: A stepping stone for research into herbal medicine

Haoyu Hu, Xiaofeng Shen, Baosheng Liao, Lu Luo, Jiang Xu* & Shilin Chen*

Institute of Chinese Materia Medica, China Academy of Chinese Medicinal Sciences, Beijing 100700, China

Received November 19, 2018; accepted December 21, 2018; published online February 27, 2019

From the prehistoric era until the publishing of the Compendium of Materia Medica and the first scientific Nobel Prize in the Chinese mainland for Tu's discovery on anti-malarial tablets, each milestone and stepping stone in the developmental history of herbal medicine involved intrepid exploration, bold hypothesis formulation, and cautious verification. After thousands of years of discovery and development, herbal research has entered a new era—the era of herbgenomics. Herbgenomics combines herbal and genomic research, bridging the gap between traditional herbal medicine and cutting-edge omics studies. Therefore, it provides a general picture of the genetic background of traditional herbs, enabling researchers to investigate the mechanisms underlying the prevention and treatment of human diseases from an omics perspective.

herbgenomics, herbal medicines, omics, potential application

Citation: Hu, H., Shen, X., Liao, B., Luo, L., Xu, J., and Chen, S. (2019). Herbgenomics: A stepping stone for research into herbal medicine. Sci China Life Sci 62, 913–920. https://doi.org/10.1007/s11427-018-9472-y

Introduction

Herbal medicines, which refer to plant- or fungi-based pharmaceuticals, have been widely used for more than 5,000 years. They continue to play an important and irreplaceable role in modern society by reducing the effects of disease on human health (Tu, 2011). Primitive humans probably started using herbs to fight disease by noticing or imitating the behavior of animals that tended to consume specific plants when suffering from an illness. Herbal research has evolved from naked eye to microscopic and chemical analyses. In the ancient era, it was referred to as "tasting hundreds of plants." However, it has subsequently transformed in the morphological and biological period. It includes thorough investigation into the biological origin, taxonomy, and macroscopical and microscopical characteristics of a plant, as well as the study of its chemical nature via pharmaceutical chemistry. With recent advances in cutting-edge biological science, particularly omics technologies, herbal research has entered a brand new era—the era of "herbgenomics" (Figure 1). Omics studies into the genetic background of traditional herbs are of great significance to determine their biological basis, medicinal quality, herbal synthetic biology, natural herb resource protection, and drug safety. Thus, the foregoing technological developments will provide a major impetus to the continuous development and progress of human health (Chen et al., 2015).

Herbgenomics provides a theoretical basis and an effective technological platform for research into model herbs, biological basis of herbalism, targeted herb breeding, molecular identity, herbal synthetic biology, and construction of herbal gene banks. This enables researchers to systematically investigate medicinal plants by sequencing, assembling, and recording their genome, and further analyzing their genetic functions. Moreover, genomic information together with transcriptomic, proteomic, and metabolomic data can now be applied to numerous areas of herb-related biological research, triggering a revolution in large-data-based research into the genetic and biological nature of herbs (Figure 2)

---

*Corresponding authors (Jiang Xu, email: jxu@icmm.ac.cn; Shilin Chen, email: slchen@icmm.ac.cn)

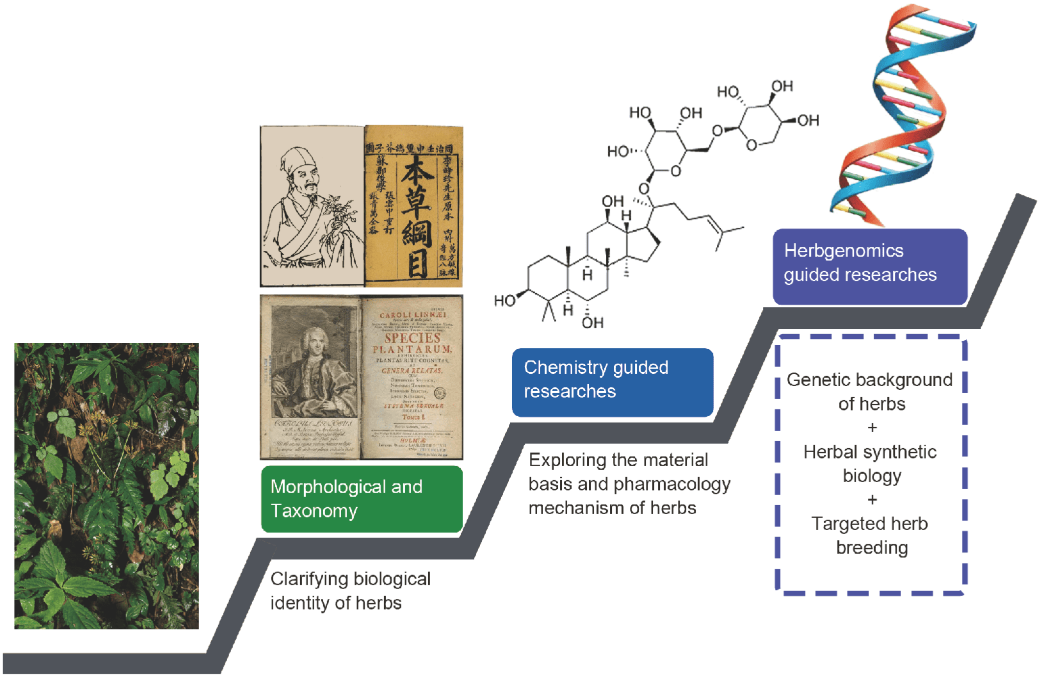

Figure 1 The three main evolvement phases of the theory concept and experimental methods of herbal research.

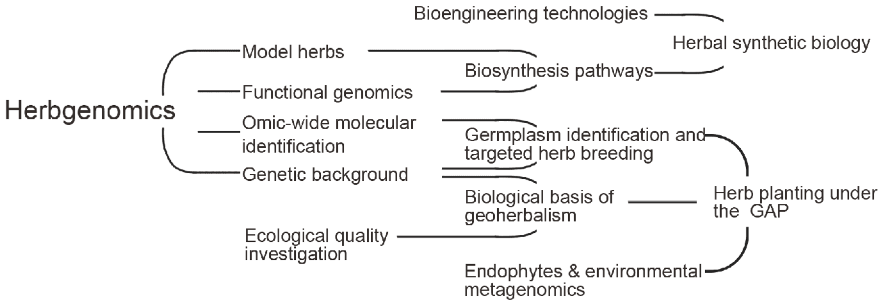

Figure 2 Research contents and applications of herbgenomics.

(Chen et al., 2015; Xin et al., 2018).

A milestone: current achievements of herbgenomics

The structural genome is the complete set of sequences in the genetic material of an organism. The typical eukaryotic genome comprises three main components: protein-coding, non-coding, and repetitive regions. Genomics research focuses on gene numbers, the gene's linear distribution and distance in the chromosome, and the coding and structure of intergenic spacer regions. The abundant diversity of plant species is reflected in their genome, which varies mainly with respect to size and complexity. This multiformity provides information regarding the genetic origin, evolution, speciation, individual development, agronomic traits, active compound biosynthesis, and all other biological characteristics of medicinal organisms. Chen et al. (2010) previously proposed a project, which aimed to sequence and investigate the whole genome of medicinal plants with a high economic value and typical secondary metabolic pathways. The key technology of this project encompasses species screening, genome pre-analysis, sequencing platform selection, genetic mapping, physical mapping, genome assembly, sequence maps, gene maps, genome recording, and bioinformatics analysis.

Many medicinal organisms are phylogenetically closely-related, sharing similar pharmacological activity. The phylogenomic approach enables the determination of phylogenetic relationships among herbs through structural comparisons of the genome. This enables researchers to identify the differences and similarities among different species, to further understand the species' evolutionary history, and to deduce the biosynthetic pathways of pharmaceutical ingredients with similar activity (Zhang et al., 2017).

The genome of a typical higher plant comprises three components: nuclear, chloroplast, and mitochondrial genomes. The nuclear genome contains most of the genetic information pertaining to the species origin, evolution, physiological regulation, growth and development, and synthetic secondary metabolite pathways. Since the first genome reporting of the herbal drug *Ricinus communis* in 2010 (Chan et al., 2010), over 40 medicinal plants and fungi have been sequenced and analyzed, including *Ganoderma lucidum* (Chen et al., 2012), *Salvia miltiorrhzia* (Xu et al., 2016), *Panax ginseng* (Xu et al., 2017; Jayakodi et al., 2018), *Panax notoginseng* (Chen et al., 2017), *Selaginella tamariscina* (Xu et al., 2018), *Artemisia annua* (Shen et al., 2018), and *Papaver somniferum* (Guo et al., 2018) (Table 1, Figure 3A). Acquiring information regarding medicinal organisms' whole genome has considerably promoted basic research, providing a foundation to investigate the biosynthetic pathways of different pharmacological active ingredients and the plant genome evolution. *Ganoderma lucidum*, frequently designated as seocho or "the mushroom of immortality", is a medicinal macrofungus, widely utilized in traditional Chinese medicine (TCM), which produces a diverse set of bioactive compounds. Chen et al. (2012) reported that the genome of *G. lucidum* comprises 43.3 Mb, encoding 16,113 predicted genes. The study also reported an impressive array of chromosome-level genomic maps, as well as the existence and regulation of a triterpenoid biosynthesis pathway in medicinal fungi (Figure 3B).

Chloroplasts (cps) are photosynthetic organelles, which transform light into chemical energy to synthesize starch, amino acids, pigments, and fatty acids (Neuhaus and Emes, 2000; Rodríguez-Ezpeleta et al., 2005). The chloroplast has its independent genome, which occurs as circular DNA and contains two copies of inverted repeats separating the large single copy and the small single copy regions (Palmer, 1991). The common angiosperm cp genomes are highly conserved in terms of gene contents and regulation (Wicke et al., 2011). Additionally, cp genomes have been found to present multiple useful characteristics, which decrease the efforts required for genetic transformation, recombination and replication, maternal inheritance, and high expression. Consequently, the cp genome constitutes an ideal data set for evolutionary research and molecular identification, possessing great potential for genetic engineering and herbal breeding. In recent years, over 7,200 plant cp genomic sequences have been published, among which over 20% are believed to have medicinal properties.

The mitochondrion is the other organelle with its own set of genetic instructions. As the center of aerobic respiration, mitochondria are the primary source of cellular energy and play a vital role in eukaryotic growth, development, and environmental adaption (Taanman, 1999; Gualberto et al., 2014). Compared with the nuclear genome, the mitochondrial genome has the following characteristics: (i) strict maternal lineage, no recombination, and less genetic resetting phenomena; (ii) fast evolution rate; and (iii) no tissue specificity or cellular abundance. In contrast to the highly-conserved vascular plants' cp genomes, the size of the mitochondrial genome of higher plants ranges between hundreds of kb to dozens of Mb, indicating toward a higher complex organization and structural diversity (Gualberto et al., 2014; Wu et al., 2016). The known molecular configurations of plant mitochondrial genomes include Y-type, H-type, multivariate linear, single linear, circular, and coexisting circular and linear. The complete mitochondrial genomes of some medicinal plants and fungi, such as *Brassica napusbn* (Handa, 2003), *Ganoderma lucidum* (Li et al., 2013), and *Ophiocordyceps sinensis* (Li et al., 2015), have been published, providing valuable information to understand mitochondrial origin and genome evolution of modern plants and fungi.

Applications of herbgenomics

Eukaryote genomics plays a central role in the biosynthesis of secondary metabolites and heredity procedures. There-

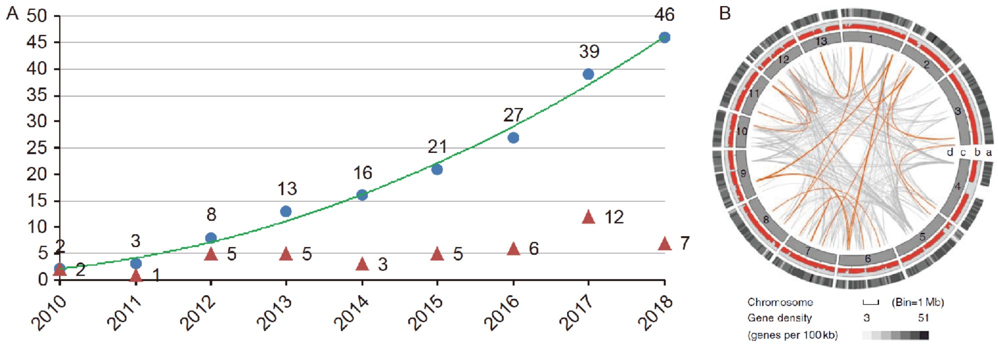

Figure 3 A trend of the number of publications on herb genome between 2010 and 2018. ● represents the total number of published herb genomes before the end of current year; ▲ represents the number of published herb genome studies during the current year. B, Chromosome-level genomic map showing the genomic features of *Ganoderma lucidum* (Chen et al. 2012).

Table 1 Published herb whole-genomes in 2017–2018[a)]

Family	Species	Assembled genome size	Sequencing platform	Reference
Araliaceae	*Panax ginseng*	3.43 Gb 3.00 Gb	Illumina HiSeq X-Ten Illumina HiSeq	(Xu et al., 2017) (Jayakodi et al., 2018)
Araliaceae	*Panax notoginseng*	1.85 Gb 2.39 Gb	Illumina HiSeq 2000 Illumina HiSeq 2000	(Chen et al., 2017) (Zhang D. et al., 2017)
Compositae	*Erigeron breviscapus*	1.20 Gb	Illumina HiSeq 2500, PacBio RS II platform	(Yang et al., 2017)
Crassulaceae	*Rhodiola crenulata*	344.5 Mb	Illumina HiSeq 2000/4000	(Fu et al., 2017)
Cucurbitaceae	*Momordica charantia*	339 Mb	Illumina MiSeq, HiSeq 2500	(Urasaki et al., 2017)
Fabaceae	*Glycyrrhiza uralensis*	379 Mb	Illumina HiSeq 2000, PacBio RS II platform	(Mochida et al., 2017)
Labiatae	*Mentha longifolia*	353 Mb	Illumina HiSeq 2000, PacBio RS platform	(Vining et al., 2017)
Nyssaceae	*Camptotheca acuminata*	384.5 Mb	Illumina HiSeq 2000	(Zhao et al., 2017)
Papaveraceae	*Macleaya cordata*	378 Mb	Illumina HiSeq 2000	(Liu et al., 2017)
Punicaceae	*Punica granatum*	328.38 Mb	Illumina HiSeq 2000	(Qin et al., 2017)
Theaceae	*Camellia sinensis*	3.02 Gb	Illumina HiSeq 2000	(Xia et al., 2017)
Asteraceae	*Artemisia annua*	1.78 Gb	Illumina HiSeq 2500 and MiSeq, PacBio RSII platform	(Shen et al., 2018)
Asteraceae	*Chrysanthemum nankingense*	2.53 Gb	Oxford Nanopore	(Song et al., 2018)
Orchidaceae	*Gastrodia elata*	1.06 Gb	Illumina HiSeq 2500	(Yuan et al., 2018)
Papaveraceae	*Papaver somniferum*	2.72Gb	PacBio SMRT, Oxford Nanopore and Illumina Hiseq2500/4000	(Guo et al., 2018)
Rosaceae	*Rosa chinensis*	503 Mb	Illumina HiSeq 2000, PacBio RSII platform	(Raymond et al., 2018)
Selaginellaceae	*Selaginella tamariscina*	301 Mb	PacBio RSII platform	(Xu et al., 2018)

a) Twenty-one herb whole genomes have been publish in 2010–2016 (Chan et al., 2010; Schmutz et al., 2010; van Bakel et al., 2011; Chen et al., 2012; Garcia-Mas et al., 2012; Guo et al., 2012; Bennetzen et al., 2012; The International Barley Genome Sequencing et al., 2012; Wang et al., 2012; Ming et al., 2013; Wang et al., 2013; Verde et al., 2013; He et al., 2013; Kitashiba et al., 2014; Liu et al., 2014; Wang et al., 2014; Zhu et al., 2015; Ye et al., 2015; Yang et al., 2015; Kellner et al., 2015; Yan et al., 2015; Guan et al., 2016; Zhang J. et al., 2016; Hoshino et al., 2016; He et al., 2016; Xu H. et al., 2016; Zhang G. Q. et al., 2016).

fore, these genomic data constitute enabling key steps for research into biosynthetic pathways, herbal synthetic biology, *omics* molecular identification, and targeted herb breeding.

Model herb research platform

The creation and analysis of model herbs enables researchers to comprehensively understand secondary metabolic activities and synthesis pathways of pharmaceutical components, through the biological systems' complex phenomena and properties. Nevertheless, the selection of model herbs is relatively more complex than that of normal model organisms. A typical model herb must comply with the following criteria: (i) have a short generation cycle; (ii) be easy to culture and genetically transform, and be environmentally friendly; (iii) have a relatively small genome size and a stable phenotype; and (iv) have a typical pathway of pharmaceutical secondary metabolite synthesis. According to these principles, several species were identified as valuable model herbs to study metabolic activities and genetics, including *Ganoderma lucidum*, *Salvia miltiorrhiza*, and *Catharanthus roseus* (Chen et al., 2015).

The chromosomal-level genome of *Ganoderma lucidum* has laid a solid foundation for its application as a medicinal model organism. Moreover, *G. lucidum* involves multiple secondary metabolite biosynthetic pathways, making it an ideal model to study the biosynthesis and regulation of terpenoids.

Salvia miltiorrhiza contains hydrophilic phenolicacids and tanshinones as pharmaceutically active components and has a high medicinal and economic value. Its genome and transcriptome have been previously reported (Cui et al., 2015). With such characteristics, *S. miltiorrhiza* meets the criteria as an ideal model herb for the study of multiple secondary metabolite biosynthetic pathways and their regulation (Wang et al., 2009; Xu Z. et al., 2016). A mutant library has already been constructed.

Herbal synthetic biology

Although herbs are the primary source of therapeutic compounds, they pose a sourcing problem, and their low production efficiency, long growth period, and farmland occupation remain common limitations in the production of natural medicaments. With the advances in herbgenomics research, herbal synthetic biology encompassing the discovery of natural product synthesis-related elements and the optimization of engineering principles, offers alternative production methods.

Herbal synthetic biology aims to decompose complex secondary metabolism pathways into simple, controllable modules or synthetic units, and to modify host cells (Chae et al., 2014), which can be used to regulate and reconstruct artificial metabolic systems. Its ultimate aim is the commercial production of TCM effective constituents.

Omics-wide molecular identification

DNA barcoding has developed rapidly, and has been used to identify medicinal species, control drug quality, and supervise the market (Erickson et al., 2008), thus promoting the standardization of TCM identification. This technology is currently recorded in the British and Chinese Pharmacopoeias (State Pharmacopoeia Committee, 2015; The British Pharmacopoeia Commission, 2015).

In recent years, the advances in next generation sequencing (NGS) have led to the introduction of omics molecular phylogenetic analysis and identification (Li et al., 2014; Coissac et al., 2016; Hollingsworth et al., 2016; Bakker, 2017). Extending concepts of DNA barcoding include superbarcoding, which introduces the whole cp genome as species-level DNA marker (Erickson et al., 2008; Sucher and Carles, 2008; Parks et al., 2009; Nock et al., 2011; Yang et al., 2013; Li et al., 2014) and genome skimming which refers to low-coverage shotgun sequencing of genomic DNA (Coissac et al., 2016; Hollingsworth et al., 2016; Bakker, 2017).

These extended molecular identification technologies can potentially overcome the limitations of traditional barcodes in the following ways: (i) they circumvent the need for a PCR stage, thus introducing the possibility to recover materials containing degraded DNA (e.g., ancient herbarium specimens up to 146 years old (Bakker et al., 2016), processed drugs or sediments); (ii) they greatly enhance the phylogenetic signal; (iii) they make way for the construction and standardization of a universal DNA marker through different taxa, whereas currently various loci are being applied by different researchers; (iv) they increase species discrimination through additional sequence data hybridization, in cases where repeated introgressions, incomplete linage sorting, reticulate evolution, and recent origins might occur.

Genomics-assisted herb breeding

As a more precisely targeted breeding strategy, genomics-assisted herb breeding results from the combination of molecular breeding and NGS. Plant genomics plays a vital role in plant improvement, including the discovery of the genetic variation underlying the enhanced effectiveness and increased efficiency of breeding (Bevan et al., 2017). Herbgenomics research and computational biology have accelerated the exploration and identification of functional genes in model herbs, strengthening the development of functional markers that provide a direct link to the production of desired compounds (Chen et al., 2015).

Genomics-assisted herb breeding involves an extensive combination of both genotype- and phenotype-related information and the design of multiple characteristics, thus breaking through the bottleneck of selection efficiency to accelerate the breeding process.

Herbal gene bank construction

With the generation of herbal genetic information at a soaring speed, a common and universally-shared platform for integrated access to all omics data is of utmost importance. Several herb-genetics-related databases involving genomic, epigenetic, transcriptomic, proteomic and metabolic pathway information, as well as DNA barcoding systems have already been developed. However, such resources are retrieved in different formats and its use requires bioinformatics skills (Chen et al., 2015). Therefore, a comprehensive and user-friendly database to organize and retrieve integrated molecular and biological data for herbal medicines is required. With advanced bioinformatics approaches and database technologies, multi-omics and chemical information can be applied to resolve a wide range of herb-related scientific issues (Yandell and Ence, 2012; Berger et al., 2013).

Conclusions and prospects

The present paper summarizes the background of herbgenomics generation, covering interdisciplinary subjects such as its theoretical basis, research contents, key technologies, practical applications, and research progress.

Despite the fact that herbgenomics research is still in its initial developmental stage, it has already obtained a great promotion and achievement. In the wake of high-throughput method developments in bioinformatics approaches and other innovative technologies, herbgenomics research will open a novel field for basic theory and applied research into natural drugs. Moreover, it will provide an unprecedented opportunity to revolutionize the use and acceptance of traditional herbal medicines.

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *This work was supported by the National Key Research and Development Program of China (2017-YFC1702100) and the Fundamental Research Funds for the Central Public Welfare Research Institutes (L2017004).*

References

Bakker, F.T., Lei, D., Yu, J., Mohammadin, S., Wei, Z., van de Kerke, S., Gravendeel, B., Nieuwenhuis, M., Staats, M., Alquezar-Planas, D.E., et al. (2016). Herbarium genomics: plastome sequence assembly from a range of herbarium specimens using an Iterative Organelle Genome Assembly pipeline. Biol J Linn Soc 117, 33–43.

Bennetzen, J.L., Schmutz, J., Wang, H., Percifield, R., Hawkins, J., Pontaroli, A.C., Estep, M., Feng, L., Vaughn, J.N., Grimwood, J., et al. (2012). Reference genome sequence of the model plant *Setaria*. Nat Biotechnol 30, 555–561.

Berger, B., Peng, J., and Singh, M. (2013). Computational solutions for omics data. Nat Rev Genet 14, 333–346.

Bevan, M.W., Uauy, C., Wulff, B.B.H., Zhou, J., Krasileva, K., and Clark, M.D. (2017). Genomic innovation for crop improvement. Nature 543, 346–354.

Chae, L., Kim, T., Nilo-Poyanco, R., and Rhee, S.Y. (2014). Genomic signatures of specialized metabolism in plants. Science 344, 510–513.

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. (2010). Draft genome sequence of the oilseed species *Ricinus communis*. Nat Biotechnol 28, 951–956.

Chen, S.L., Sun, Y.Z., Xu, J., Luo, H.M., Sun, C., He, L., Cheng, X.L., Zhang, B.L., and Xiao, P.G. (2010). Strategies of the study on Herb Genome Program (in Chinese). Acta Pharm Sin, 45, 807−812.

Chen, S., Xu, J., Liu, C., Zhu, Y., Nelson, D.R., Zhou, S., Li, C., Wang, L., Guo, X., Sun, Y., et al. (2012). Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. Nat Commun 3, 913.

Chen, S., Song, J.Y., Sun, C., Xu, J., Zhu, Y.J., and Verpoorte, R. (2015). Herbal genomics: examining the biology of traditional medicine. Science 347 (6219 suppl), S27–S29.

Chen, W., Kui, L., Zhang, G., Zhu, S., Zhang, J., Wang, X., Yang, M., Huang, H., Liu, Y., Wang, Y., et al. (2017). Whole-genome sequencing and analysis of the Chinese herbal plant *Panax notoginseng*. Mol Plant 10, 899–902.

Coissac, E., Hollingsworth, P.M., Lavergne, S., and Taberlet, P. (2016). From barcodes to genomes: extending the concept of DNA barcoding. Mol Ecol 25, 1423–1428.

Cui, G., Duan, L., Jin, B., Qian, J., Xue, Z., Shen, G., Snyder, J.H., Song, J., Chen, S., Huang, L., et al. (2015). Functional divergence of diterpene syntheses in the medicinal plant *Salvia miltiorrhiza*. Plant Physiol 169, pp.00695.2015.

Erickson, D.L., Spouge, J., Resch, A., Weigt, L.A., and Kress, J.W. (2008). DNA barcoding in land plants: developing standards to quantify and maximize success. TAXON 57, 1304–1316.

Bakker, F.T. (2017). Herbarium genomics: skimming and plastomics from archival specimens. Webbia 72, 35–45.

Fu, Y., Li, L., Hao, S., Guan, R., Fan, G., Shi, C., Wan, H., Chen, W., Zhang, H., Liu, G., et al. (2017). Draft genome sequence of the Tibetan medicinal herb *Rhodiola crenulata*. GigaScience 6, 1–5.

Garcia-Mas, J., Benjak, A., Sanseverino, W., Bourgeois, M., Mir, G., González, V.M., Hénaff, E., Câmara, F., Cozzuto, L., Lowy, E., et al. (2012). The genome of melon (*Cucumis melo* L.). Proc Natl Acad Sci USA 109, 11872–11877.

Gualberto, J.M., Mileshina, D., Wallet, C., Niazi, A.K., Weber-Lotfi, F., and Dietrich, A. (2014). The plant mitochondrial genome: dynamics and maintenance. Biochimie 100, 107–120.

Guan, R., Zhao, Y., Zhang, H., Fan, G., Liu, X., Zhou, W., Shi, C., Wang, J., Liu, W., Liang, X., et al. (2016). Draft genome of the living fossil *Ginkgo biloba*. GigaScience 5, 1–14.

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

Guo, S., Zhang, J., Sun, H., Salse, J., Lucas, W.J., Zhang, H., Zheng, Y., Mao, L., Ren, Y., Wang, Z., et al. (2012). The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. Nat Genet 45, 51–58.

Handa, H. (2003). The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. Nucleic Acids Res 31, 5907–5916.

He, N., Zhang, C., Qi, X., Zhao, S., Tao, Y., Yang, G., Lee, T.H., Wang, X., Cai, Q., Li, D., et al. (2013). Draft genome sequence of the mulberry tree *Morus notabilis*. Nat Commun 4, 2445.

He, Y., Xiao, H., Deng, C., Xiong, L., Nie, H., and Peng, C. (2016). Survey of the genome of *Pogostemon cablin* provides insights into its evolutionary history and sesquiterpenoid biosynthesis. Sci Rep 6, 26405.

Hollingsworth, P.M., Li, D.Z., van der Bank, M., and Twyford, A.D. (2016). Telling plant species apart with DNA: from barcodes to genomes. Phil Trans R Soc B 371, 20150338.

Hoshino, A., Jayakumar, V., Nitasaka, E., Toyoda, A., Noguchi, H., Itoh, T., Shin-I, T., Minakuchi, Y., Koda, Y., Nagano, A.J., et al. (2016). Genome sequence and analysis of the Japanese morning glory *Ipomoea nil*. Nat Commun 7, 13295.

Jayakodi, M., Choi, B.S., Lee, S.C., Kim, N.H., Park, J.Y., Jang, W., Lakshmanan, M., Mohan, S.V.G., Lee, D.Y., and Yang, T.J. (2018). Ginseng Genome Database: an open-access platform for genomics of *Panax ginseng*. BMC Plant Biol 18, 62.

Kellner, F., Kim, J., Clavijo, B.J., Hamilton, J.P., Childs, K.L., Vaillancourt, B., Cepela, J., Habermann, M., Steuernagel, B., Clissold, L., et al. (2015). Genome-guided investigation of plant natural product biosynthesis. Plant J 82, 680–692.

Kitashiba, H., Li, F., Hirakawa, H., Kawanabe, T., Zou, Z., Hasegawa, Y., Tonosaki, K., Shirasawa, S., Fukushima, A., Yokoi, S., et al. (2014). Draft sequences of the radish (*Raphanus sativus* L.) genome. DNA Res 21, 481–490.

Li, X., Yang, Y., Henry, R.J., Rossetto, M., Wang, Y., and Chen, S. (2014). Plant DNA barcoding: from gene to genome. Biol Rev 90, 157–166.

Li, Y., Hu, X.D., Yang, R.H., Hsiang, T., Wang, K., Liang, D.Q., Liang, F., Cao, D.M., Zhou, F., Wen, G., et al. (2015). Complete mitochondrial genome of the medicinal fungus *Ophiocordyceps sinensis*. Sci Rep 5, 13892.

Liu, M.J., Zhao, J., Cai, Q.L., Liu, G.C., Wang, J.R., Zhao, Z.H., Liu, P., Dai, L., Yan, G., Wang, W.J., et al. (2014). The complex jujube genome provides insights into fruit tree biology. Nat Commun 5, 5315.

Liu, X., Liu, Y., Huang, P., Ma, Y., Qing, Z., Tang, Q., Cao, H., Cheng, P., Zheng, Y., Yuan, Z., et al. (2017). The genome of medicinal plant *Macleaya cordata* provides new insights into benzylisoquinoline alkaloids metabolism. Mol Plant 10, 975–989.

Ming, R., VanBuren, R., Liu, Y., Yang, M., Han, Y., Li, L.T., Zhang, Q., Kim, M.J., Schatz, M.C., Campbell, M., et al. (2013). Genome of the long-living sacred lotus (*Nelumbo nucifera* Gaertn.). Genome Biol 14, R41.

Mochida, K., Sakurai, T., Seki, H., Yoshida, T., Takahagi, K., Sawai, S., Uchiyama, H., Muranaka, T., and Saito, K. (2017). Draft genome assembly and annotation of *Glycyrrhiza uralensis*, a medicinal legume. Plant J 89, 181–194.

Neuhaus, H.E., and Emes, M.J. (2000). Nonphotosynthetic metabolism in plastids. Annu Rev Plant Physiol Plant Mol Biol 51, 111–140.

Nock, C.J., Waters, D.L.E., Edwards, M.A., Bowen, S.G., Rice, N., Cordeiro, G.M., and Henry, R.J. (2011). Chloroplast genome sequences from total DNA for plant identification. Plant Biotech J 9, 328–333.

Palmer, J.D. (1991). CHAPTER 2—Plastid Chromosomes: Structure and Evolution. In Bogorad, L., Vasil, I.K., eds. The Molecular Biology of Plastids (Academic Press), pp. 5–53.

Parks, M., Cronn, R., and Liston, A. (2009). Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. BMC Biol 7, 84.

Qin, G., Xu, C., Ming, R., Tang, H., Guyot, R., Kramer, E.M., Hu, Y., Yi, X., Qi, Y., Xu, X., et al. (2017). The pomegranate (*Punica granatum* L.) genome and the genomics of punicalagin biosynthesis. Plant J 91, 1108–1128.

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

Rodríguez-Ezpeleta, N., Brinkmann, H., Burey, S.C., Roure, B., Burger, G., Löffelhardt, W., Bohnert, H.J., Philippe, H., and Lang, B.F. (2005). Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. Curr Biol 15, 1325–1330.

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. (2010). Genome sequence of the palaeopolyploid soybean. Nature 463, 178–183.

Shen, Q., Zhang, L., Liao, Z., Wang, S., Yan, T., Shi, P., Liu, M., Fu, X., Pan, Q., Wang, Y., et al. (2018). The genome of *Artemisia annua* provides insight into the evolution of Asteraceae family and artemisinin biosynthesis. Mol Plant 11, 776–788.

Song, C., Liu, Y., Song, A., Dong, G., Zhao, H., Sun, W., Ramakrishnan, S., Wang, Y., Wang, S., Li, T., et al. (2018). The *Chrysanthemum nankingense* genome provides insights into the evolution and diversification of chrysanthemum flowers and medicinal traits. Mol Plant 11, 1482–1491.

State Pharmacopoeia Committee. (2015). Pharmacopoeia of the People's Republic of China, Part IV. pp. 383–385.

Sucher, N.J., and Carles, M.C. (2008). Genome-based approaches to the authentication of medicinal plants. Planta Med 74, 603–623.

Taanman, J.W. (1999). The mitochondrial genome: structure, transcription, translation and replication. Biochim Biophys Acta 1410, 103–123.

The British Pharmacopoeia Commission. (2015). British Pharmacopoeia Appendix XI V. Deoxyribonucleic Acid (DNA) Based Identification Techniques for Herbal Drugs.

The International Barley Genome Sequencing. (2012). A physical, genetic and functional sequence assembly of the barley genome. Nature 491, 711.

Tu, Y. (2011). The discovery of artemisinin (Qinghaosu) and gifts from Chinese medicine. Nat Med 17, 1217–1220.

Urasaki, N., Takagi, H., Natsume, S., Uemura, A., Taniai, N., Miyagi, N., Fukushima, M., Suzuki, S., Tarora, K., Tamaki, M., et al. (2017). Draft genome sequence of bitter gourd (*Momordica charantia*), a vegetable and medicinal plant in tropical and subtropical regions. DNA Res 24, 51–58.

van Bakel, H., Stout, J.M., Cote, A.G., Tallon, C.M., Sharpe, A.G., Hughes, T.R., and Page, J.E. (2011). The draft genome and transcriptome of *Cannabis sativa*. Genome Biol 12, R102.

Verde, I., Abbott, A.G., Scalabrin, S., Jung, S., Shu, S., Marroni, F., Zhebentyayeva, T., Dettori, M.T., Grimwood, J., Cattonaro, F., et al. (2013). The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. Nat Genet 45, 487–494.

Vining, K.J., Johnson, S.R., Ahkami, A., Lange, I., Parrish, A.N., Trapp, S. C., Croteau, R.B., Straub, S.C.K., Pandelova, I., and Lange, B.M. (2017). Draft genome sequence of *Mentha longifolia* and development of resources for mint cultivar improvement. Mol Plant 10, 323–339.

Wang, L., Yu, S., Tong, C., Zhao, Y., Liu, Y., Song, C., Zhang, Y., Zhang, X., Wang, Y., Hua, W., et al. (2014). Genome sequencing of the high oil crop sesame provides insight into oil biosynthesis. Genome Biol 15, R39.

Wang, Q., Chen, A., and Zhang, B. (2009). *Salvia miltiorrhzia*, a model organism for Chinese traditional medicine genomic studies. Acta Chin Med Pharmacol 37, 2.

Wang, Y., Fan, G., Liu, Y., Sun, F., Shi, C., Liu, X., Peng, J., Chen, W., Huang, X., Cheng, S., et al. (2013). The sacred lotus genome provides insights into the evolution of flowering plants. Plant J 76, 557–567.

Wang, Z., Hobson, N., Galindo, L., Zhu, S., Shi, D., McDill, J., Yang, L., Hawkins, S., Neutelings, G., Datla, R., et al. (2012). The genome of flax (*Linum usitatissimum*) assembled *de novo* from short shotgun sequence reads. Plant J 72, 461–473.

Wicke, S., Schneeweiss, G.M., dePamphilis, C.W., Müller, K.F., and Quandt, D. (2011). The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol Biol 76, 273–297.

Wu, L., Wang, B., Zhao, M., Liu, W., Zhang, P., Shi, Y., Xiong, C., Wang, P., Sun, W., and Chen, S. (2016). Rapid identification of officinal Akebiae Caulis and its toxic adulterant Aristolochiae Manshuriensis Caulis (*Aristolochia manshuriensis*) by loop-mediated isothermal amplification. Front Plant Sci 7, 887.

Xia, E.H., Zhang, H.B., Sheng, J., Li, K., Zhang, Q.J., Kim, C., Zhang, Y., Liu, Y., Zhu, T., Li, W., et al. (2017). The tea tree genome provides insights into tea flavor and independent evolution of caffeine biosynthesis. Mol Plant 10, 866–877.

Xin, T., Zhang, Y., Pu, X., Gao, R., Xu, Z., and Song, J. (2018). Trends in herbgenomics. Sci China Life Sci, doi: 10.1007/s11427-018-9352-7.

Xu, H., Song, J., Luo, H., Zhang, Y., Li, Q., Zhu, Y., Xu, J., Li, Y., Song, C., Wang, B., et al. (2016). Analysis of the genome sequence of the medicinal plant *Salvia miltiorrhiza*. Mol Plant 9, 949–952.

Xu, J., Chu, Y., Liao, B., Xiao, S., Yin, Q., Bai, R., Su, H., Dong, L., Li, X., Qian, J., et al. (2017). Panax ginseng genome examination for ginsenoside biosynthesis. GigaScience 6.

Xu, Z., Ji, A., Zhang, X., Song, J., and Chen, S. (2016). Biosynthesis and regulation of active compounds in medicinal model plant *Salvia miltiorrhiza*. Chin Herb Med 8, 3–11.

Xu, Z., Xin, T., Bartels, D., Li, Y., Gu, W., Yao, H., Liu, S., Yu, H., Pu, X., Zhou, J., et al. (2018). Genome analysis of the ancient tracheophyte *Selaginella tamariscina* reveals evolutionary features relevant to the acquisition of desiccation tolerance. Mol Plant 11, 983–994.

Yan, L., Wang, X., Liu, H., Tian, Y., Lian, J., Yang, R., Hao, S., Wang, X., Yang, S., Li, Q., et al. (2015). The genome of *Dendrobium officinale* illuminates the biology of the important traditional chinese orchid herb. Mol Plant 8, 922–934.

Yandell, M., and Ence, D. (2012). A beginner's guide to eukaryotic genome annotation. Nat Rev Genet 13, 329–342.

Yang, J.B., Tang, M., Li, H.T., Zhang, Z.R., and Li, D.Z. (2013). Complete chloroplast genome of the genus *Cymbidium*: lights into the species identification, phylogenetic implications and population genetic analyses. BMC Evol Biol 13, 84.

Yang, J., Zhang, G., Zhang, J., Liu, H., Chen, W., Wang, X., Li, Y., Dong, Y., and Yang, S. (2017). Hybrid *de novo* genome assembly of the Chinese herbal fleabane *Erigeron breviscapus*. GigaScience 6, 1–7.

Yang, K., Tian, Z., Chen, C., Luo, L., Zhao, B., Wang, Z., Yu, L., Li, Y., Sun, Y., Li, W., et al. (2015). Genome sequencing of adzuki bean (*Vigna angularis*) provides insight into high starch and low fat accumulation and domestication. Proc Natl Acad Sci USA 112, 13213–13218.

Ye, N., Zhang, X., Miao, M., Fan, X., Zheng, Y., Xu, D., Wang, J., Zhou, L., Wang, D., Gao, Y., et al. (2015). Saccharina genomes provide novel insight into kelp biology. Nat Commun 6, 6986.

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

Zhang, D., Li, W., Xia, E.H., Zhang, Q.J., Liu, Y., Zhang, Y., Tong, Y., Zhao, Y., Niu, Y.C., Xu, J.H., et al. (2017). The medicinal herb *Panax notoginseng* genome provides insights into ginsenoside biosynthesis and genome evolution. Mol Plant 10, 903–907.

Zhang, G.Q., Xu, Q., Bian, C., Tsai, W.C., Yeh, C.M., Liu, K.W., Yoshida, K., Zhang, L.S., Chang, S.B., Chen, F., et al. (2016). The *Dendrobium catenatum* Lindl. genome sequence provides insights into polysaccharide synthase, floral development and adaptive evolution. Sci Rep 6, 19029.

Zhang, J.J., Su, H., Zhang, L., Liao, B.S., Xiao, S.M., Dong, L.L., Hu, Z.G., Wang, P., Li, X.W., Huang, Z.H., et al. (2017). Comprehensive characterization for ginsenosides biosynthesis in Ginseng root by integration analysis of chemical and transcriptome. Molecules 22, 889.

Zhang, J., Tian, Y., Yan, L., Zhang, G., Wang, X., Zeng, Y., Zhang, J., Ma, X., Tan, Y., Long, N., et al. (2016). Genome of plant Maca (*Lepidium meyenii*) illuminates genomic basis for high-altitude adaptation in the central Andes. Mol Plant 9, 1066–1077.

Zhao, D., Hamilton, J.P., Pham, G.M., Crisovan, E., Wiegert-Rininger, K., Vaillancourt, B., DellaPenna, D., and Buell, C.R. (2017). *De novo* genome assembly of *Camptotheca acuminata*, a natural source of the anti-cancer compound camptothecin. GigaScience 6.

Zhu, Y., Xu, J., Sun, C., Zhou, S., Xu, H., Nelson, D.R., Qian, J., Song, J., Luo, H., Xiang, L., et al. (2015). Chromosome-level genome map provides insights into diverse defense mechanisms in the medicinal fungus *Ganoderma sinense*. Sci Rep 5, 11087.