

Review

Herbal genomics as tools for dissecting new metabolic pathways of unexplored medicinal plants and drug discovery

Prasanta Chakraborty*

Kalpana Chawla Center for Space and Nanosciences, Indian Institute of Chemical Biology (retd.), Kolkata, 700032, India

Received 12 July 2017; accepted 28 December 2017

Available online 10 January 2018

Abstract

Herbal drugs, on which 80% of the world's population rely, are relatively safe over conventional drugs. Conventional drugs are costly, have serious side effects and hence over the past few decades researchers have focused on drug discovery from herbal medicines or botanical sources. The majority of new herbal drugs have been generated from secondary metabolites (alkaloids, terpenoids and phenolic compounds) of plant metabolism. Till date, only a small fraction of the vast diversity of plant metabolism has been explored for the production of new medicines and other products. The emergence of new herbal genomics research, medicinal plant genomics consortium, together with advances in other omics information may help for the speedy discovery of previously unknown metabolic pathways and enzymes. This review highlights the importance of genomics research in the discovery of some previously unknown enzymes/pathways which may make significant contributions in plant metabolic biology and may be used for the future discovery of many new pharmaceutical agents.

© 2018 Published by Elsevier B.V. on behalf of Société Française de Biochimie et Biologie Moléculaire (SFBBM). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Herbal; Genomics; Metabolite; Drug

1. Introduction (why herbal genomics?)

Traditional herbal medicines, basically plant-based therapies have been used for more than 5000 years [1]. At the time of invention of medicinal properties of herbs, people could not think of scientific evidence, philosophical and experimental basis, molecules responsible for medicinal value and of course the currently emerging herbal genomics. Constant and renewed public interest on alternative and complementary medicine lies mainly due to high cost of new drugs, increased side effects, microbial resistance and lack of curative treatment for several chronic diseases. Countries like India, China, Korea and Japan are now taking lead role and continuously investing in research on evidence-based traditional medicines and scientific validation of fundamental principles. China, the traditional medicine giant has successfully promoted its own

therapies over the globe with a science-based approach and evidenced by the rapid increase in number of licensed Chinese medicine providers in the United states [2]. India is also gearing up, and there has been a steep rise in the global acceptance of Ayurveda, the traditional Indian medicine. The World Health Organization has listed 21,000 medicinal plants, among which 2500 species are in India and now India is known as the largest producer of medicinal herbs [3]. Collectively, global data shows that 80% of world's population, rely primarily on ethnobotanical remedies and plant drugs, e.g. antineoplastic: camptothecin, Taxol, antimalarial: artemisinin, quinine, antigout: colchicine, analgesic: codeine, morphine, cardiac depressant: quinidine, antidiabetic: allicin, and for brain functions: caffeine, nicotine are the well known curative agents (Table 1, Fig. 5, modified from [4,5]).

Although, plant derived natural products remain rich resources for drug development and have had a profound and lasting impact on human health with almost 100 plant-derived compounds in clinical trials as of late 2007, the clinical

* Akashdeep, 184, Santoshpur Avenue, Kolkata, 700075, India.

E-mail address: prasanta3274@yahoo.in.

Table 1
Recognized drugs from medicinal plants and their actions against various human diseases.

Drugs	Medicinal plants	Actions against
vinblastine	<i>Catharanthus roseus</i>	Cancer
vincristine	<i>Catharanthus roseus</i>	
camptothecin	<i>Camptotheca acuminata</i>	
taxol	<i>Taxus baccata</i>	
podophyllotoxin	<i>Podophyllum peltatum</i>	
artemisinin	<i>Artemisia annua</i>	Malaria
quinine	<i>Cinchona ledgeriana</i>	
quinoline	<i>Cinchona ledgeriana</i>	
digoxin	<i>Digitalis purpurea</i>	Cardiac disorder
digitoxin	<i>Digitalis purpurea</i>	
quinidine	<i>Cinchona ledgeriana</i>	
codeine	<i>Papaver somniferum</i>	Analgesic
morphine	<i>Papaver somniferum</i>	
allicin	<i>Allium sativum</i>	Diabetes
SMCS (S-methyl cysteine sulfoxide)	<i>Allium cepa</i>	
nicotine	<i>Nicotiana tabacum</i>	Brain disorder
caffeine	<i>Coffea canephora</i>	
diosgenin	<i>Dioscorea Mexicana</i>	sexual problems
stigmasterol	<i>Glycine max</i>	
cannabinoids	<i>Cannabis sativa</i>	psycho-disorder
tubocurarine	<i>Chondodendron tomentosum</i>	muscle disorder
atropine	<i>Atropa belladonna</i>	nervous system
hyoscyamine	<i>Hysocyamus niger</i>	

potential of these compounds is often curtailed due to low production levels in plant species or due to loss of source for extinction [6]. For example, use of the blockbuster drug Taxol almost stopped in the early 1990's because the primary source, yew tree bark, could not be used as a sustainable source of the drug. A taxol precursor and a semi-synthetic protocol could have been the ideal way to convert it into the active drug. More generalized solutions, such as metabolic engineering of effective plant and microbial production platforms, are urgently needed to ensure that the plant-derived compounds having enormous structural diversity and biological activities enter the clinical pipeline and find widespread use in medicine.

Considering the vast chemical biodiversity of the plant world, and widespread medicinal values of plant derived natural products, researchers are now focusing on the poorly understood areas of herbs such as genetic background, the agricultural traits, and the medicinal quality. With rapid advances in high throughput sequencing technologies and greatly reduced costs, a new discipline called “herbal genomics” is now emerging. Systematic analysis of medicinal herbs genes' functions are necessary and achievable through sequencing, assembling and annotating their genomes. There have been only a few well assembled herbal genomes released to date, partly because of their complexity. Genomic information, together with transcriptomic, proteomic and metabolomic data, can therefore be used to predict the secondary metabolic pathways of herbs. Moreover, functional herbal genomics can contribute to model herb research platforms, geoherb

research, and herbal synthetic biology, all of which are important for securing the sourcing of the medicinal plants and their active compounds in the future.

2. Genomes and metabolic activities/pathways of herbs

Plants synthesize an abundance of metabolites that can be exploited for pharmacological purposes. Till today, only a small fraction of the immense diversity of plant metabolism has been explored for the production of new medicines and other products important to human well-being. All plants synthesize basic metabolites through primary metabolism needed for their survival but different taxa produce distinct metabolites through secondary or specialized metabolism that are specialized for some specific reactions. As stationary autotrophs, plants have to cope with a number of challenges such as local fluctuations of the simple nutrients they require to synthesize their foods, coexistence of herbivores and pathogens in their immediate environment. Plants have therefore evolved secondary biochemical pathways that allow them to synthesize a diverse array of organic compounds to counter-attack to specific environmental stimuli [7,8]. The palette of secondary metabolites are subdivided into a number of distinct groups, on the basis of their chemical structure and synthetic pathways and those groups are the alkaloids, terpenoids and the phenolic compounds. With a view to understand the full metabolic potential/pathways of the plant/herbs, the whole nuclear and chloroplast genomes needed to be sequenced. In this regard, Chen et al. [44,45] initiated a project, “Herb Genome Programme” for the genome sequencing of various medicinal plants and post genomic functional analysis of various secondary metabolite biosynthetic pathways. Unfortunately, there have been only a few well assembled herbal genomes released to date, partly because of their complexity. The genomes of some commonly used herbs such as *Ganoderma lucidum*, *Salvia miltiorrhiza* and *Catharanthus roseus* have already been sequenced and they emerged as valuable models for studying the genetics and metabolic activities of herbs [1,9,10,50]. These species have been shown to synthesize active pharmaceutical components, including triterpenes, diterpene quinone and indole alkaloids. Analysis of a draft genome sequence of *C. roseus* provided evidence for partial clustering of genes for the biosynthesis of the monoterpene indole alkaloids vinblastine and vincristine. With the help of bacterial artificial chromosome (BAC) sequencing, Kellner et al. [50] showed seven small clusters each of two to three genes that contained genes encoding enzymes for vinblastine/vincristine biosynthesis pathway and other genes for other pathway. In addition to those, the whole genome of medicinal plant, *Z. jujuba* [46,51], *A. Indica*(Neem) [47], and chloroplast genomes of *P. cablin* [48] and *S. miltiorrhiza* [49] have been sequenced successfully. *Z. jujuba* has got significant medicinal value and its fruits are rich in vitamin C and sugar, and the plant contains various therapeutically important alkaloids, flavonoids and phenolics. De novo assembly of its complex genome and transcriptomics data established that L-galactose pathway is the major synthesis pathway for vitamin C and

consistently higher expression of the genes for the enzymes, GDP-D-mannose 3,5 epimerase, and GDP-L-galactose phosphorylase contributes for sugar metabolism. *A. Indica* (Neem) is important for its huge medicinal value and bioactivities against malaria, diabetes and tumor. The genomes and transcriptomes analysis of neem shows that its genome is AT-rich, bears little repetitive DNA elements and comprises about 20,000 genes. Comparative transcript expression analysis showed either exclusive or enhanced expression of known genes involved in neem terpenoid biosynthesis. Draft sequence of whole genome of *S. miltiorrhiza* by Xu et al. [52]. shows that plant genome size is ~600 MB and contains 30,478 protein coding genes, and 1620 genes for transcription factors, and several of these transcription factors revealed to be involved in the biosynthesis of tanshinone and phenolic acids used in several cardiovascular, cerebrovascular and hyperlipidemia disease. While, chloroplast genome of this plant is 151,328 bp in length and it contains 114 unique genes including 80-protein coding genes, 30-tRNA genes and four rRNA genes [49]. The chloroplast genome sequences will facilitate population, phylogenetic and genetic engineering studies of this medicinal plant. *Pogostemon cablin*, an important herb in the Lamiaceae family contains more than 40 major components, including flavonoids, terpenoids, alkaloids and phenylpropanoid glycosides. It has got wide range of applications including medicinal effects like anti-inflammatory, antidepressant activity. The entire chloroplast genome with 38.24% GC content, is 152,460 bp in length and contains 127 genes of which 107 genes are single-copy, including 79 protein-coding genes, four rRNA genes, and 24 tRNA genes. Phylogenetic analysis reveals that *P. cablin* diverged from the *Sentellarioideae* clade about 29.45 million year ago. Complete sequences and annotation of *P. cablin* genome will help in the understanding of population, phylogenetic and genetic engineering research [48]. *Ocimum sanctum* L. is an important sacred medicinal plant of India, whose whole nuclear and chloroplast genomes recently been sequenced [11]. The genome sequence and annotation of *O. sanctum* and another species *O. basilicum* indicates higher expression of phenylpropanoid/terpenoid pathway genes, expression of several cytochrome P450s and transcription factor families and the information thus provides new insights for mining biosynthetic pathways for important metabolites of medicinal nature in related species [11,12].

Metabolic diversification across plants different taxa (from algae to angiosperms) and their pathway studies may further provide insights for the potential discovery of novel specialized metabolic processes. Genes coding for specialized metabolic functions, their proliferations and physical clustering within the genome when compared with their relative primary metabolic counterparts, some important observations were made [13]. Primary metabolism related reactions, such as carbohydrate metabolism were enriched in early land plant, hormone-related reactions were enriched in algae, whereas angiosperms were enriched for specialized metabolism reactions (amino acid, carbohydrate and nucleotide metabolism reactions significantly underrepresented in angiosperm). In

specialized metabolism, gene duplication such as whole-genome duplications (WGDs) and local (tandem) duplication (LDs) play an important role for increase gene content in genomes. For example in *Arabidopsis*, specialized metabolic genes were significantly enriched in LD genes, but other suggested that whole genome duplications contributed significantly to these metabolic pathways [14]. In soybean and sorghum species, significant depletions of WGD-derived specialized metabolic genes were observed, but displayed significant enrichment in LD-derived specialized metabolic genes [15]. These local duplications of genes appeared to affect a number of flavonoid related genes in each species. Gene clustering also plays a major role in specialized metabolic events, discussed above little bit for the medicinal plant *C. roseus*. For soybean, sorghum and *Arabidopsis*, it was shown that approximately one-third of the metabolic genes and one-fifth of the genes in rice were situated in clusters. The *Arabidopsis* and soybean clusters were significantly enriched for specialized metabolic genes. Clustered genes in *Arabidopsis* were enriched in phenylpropanoid and terpenoid metabolism, and clustered genes in sorghum were enriched in terpenoid metabolism but the enrichment pattern differed in soybean. The differing results among the species suggest that specialized metabolic gene clusters are a product of independent mechanism rather than a broad mechanism underlying specialized metabolism evolution.

Furthermore, additional research for digging out medicinally important plant metabolic pathways involves various processes including i) gene cloning approaches, ii) RNA interference (RNAi), and iii) virus-induced gene silencing (VIGS) technology. i) Gene cloning approaches involves protein purification, protein sequencing, and use of protein sequence to clone desired genes. A number of genes involved in the biosynthesis of monoterpenoid indole alkaloids (MIAs) vindoline, camptothecin and catharanthine have been identified and most of the steps in taxol and morphine biosynthesis have been elucidated. Unfortunately, this approach is slow and tedious and most importantly, this requires random testing of many members of the gene family, together with substrate availability, to determine whether a candidate gene is involved in the target pathway [16]. ii) RNA interference is a biological mechanism that eliminates targeted mRNAs through homology dependent gene silencing and has been successfully used in the transformation of plants through the introduction of short double stranded RNA that interferes with the expression of the targeted gene. Transgenic poppies were created through this technology which accumulates the benzyloquinoline alkaloid reticuline rather than the narcotic morphine [17]. iii) Virus-induced gene silencing technology has been shown to be useful for an increasing number of plant hosts and is particularly useful for plants that are difficult to transform [18]. VIGS technology has identified genes involved in benzyloquinoline alkaloid (BIA) biosynthesis and has helped identify the two reactions involved in the consecutive O-demethylation of thebaine to codeine and codeine to morphine in opium poppy [19]. Furthermore, although tropane alkaloids (hyoscyamine, atropine and scopolamine) and their biosynthetic

pathways have been well characterized, testing candidate genes encoding cytochrome P450s (CYP P450) of *Hyoscyamus niger*, representing tropane alkaloid pathway-enriched transcripts with VIGS technology identified a CYP80F1 family member that when silenced showed preferential accumulation of litorine rather than hyoscyamine [20]. MIA pathway genes from *Catharanthus roseus*, though mostly have been characterized, the successful VIGS-mediated suppression of 16-methoxy-2,3-dihydro-3-hydroxytabersonine N-methyltransferase (NMT) in *C. roseus* and targeted metabolic profiling of NMT-VIGS-silenced plants showed preferential accumulation of NMT substrate 16-methoxy-2,3-dihydro-3-hydroxytabersonine rather than the downstream MIA vindoline [21,22]. VIGS technology may thus be used as a gene discovery tool, and should greatly speed up the discovery of the remaining MIA pathway genes in *C. roseus* and other medicinal herbs.

3. Identification of unknown pathways/enzymes for new drug discovery

More than one-third of human drugs have originated from products synthesized by plant metabolism. Understanding the origins and vast diversification of plant metabolism has been a long-standing goal in plant biology. Since, only a fraction of the vast chemical biodiversity of the plant world has been explored, its worth stressing on high-throughput sequencing technology to examine increasing number of species for the speedy discovery of previously unknown enzymes and pathways for the production of new medicines. Till date, very little is known about biosynthetic genes of clinically used existing drugs which prevents access to engineered hosts for their production. Very few complete pathways exist, and only three-

artemisinic acid and the benzyloisoquinoline alkaloids [23,24], and monoterpene indole alkaloids [25,26] have heterologous host for industrial production. Now, the availability of inexpensive large-scale sequencing tools, together with metabolomic and proteomic information will help to identify candidate genes involved in the biosynthesis of biologically active metabolites. The Medicinal Plant Genomics Consortium and genome-guided investigation [27,50], the Medicinal Plant Transcriptome Project [28], the 1000 Green Plant Transcriptome Project [29] etc. will add to identify biosynthetic pathways and their evolution in plants. These databases may be specially useful for new pathway discovery if a number of them are found as operon like gene clusters as described in barley, rice [30] and in opium poppy, where a 10-gene cluster involved in the biosynthesis of the antitumor alkaloid noscapine was located over 401 kb of genomic sequence [31]. Recently, using transcriptome data from the plant mayapple (*Podophyllum hexandrum*) and selecting candidate genes to combinatorially express in tobacco (*Nicotiana benthamiana*), six pathway enzymes of podophyllotoxin pathway to etoposide aglycone were identified [32]. Podophyllotoxin is the natural product precursor of the chemotherapeutic ‘unnatural’ anticancer etoposide, yet only part of its biosynthetic pathway is known (Fig. 1, modified from [32]). By coexpressing altogether 10 genes in tobacco, the pathway to etoposide aglycone, a naturally occurring lignan and intermediate precursor of etoposide was reconstituted. These works not only shows the expression of genes of etoposide precursor in a different plant species but also circumvent the need for cultivation of mayapple. Identification of another plant enzyme, olivetolic acid cyclase, a polyketide cyclase-like enzyme from transcriptome data of glandular trichomes on female cannabis flowers suggests that these

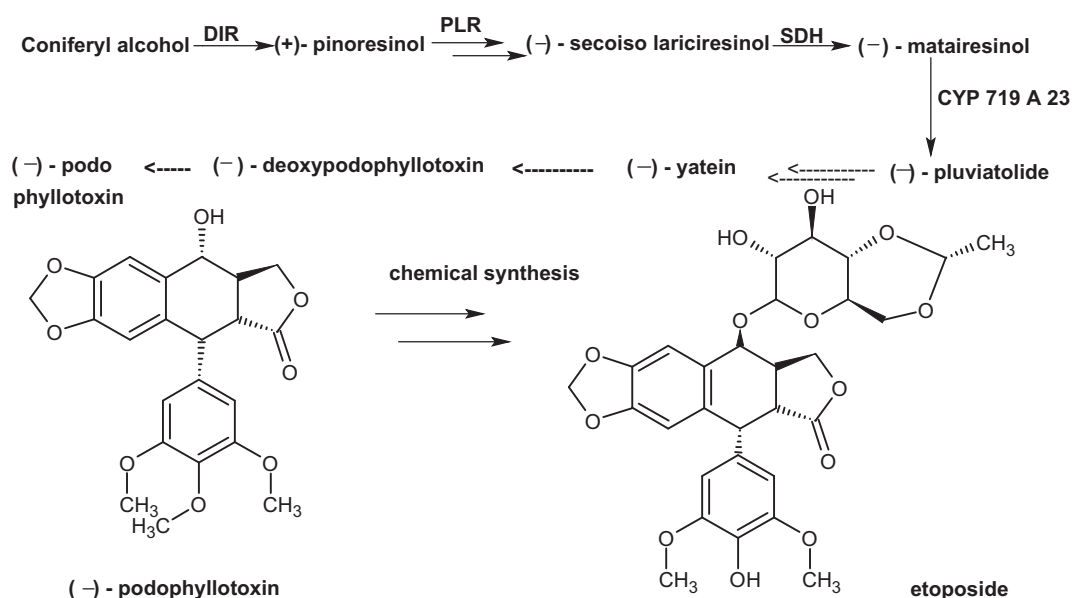


Fig. 1. Biosynthetic pathway of podophyllotoxin in *P. hexandrum*. Chemical conversion of anticancer etoposide from podophyllotoxin is also shown. Uncharacterized steps in the pathway are indicated by dashed lines. DIR: dirigent protein, PLR: pinoresinol lariciresinol reductase, SDH: secoisolariciresinol dehydrogenase CYP: cytochrome P 450.

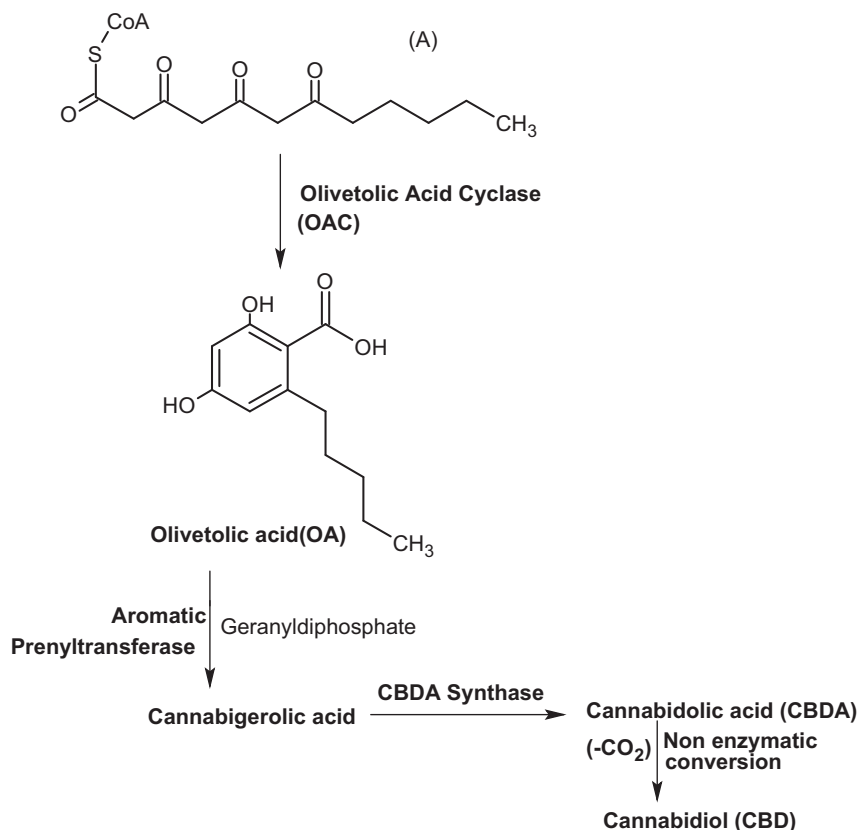


Fig. 2. Showing Olivetolic Acid Cyclase in the biosynthesis of cannabinoids, a unique biosynthetic route to plant polyketides. The enzyme cyclase catalyzes the formation of OA from hexanoyl CoA plus malonyl CoA (A). OA then converted to cannabigerolic acid and then to major cannabinoids, CBDA through enzymatic conversions. Finally, CBDA upon decarboxylation produces their neutral form CBD, cannabidiol.

cyclases may play an overlooked role in generating plant chemical diversity (Fig. 2, modified form of [34]). Humans have used *Cannabis sativa* L. (marijuana) as a medicinal and psychoactive herbal drug, and now it is the most widely consumed illicit drug worldwide [33]. Its unique effects are due to the presence of cannabinoids, and the first intermediate in the cannabinoid biosynthetic pathway is proposed to be olivetolic acid (OA), that forms the polyketide nucleus of the cannabinoids. During searching for polyketide cyclase-like enzymes, that could assist in OA cyclization, the enzyme, olivetolic acid cyclase was discovered [34]. Continued search on new enzymatic pathways and compounds having wide range of pharmacological activities, a plant-derived iridoid biosynthetic pathway enzyme, iridoid synthase was discovered [35]. To identify iridoid synthase, recently available transcriptomic data was used from a medicinal plant, *Catharanthus roseus* that produces a variety of iridoid-derived monoterpene indole alkaloids, including the anticancer alkaloid vinblastine [36]. The iridoids comprise a large family of distinctive bicyclic monoterpenes that possess anticancer, anti-inflammatory and anti-bacterial activities [37–39]. During iridoid cyclization step, iridoid synthase uses linear monoterpene 10-oxogeranial as substrate which is in contrast to all known monoterpene cyclases that uses geranyl diphosphate as substrate, and the enzyme synthase probably couples an initial NAD(P)H-dependent reduction via a Diels-Alder

cycloaddition or a Michael addition (Fig. 3, modified from [36]). This work not only suggests alternative biochemical mechanism for the biosynthesis of cyclic terpenes but also enable the large-scale heterologous production of iridoids in plants. Furthermore, using medicinal plant metabolic pathway enzymes other works are carried out for new drug discovery. For example, through the use of recombinant technology, combining enzyme engineering, genes from microorganisms and plant genetic engineering, search for new classes of compounds and their potentials as new drugs were carried out. For this, an increasing number of enzymes of genes involved in the biosynthesis of plant medicinal products have been crystallized. From the information of the crystal structure of strictosidine synthase, mutants having catalytic activity for broader substrates, together with soil bacteria halogenases were expressed in *Catharanthus hairy roots*. Transformed hairy roots made chlorinated tryptophan, which was converted via decarboxylase reaction into chlorotryptamine derivatives and into chlorinated monoterpene indole alkaloids including 10-chlorocatharanthine and 15-chlorotabersonine [40](Fig. 4). This work showed that plant can serve as chemical factories to produce many non-natural monoterpene indole alkaloids that can be used to produce new monoterpene indole alkaloid-based drugs.

Thus, the creation of new pathways in plants and the introduction of new reactions [40] or suppression of existing

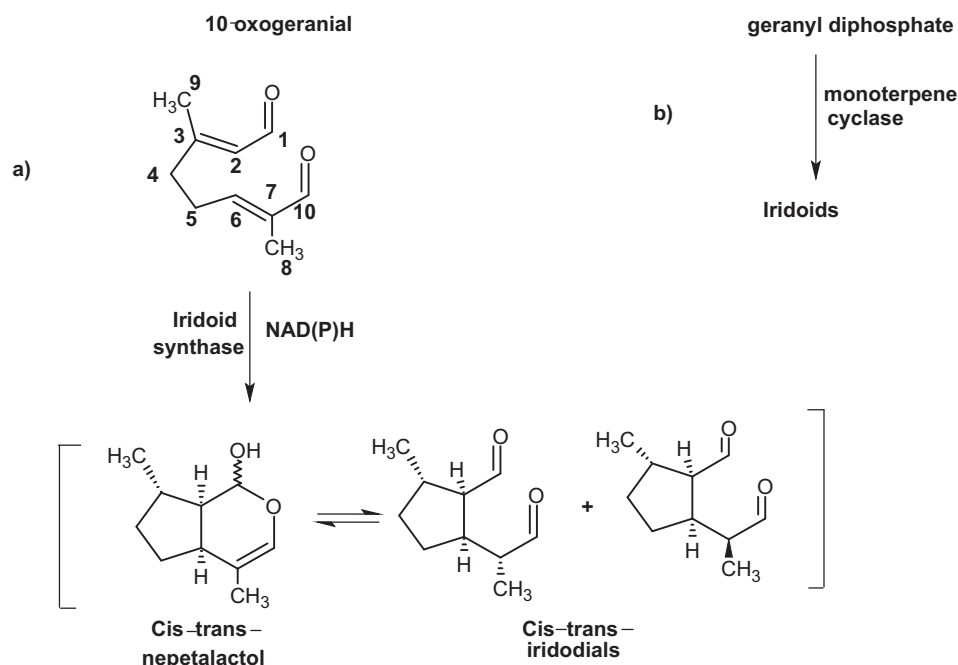


Fig. 3. Iridoid synthase from *C. roseus* in alternate biochemical mechanism of iridoid biosynthesis. a) The synthase uses monoterpene 10 oxogeraniol as substrate and NAD(P)H reduction step for cyclization, whereas b) monoterpene cyclases uses geranyl diphosphate as substrate in the usual process of cyclization.

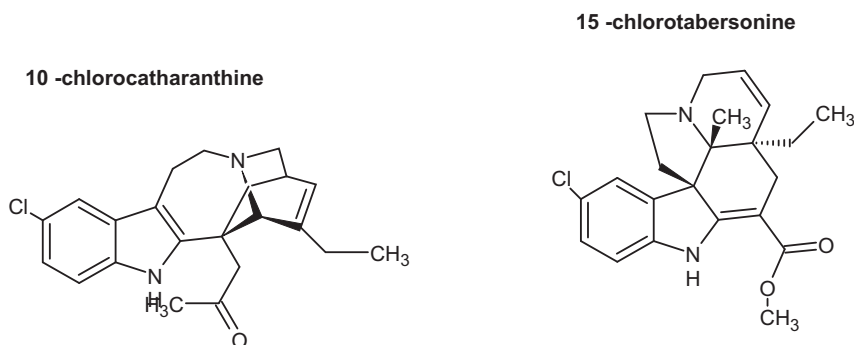


Fig. 4. Valuable secondary metabolites produced by pathway engineering in plants. Chlorinated monoterpene indole alkaloids are generated from transformed hairy roots of *Catharanthus*.

ones [22,41,42] can significantly affect for randomly generating previously unknown molecules. These could be biosynthetic intermediates from an established pathway, or totally new products could be developed from these intermediates [43]. The development of inexpensive genome sequencing together with proteomic and metabolomic information adds to new pathway discovery and may help to identify gene function across plant species. Such studies will definitely help in future to reveal new biologically active secondary metabolites and will take care of nature's plant biodiversity for new drug discovery.

4. Concluding remarks

There are now slogans that with tens of thousands of plant species on earth, we are blessed with an enormous wealth of

herbal medicine originated from nature and hence under green therapy cover and with no side effect, we can take it in the long term and so forth. Therefore, more and more attention in the field of drug discovery has been focused on the herbal medicine, considering low success rate, high cost and long time of conventional drug development. Plants display an immense diversity of specialized metabolites many of which have been important to humanity as medicines, flavors, fragrances etc. However, only a small fraction of the vast diversity of plant metabolism has been explored. For the speedy discovery of unknown metabolic pathways/enzymes, interdisciplinary research on herbal genomics together with large-scale sequencing tools and metabolomics and proteomics is absolutely necessary. We hope, such studies commits to reveal new biologically active secondary metabolites, making use of vast aspects of plant biodiversity for new drug discovery.

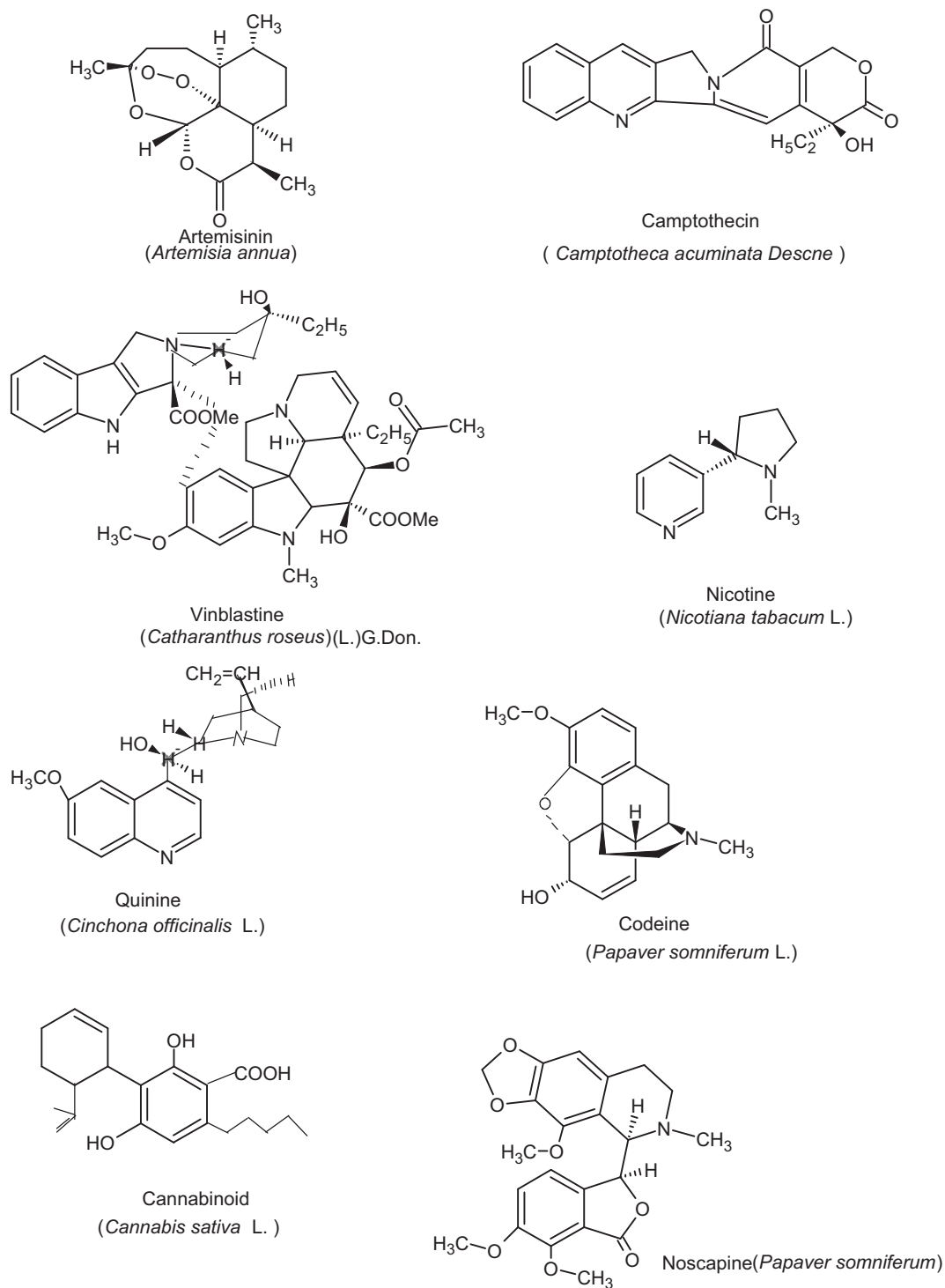


Fig. 5. Some important structures from medicinal plants and the names of plants that produce them.

Conflicts of interest

The author declared no conflict of interest with respect to the authorship and/or publication of this article.

Acknowledgments

The author gratefully acknowledges the overwhelming support from American Center Library, Kolkata, and National Library, Kolkata.

References

- [1] S. Chen, J. Song, C. Sun, J. Xu, Y. Zhu, R. Verpoorte, T.-P. Fan, Herbal genomics: examining the biology of traditional medicines, *Science* 347 (2015) 527–529.
- [2] B. Patwardhan, D. Warude, P. Pushpangadan, N. Bhatt, Ayurveda and traditional Chinese medicine: a comparative overview, *Evid Based Comple. Alter Med.* 2 (2005) 465–473.
- [3] M. Modak, P. Dixit, J. Londhe, S. Ghaskadbi, T.P.A. Devasagayam, Indian herbs and herbal drugs used for the treatment of diabetes, *J. Clin. Biochem. Nutr.* 40 (2007) 163–173.

- [4] V.D. Luca, V. Salim, S. Masada, F. Yu, Mining the biodiversity of plants: a revolution in the making, *Science* 336 (2012) 1658–1661.
- [5] C. Chang, J.C. Bowman, E.M. Meyerowitz, Field guide to plant model system, *Cell* 167 (2016) 325–339.
- [6] V. Brower, Back to Nature: extinction of medicinal plants threatens drug discovery, *J. Natl. Cancer Inst.* 100 (2008) 838–839.
- [7] P. Reymond, H. Weber, M. Damond, E.E. Farmer, Differential gene expression in response to mechanical wounding and insect feeding in *Arabidopsis*, *Plant Cell* 12 (2000) 707–719.
- [8] D. Hermsmeier, U. Schittko, I.T. Baldwin, Molecular interactions between the specialist herbivore *Manduca sexta* and its natural host *Nicotiana attenuata*. I. Large- scale changes in the accumulation of growth- and defense- related plant mRNAs, *Plant Physiol.* 125 (2001) 683–700.
- [9] S. Chen, J. Xu, C. Liu, Genome sequence of the model medicinal mushroom *Ganoderma Lucidum*, *Nat. Commun.* 3 (2012) 913–922.
- [10] L.A. Giddings, A stereoselective hydroxylation step of alkaloid biosynthesis by a unique cytochrome P450 in *Catharanthus roseus*, *J. Biol. Chem.* 286 (2011) 16751–16757.
- [11] S. Rastogi, A. Kalra, V. Gupta, Unravelling the genome of Holy basil an incomparable “elixir of life” of traditional Indian medicine, *BMC Genom.* 16 (2015) 413–431.
- [12] S. Rastogi, S. Meena, A. Bhattacharya, De novo sequencing and comparative analysis of holy and sweet basil transcriptomes, *BMC Genom.* 15 (2014) 588–603.
- [13] L. Chae, T. Kim, R. Nilo-Poyanco, Genomic signatures of specialized metabolism in plants, *Science* 344 (2014) 510–513.
- [14] D.J. Kliebenstein, A role for gene duplication and natural variation of gene expression in the Evolution of metabolism, *PLoS One* 3 (2008) e1838, 1–13.
- [15] Supplementary materials of Ref. [12], *Science* 344 (2014) 1–31.
- [16] V.D. Luca, V. Salim, S.M. Atsumi, Biodiversity of plants: a revolution in the plant biology, *Science* 336 (2012) 1658–1661.
- [17] R.S. Allen, A.G. Millgate, J.A. Chitty, RNAi-mediated replacement of morphine with the non narcotic alkaloid reticuline in opium poppy, *Nat. Biotechnol.* 22 (2004) 1559–1566.
- [18] F.G. Ratcliff, S.A. MacFarlane, D.C. Baulcombe, Gene silencing without DNA: RNA mediated cross-protection between viruses, *Plant Cell* 11 (1999) 1207–1215.
- [19] J.M. Hagel, P.J. Facchini, Dioxygenases catalyze the O-demethylation steps of morphine biosynthesis in opium poppy, *Nat. Chem. Biol.* 6 (2010) 273–275.
- [20] R. Li, D.W. Reed, E. Liu, Functional genomic analysis of alkaloid biosynthesis in *Hyoscyamus Niger* reveals a cytochrome P450 involved in littorine rearrangement, *Chem. Biol.* 13 (2006) 513–520.
- [21] J. Roepke, V. Salim, M. Wu, Vinca drug components accumulate exclusively in leaf exu-dates of Madagascar periwinkle, *Proc. Natl. Acad. Sci. USA* 107 (2010) 15287–15292.
- [22] D.K. Liscombe, S.E. O'Connor, A virus-induced gene silencing approach to understanding alkaloid metabolism in *Catharanthus roseus*, *Phytochemistry* 72 (2011) 1969–1977.
- [23] C.J. Paddon, P.J. Westfall, D.J. Pitera, High-level semi-synthetic production of the potent antimalarial artemisinin, *Nature* 496 (2013) 528–532.
- [24] K. Thodey, S. Galanie, C.D. Smolke, A microbial biomanufacturing platform for natural and semisynthetic opioids, *Nat. Chem. Biol.* 10 (2014) 837–844.
- [25] S. Brown, M. Clastre, V. Courdavault, De novo production of the plant-derived alkaloid strictosidine in yeast, *Proc. Natl. Acad. Sci. USA* 112 (2015) 3205–3210.
- [26] Y. Qu, M.L.A.E. Easson, J. Froese, Completion of the seven step pathway from taberson- nine to the anticancer drug precursor vindoline and its assembly in yeast, *Proc. Natl. Acad. Sci. USA* 112 (2015) 6224–6229.
- [27] A. Kellner, Genome sequence of *Catharanthus roseus*, 2015. <http://medicinalplantgenomics.msu.edu/>.
- [28] Transcriptome characterization, sequencing, and assembly of medicinal plants relevant to human health, <http://uic.edu/pharmacy/med pl transcriptome/index.html>.
- [29] Plants 1000 (one KP or 1 KP) gene sequencing data, www.onekp.com.
- [30] H.Y. Chu, E. Wegel, A. Osbourn, From hormones to secondary metabolism : the emergence of metabolic gene clusters in plants, *Plant J.* 66 (2011) 66–79.
- [31] T. Wizer, V. Gazda, Z. He, A *Papaver somniferum* 10-Gene cluster for synthesis of the anticancer alkaloid noscapine, *Science* 336 (2012) 1704–1708.
- [32] W. Lau, E.S. Sattely, Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone, *Science* 349 (2015) 1224–1228.
- [33] United Nations office on drugs and crime, World Drug Report 2010, United Nations office on drugs and crime, Vienna, Austria, 2010.
- [34] S.J. Gagne, J.M. Stout, E. Liu, Identification of olivetolic acid cyclase from *Cannabissativa* reveals a unique catalytic route to plant polyketides, *Proc. Natl. Acad. Sci. USA* 109 (2012) 12811–12816.
- [35] F. Geu-Flores, N.H. Sherden, V. Courdavault, An alternative route to cyclic terpenes by reductive cyclization in iridoid biosynthesis, *Nature* 492 (2012) 138–142.
- [36] B. Dinda, R. Chowdhury, D. Mohanta, Naturally occurring iridoids, secoiridoids and their bioactivity, *Chem. Pharm. Bull. (Tokyo)* 57 (2009) 765–796.
- [37] B. Dinda, S. Debnath, R. Banik, Naturally occurring iridoids and secoiridoids, *Chem. Pharm. Bull. (Tokyo)* 59 (2011) 803–833.
- [38] B. Dinda, S. Debnath, Y. Harigaya, Naturally occurring secoiridoids and bioactivity natu- rally occurring iridoids and secoiridoid, *Chem. Pharm. Bull. (Tokyo)* 55 (2007) 689–728.
- [39] R. Tundis, M.R. Loizzo, F. Menichini, Biological and pharmacological activities of iridoids: recent developments, *Med. Chem.* 8 (2008) 399–420.
- [40] W. Runguphan, X. Qu, S.E. O'Connor, Integrating carbon-halogen bond formation into medicinal plant metabolism, *Nature* 468 (2010) 461–464.
- [41] L.C. Hileman, S. Drea, G. Martino, Virus-induced gene silencing is an effective tool for assaying gene function in the basal eudicot species *Papaver somniferum* (opium poppy), *Plant J.* 44 (2005) 334–341.
- [42] S. Wege, A. Scholz, S. Gleissberg, Highly efficient virus-induced gene silencing in California poppy (*Eschscholzia californica*): an evaluation of VIGS as a strategy to obtain functional data from non-model plants, *Ann. Bot. (Lond)* 100 (2007) 641–649.
- [43] P.J. Facchini, J. Bohlmann, P.S. Covello, Synthetic biosystems for the production of high-value plant metabolites, *Trends Biotech.* 30 (2012) 127–131.
- [44] S. Chen, Y.Z. Sun, J. Xu, et al., Strategies of the study on herb genome program, *Yao Xue Xue Bao* 45 (2010) 807–812.
- [45] S. Chen, L.X. Xu, Q.L. Guo, An introduction to the medicinal plant genome project, *Front. Med.* 5 (2011) 178–184.
- [46] R.T.C.M. Mahajan, Phyto-pharmacology of *Ziziphus Jujuba* Mill-a plant review, *Pharmacol. Rev.* 3 (2009) 320–329.
- [47] Y. Shivaraj, S. Govind, S. Jogaiah, et al., Functional analysis of medicinal plants using system biology approaches, *Int. J. Pharm. Pharmaceut. Sci.* 7 (2015) 41–43.
- [48] Y. He, H. Xiao, C. Deng, et al., The complete chloroplast genome sequences of the medicinal plant *Pogostemon cablin*, *Int. J. Mol. Sci.* 17 (2016) 820–830.
- [49] J. Qian, J. Song, H. Gao, et al., The complete chloroplast genome sequences of the medicinal plant *Salvia miltiorrhiza*, *PLoS One* 8 (2013) e57607.
- [50] F. Kellner, J. Kim, B.J. Clavijo, et al., Genome-guided investigation of plant natural product biosynthesis, *Plant J.* 82 (2015) 680–692.
- [51] Y. Li, C. Xu, X. Lin, De novo assembly and characterization of the fruit transcriptome of Chinese *Jujuba* (*Ziziphus Jujuba* Mill) using 454 pyrosequencing and the development of novel trinucleotide SSR markers, *PLoS One* (2014c), <https://doi.org/10.1371/journals.Pone.0106438>.
- [52] H. Xu, J. Song, H. Luo, et al., Analysis of the genome sequence of the medicinal plant *Salvia miltiorrhiza*, *Mol. Plant* 9 (2016) 949–952.