Methods for Molecular Identification of Biosynthetic Enzymes in Plants

Chapter · January 2009

DOI: 10.1007/978-0-387-85498-4_7

1 author:

Sarah o'connor
John Innes Centre
135 PUBLICATIONS **3,532** CITATIONS

SEE PROFILE

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

Methods for Molecular Identification of Biosynthetic Enzymes in Plants

7

Sarah O'Connor

AU1 **Abstract**

Introduction

Plants produce some of the most important natural products used in medicine (Fig. 1). Plant-derived compounds have been used as powerful pharmaceuticals throughout the course of human history. For example, opium poppy has been used since neolithic times [1]; aspirin, synthesized by acetylation of salicyclic acid from willow bark, was discovered in the late nineteenth century [2]; and the powerful chemotherapy agent taxol, used in the treatment of advanced breast cancer, was discovered from the yew tree in the mid-twentieth century [3].

The study of microbe derived natural products underwent a revolution in the 1980s as genetic strategies allowed the rapid discovery of the genes that encode natural product biosynthetic pathways [4]. This genetic information allows a detailed understanding of the chemical and biochemical mechanisms that Nature uses to construct complex molecules. Microbial metabolic pathways can now be reengineered at the genetic level to produce higher levels of natural products, or novel natural product derivatives with improved medicinal properties [5].

Analogously, if the biosynthetic enzymes leading to plant derived natural products are cloned, we can begin to understand how plants construct complex molecules [6]. Furthermore, genetic information is crucial for metabolic engineering efforts to increase the production levels of these compounds. For example, expression of plant biosynthetic pathways in fast growing organisms such as *Escherichia coli* or *Saccharomyces cerevisiae* could be explored, as it has been successfully achieved for several plant derived flavonoid [7] and terpenoid natural products [8]. Additionally, many plant biosynthetic pathways are highly branched, leading to the formation of "side products" that lack the desired bioactivity of the target natural product. If the enzymes at the branch points responsible for side product biosynthesis are cloned, they could be downregulated and production of the desired products in plants could be improved [9]. Alternatively, genes encoding biosynthesis of the desired product could be overexpressed [10]. Additionally, isolation of compounds from natural sources provides limited opportunities to modify their chemical and biological properties. Cloning the enzymes that catalyze natural product biosynthesis may allow reprogramming of the biosynthetic pathway to produce new unnatural products with improved

Sarah O'Connor (✉)
Massachusetts Institute of Technology,
Department of Chemistry,
18-592, 77 Massachusetts Ave,
Cambridge MA 02139, USA
e-mail: soc@mit.edu

A.E. Osbourn and V. Lanzotti (eds.), *Plant-derived Natural Products: Synthesis, Function, and Application*,
DOI 10.1007/978-0-387-85498-4_7,

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

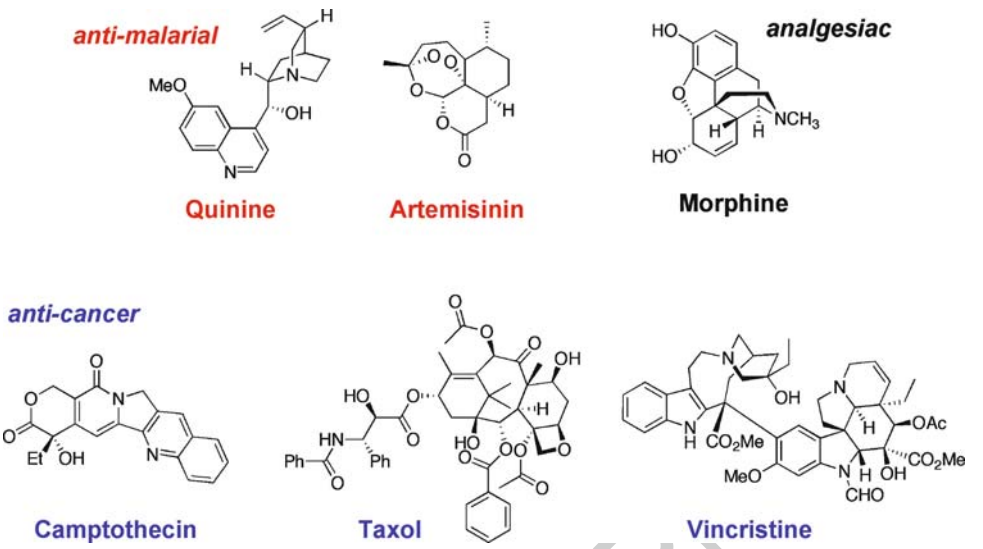

Fig. 1 Plant natural products. A selection of representative plant derived natural products that have important pharmaceutical activities

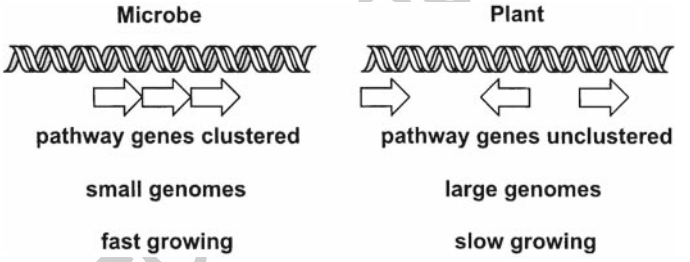

Fig. 2 Challenges of plants. Several challenges are encountered in plant genetics that have made pathway elucidation more difficult than pathway elucidation in microbes

pharmaceutical activities [11]. Finally, the expression levels of biosynthetic genes, and the levels of the corresponding natural products, are controlled by transcription factor proteins, which interact with the promoter regions of target genes and modulate the rate of mRNA production. If the transcription factors that control expression of biosynthetic genes are identified, then these proteins could be used to upregulate biosynthetic gene expression and natural product production levels [12].

Although plants produce a large fraction of known natural products, the number of plant pathways that have been elucidated pales in comparison to the number of well-characterized microbial pathways [13]. Elucidating the genes of a plant pathway is a task fraught with significant challenges, and the genetic tools that have unlocked the biosynthesis of countless bacterial and fungal natural products are not easily applied to plants. In contrast to microbes, the genes of plant pathways – with few exceptions - [14] are not clustered on the genome, so each gene of a plant pathway must be discovered individually (Fig. 2). Additionally, the genome sizes of medicinal plants are much larger

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

(>1,000 Mbp) than the typical natural product producing bacteria (~8 Mbp), which makes finding and screening putative biosynthetic genes a daunting undertaking [15]. Direct purification of plant enzymes is complicated by low levels of protein expression, the presence of phenolic compounds and proteases that readily inactivate enzyme activity and the frequent requirement of accessory proteins or cofactors for catalysis. Finally, plants are slow growing, and the process of genetic transformation and subsequent selection is time-consuming. Although spectacular successes have been achieved in elucidating plant pathways, the challenges of plant biology have hindered the study of plant secondary metabolism (Fig. 2).

This review provides an overview of the major strategies used to identify the genes that comprise plant natural product biosynthetic pathways. One or two key examples that illustrate these strategies are provided though many more examples are available for most of the techniques outlined. A range of enzyme classes been identified using the approaches described in this review including P450 enzymes [16], terpene synthases [17], and glycosyltransferases [18]. Notably, many of the strategies for gene identification described below require prior knowledge of the biochemical transformation that the corresponding gene product catalyzes. This chemical knowledge can allow prediction of the enzyme class so that the gene can be identified in homology based cloning strategies. Additionally, the biochemical reaction must be known for design of enzymatic assay for *in vitro* enzymology. The biochemistry is typically elucidated by feeding isotopically labeled precursors to the plant or plant culture and mapping the placement of the isotopes within the final natural product. Additionally, isolation of biosynthetic intermediates as well as biomimetic synthetic chemistry provides insights into the chemical transformation that occur.

Methods of Pathway Elucidation in Plants

Traditional Biochemical Approaches

Countless plant biosynthetic enzymes have been identified using a classical approach (Fig. 3). The desired enzyme is purified from a crude plant or plant cell culture lysate by traditional protein chromatography. The enzyme is typically monitored throughout the purification procedure by an *in vitro* enzymatic assay. Once a homogenous preparation of enzyme is prepared, amino acid sequences are obtained by subjecting the protein, or tryptic fragments of the protein, to N-terminal

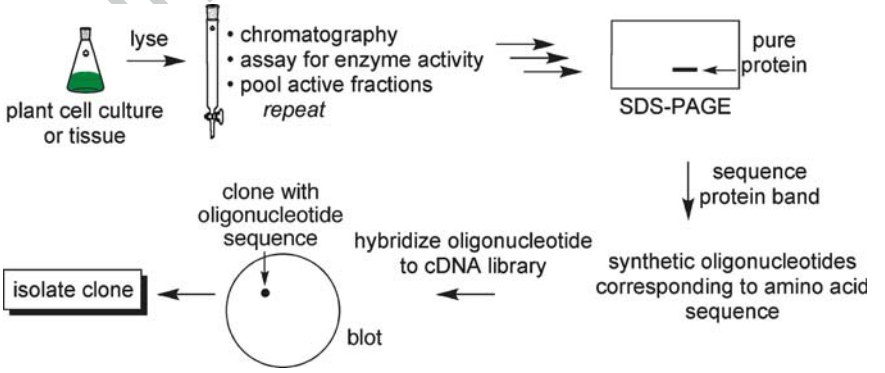

Fig. 3 Classical purification approaches. In this strategy, the desired enzyme is purified by traditional biochemical chromatography, and protein sequence information is used to isolate the desired gene from plant cDNA

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

sequencing or mass spectrometry. Oligonucleotide primers that correspond to the identified regions of protein sequence are designed, and these primers are used to identify the gene encoding the desired enzyme in a cDNA library. Usually, the identified gene is then expressed in a heterologous expression system and assayed *in vitro* to confirm that the gene encodes the enzyme having the expected function.

Example

The berberine bridge enzyme utilized in tetrahydroisoquinoline alkaloid pathways was cloned by this method [19]. This enzyme was purified from elicited *Eschscholzia californica* cell suspension cultures in six to seven steps, where approximately 10 µg of pure protein was obtained from 700 mg of crude protein. The purified protein was digested with trypsin and the resulting peptides were subjected to N-terminal sequencing. An oligonucleotide primer based on one of these amino acid sequ-ences was shown to hybridize to a clone from an *E. californica* cDNA library. This clone was sequenced and shown to encode the amino acid sequences observed by N-terminal sequencing of the purified protein. The clone was then heterologously expressed in *S. cerevisea* and was shown to catalyze the expected biochemical activity – formation of (*S*)-reticuline to (S)-scoulerine – as evidenced by a number of spectroscopic characterizations.

Notably, natural product biosynthesis is often localized to one cell type of the plant. This compartmentalization can be exploited to obtain cell lysates that are highly enriched in the cell type harboring the enzyme of interest. For example, many of the enzymes involved in vinblastine biosynthesis are localized to the epidermal layer of *Catharanthus roseus* [20]. By selectively harves-ting the epidermal cells of *C. roseus* leaves using an abrasion technique, a highly enriched fraction of 16-hydroxytabersonine-*O*-methyltransferase could be purified in just three chromatographic steps [21].

Scope and Limitations

This classical cloning strategy has proven to be successful in many cases, but it is time-consuming and laborious. Purification of a native enzyme to homogeneity is difficult, and any undesired contaminant protein can be mistakenly subjected to sequencing. Cloning efforts from plant material expressing low levels of the desired enzyme are often unsuccessful. Plant tissues that are known to be enriched in the desired activity provide the best results. Furthermore, the enzymatic activity must be maintained *in vitro* as the desired enzyme is diluted and purified. If accessory proteins or cofactors are required for stability or functional activity, the purified protein will become inactive as it is isolated from the crude cell extract. A robust *in vitro* assay with correct substrates is also required, and this may not be possible if the substrates are unavailable or if the enzymatic products are unstable.

An additional limitation applies to *in vitro* enzyme assays described here and in all subsequent sections. Results of *in vitro* assays often contradict results obtained in *in vivo* studies. For example, although transcript and metabolite profiling indicated that a glycosyltransferase from *Medicago truncatula* is involved in the biosynthesis of triterpene saponins, this enzyme glycosylated certain phenolic compounds with higher efficiency than triterpenes *in vitro* [22].

Homology-Based Screens

Enzymes within a given class often have highly conserved regions in the protein sequence. Oligonucleotide primers complementary to these consensus sequences can be designed. These

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

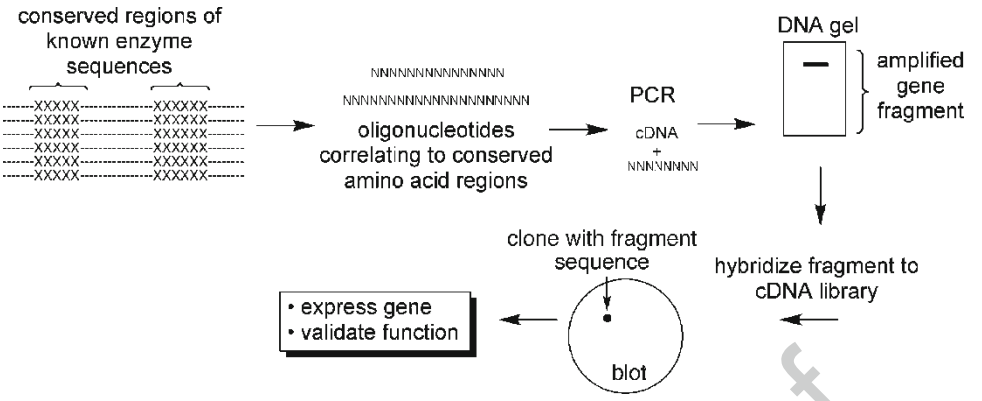

Fig. 4 Homology based cloning. Conserved regions of known enzyme classes are used to design oligonucleotide primers, which are used to isolate the desired gene from plant cDNA

primers can be used to amplify the genes having the corresponding consensus sequence from cDNA libraries by polymerase chain reaction (PCR) (Fig. 4). This method, known as homology based cloning, has been widely used to elucidate genes of biosynthetic pathways. P450 enzymes, glycosyltransferases, acetyltransferases, and prenyl transferases have all been successfully cloned from plant genetic material using primers designed to recognize the known consensus sequences of these enzyme classes.

Example

A cDNA clone encoding a taxane 2a-*O*-benzoyltransferase that catalyzes a late-stage acylation step in the taxol biosynthetic pathway was isolated from *Taxus cuspidata* by employing a homology-based PCR cloning strategy [23]. After generating oligonucleotide primers specific for acyltransferases, several gene fragments were amplified by PCR and these fragments were used to screen a cDNA library constructed from mRNA isolated from induced *Taxus* cells. Several full-length acyltransferases were obtained and each was individually expressed in *E. coli* and assayed for function *in vitro*. These assays indicated that one of these genes encoded an enzyme having taxane 2a-*O*-benzoyltransferase activity.

P450 enzymes, which catalyze a wide range of oxidative transformations in many plant biosynthetic pathways, contain highly conserved regions that can be used to identify the corresponding genes from a library of clones [24]. For example, the P450 enzyme (*S*)-*N*-methylcoclaurine-3'-hydroxylase of tetrahydroisoquinoline alkaloid biosynthesis was cloned from *E. californica* cDNA using a primer based on a P450 consensus sequence [25]. Notably, membrane bound P450 enzymes are usually present in plant cell cultures in very low quantities. Consequently, purified preparation of these enzymes are not readily obtained, so biochemical purification approaches (Section 2.1) are often not possible with enzymes of this class [26]. Homology based cloning strategies, which circumvent the native purification process, have therefore been particularly crucial for discovery of new P450 enzymes involved in plant biosynthetic pathways.

Scope and Limitations

Homology-based cloning is limited to enzymes with a known conserved sequence. If no consensus

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

sequence for the enzyme has been reported – as is the case for rare or novel enzymes – homology based cloning strategies cannot be used. Moreover, the biochemistry of the biosynthetic pathway must be sufficiently well understood to hypothesize which classes of enzyme are involved in the biosynthetic transformations. Additionally, certain enzyme classes, such as P450s, are ubiquitous and homology cloning can lead to the amplification of many genes. Therefore, the gene encoding the desired biosynthetic activity must be identified out of many that are cloned. Again, the quality of the plant material from which the mRNA is harvested is critical to ensure that the desired enzyme is present in high quantities in the library. The enzyme function and specific substrate transformation also need to be validated after isolation of the gene, either by *in vitro* enzyme assay, or by knockdown of the gene in the producing plant (see Section 2.6).

Functional Genomic Approach

As DNA sequencing technology has improved, numerous cDNA libraries can be sequenced within a relatively short time frame. The sequence information collected from the plant derived cDNA can then be compared to a gene database. If homologous genes have been functionally characterized, then a prediction about the biochemical function of the newly sequenced plant gene can be made. Often only short fragments of DNA, termed expressed sequence tags (EST), are sequenced. The EST, typically about 500 base pairs, provides enough sequence information for a protein function to be predicted by homology. The fragments that appear to encode enzymes involved in the natural product biosynthetic pathway are then used to clone the full length genes from a cDNA library. The open reading frame of an isolated cDNA can then be overexpressed in a heterologous organism such as *E. coli* and the recombinant protein can be assayed for function.

Example

The taxol biosynthetic pathway has been partially elucidated by a highly successful random library sequencing effort [27]. In this study, sequencing of 10,000 transcripts derived from an induced *Taxus cuspita* cell suspension culture led to the identification of approximately 3,500 unique ESTs. Sequence analysis indicated that several of the known genes of taxol biosynthesis were present in high abundance in this EST collection. Since many of the taxol biosynthetic steps are predicted to be catalyzed by P450 enzymes, the EST collection was searched for clones having homology to known P450 sequences. Nearly 100 unique P450s were identified by this homology search, out of which 19 displayed homology to previously identified enzymes involved in the hydroxylation steps of the taxol biosynthetic pathway. Ten of these hydroxylase-like P450 genes had never been observed in previous attempts to discover new taxol biosynthetic genes. Each of the ten full length genes was heterologously expressed in yeast. Subsequent *in vitro* enzyme assays suggested that one of these genes encoded a novel taxoid 10-hydroxylase while another encoded a taxoid 2-hydroxylase enzyme.

Importantly, these genomic libraries can be enriched for the appropriate gene transcripts if the tissue or cell type from which the genetic material is harvested corresponds to the site of natural product production. For example, 25% of the clones in cDNA obtained exclusively from the oil gland secretory cell of peppermint (*Mentha* × *piperita*) appear to be involved in oil metabolism [28]. In another example, mRNA for construction of a *C. roseus* cDNA library obtained from epidermal cells resulted in a collection enriched in genes involved in alkaloid biosynthesis [21].

Scope and Limitations

Until recently, a major disadvantage of this strategy was the expense and time incurred in large scale sequencing efforts, but modern

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

high-throughput facilities can rapidly sequence thousands of clones per day. However, after the genetic information is obtained, strategies to identify the relatively small number of biosynthetic genes out of the large amount of sequenced clones must be developed. As with homology based cloning (Section 2.2), prediction of gene function by homology is not a useful way to identify enzymes that catalyze novel biochemical transformations; for example, in the taxol study, since little is known about the biochemical mechanism of a ring expansion that occurs late in taxol biosynthesis, analysis of the ESTs failed to provide any insight for this intriguing step. As with the other methods outlined, plant material with high levels of secondary metabolite enzyme expression provide the best results.

Analysis of Metabolic, Proteomic and Genomic Networks

The information that is obtained from large scale library sequencing as described in Section 2.3 can be further refined by clustering gene expression levels with metabolite production. By comparing the appearance of metabolites with gene expression levels (as indicated by the abundance of mRNA), a metabolic network that correlates genes with natural product biosynthesis can be obtained. This correlation network introduces an additional layer of refinement when analyzing the ESTs for genes involved in natural product production; transcripts that correlate with the appearance of natural products are the ones most likely to be part of the biosynthetic pathway [29].

Genes are randomly sequenced from the desired plant genetic material, and gene functions are predicted by homology when possible. The desired natural product, or, as is the case with many plant product pathways, a mixture of biosynthetically related natural products, is simultaneously monitored by mass spectrometry coupled with liquid chromatography. Genes that are upregulated with the advent of metabolite appearance are presumed to be involved in the biosynthesis or regulation of the natural product production (Fig. 5).

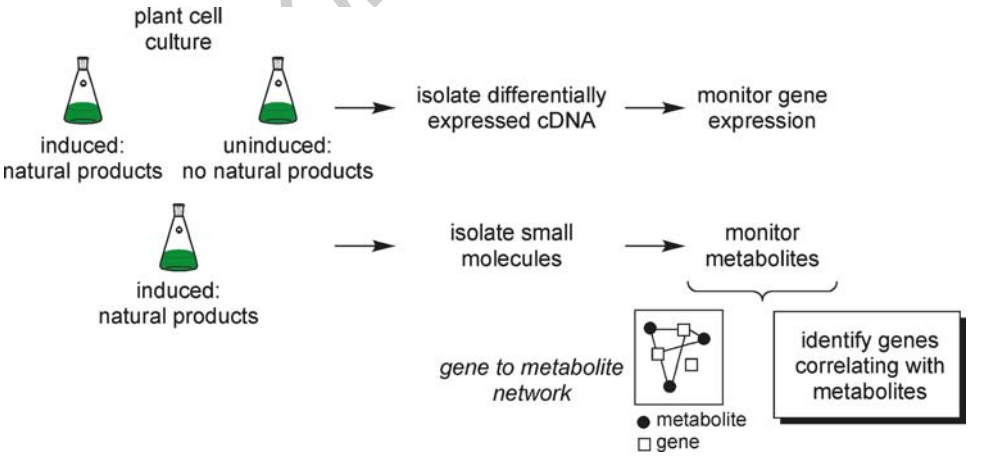

Fig. 5 Metabolic networks. Gene expression levels are correlated with the appearance of natural product production. This correlation provides clues as to which genes of the plant are involved in natural product biosynthesis

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

Example

A recent example illustrating the power of the genomics and metabolomics combination was recently demonstrated for *C. roseus*, which produces approximately 100 biosynthetically related alkaloid natural products [30]. A total of 178 metabolites were observed by mass spectrometry and nine of these metabolites were positively identified as alkaloid natural products by comparison with authentic standards. In the genomic analysis, approximately 400 unique genes were found from *C. roseus* mRNA. Almost 40% of these ESTs had no sequence homology to any genes in public databases, suggesting that novel enzymes are found within the *C. roseus* plant. Gene to metabolite networks were constructed from this information (as depicted in Fig. 5), making it clear which gene sequences were correlated with the appearance of natural products. Gene to gene networks provided information about groups of genes that had correlated expression levels. Although no conclusions regarding the mechanism of alkaloid biosynthesis were provided at the time of this report, this study provided a wealth of information that can be used to identify new genes involved in this biosynthetic pathway.

Scope and Limitations

As in Section 2.3, major disadvantages of this strategy includes the expense and time incurred in all large scale sequencing efforts. Additionally, plant material with different production levels of the desired metabolites must be available, so that gene expression levels from plant material with differential metabolite levels can be compared. Prediction of gene function by homology is not always possible, as described in previous sections, and the specific function of any putative biosynthetic enzyme candidates must be validated experimentally.

Subtracted cDNA Libraries

Genetic material from two related types of tissue can be "subtracted" from one another (Fig. 6). The genes that both sets of tissue have in common are identified, and the genes unique to each tissue type are readily obtained [31]. Subtractive hybridization techniques can be used to identify genes involved in plant secondary metabolism, provided that the appropriate plant tissue is available. For example, the genes in induced cell cultures – which produce natural products at high levels – can be compared with gene expression levels in uninduced cell cultures that produce low levels of natural products. Genes unique to the induced cell cultures will be likely to play a role in natural product biosynthesis catalysis or regulation. Additionally, different tissue types, such as root or leaf, often have different natural product production levels. Comparison of gene expression profiles from

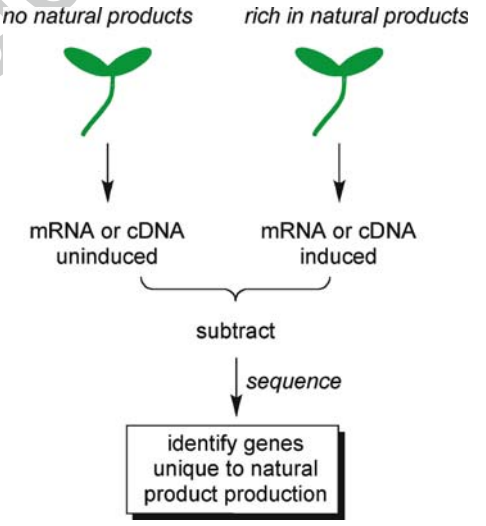

Fig. 6 Subtracted cDNA libraries. The gene expression levels of two plant tissue types, in which one produces natural products while the second does not, are compared. Genes found exclusively in the natural product producing tissue are likely to be involved in natural product production

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

two types of tissues from the same plant can also be used to predict which genes might play important roles in the biosynthesis of a particular natural product.

Example

A set of 60 cDNAs were isolated by subtractive hybridization of *Nicotiana tabacum* roots before and after removal of the flowers and young leaves, a process known as "topping" [32]. Topping is known to increase the amount of nicotine and other related alkaloids in the plant. A number of the subtracted cDNAs that were unique to the topped plants corresponded to known alkaloid biosynthetic enzymes. The functions of additional cDNAs observed in the subtracted library were not clear and remain under investigation. Further study of these new clones may provide insights into alkaloid natural product biosynthesis in *N. tabacum*.

A second example of subtractive hybridization took advantage of the variation of natural product production in two types of *C. roseus* tissue [33]. In *C. roseus*, roots and leaves produce a distinct spectrum of alkaloids. The leaf and root gene expression profiles were comparatively analyzed using subtractive hybridization. A total of 155 ESTs were subjected to homology-based classification and 16 EST sequences that had never been previously observed were obtained. Again, although new insights into the mechanism of the alkaloid pathway were not reported, functional analysis of these novel ESTs may provide new information into the mechanism of this alkaloid biosynthesis in *C. roseus*.

Scope and Limitations

This approach depends on the availability of tissues that have differential natural product expression. If natural products cannot be induced by an elicitor, or if various tissue types of the plant do not have different levels of natural product expression, then subtractive hybridization strategies cannot be applied. Additionally, subtractive hybridization, as with homology based cloning (Section 2.2) and large scale screening efforts (Sections 2.3 and 2.4), only provides putative biosynthetic genes that must be subjected to additional study. Validating the function of candidate genes must be addressed before the genes can be assigned a defined role in the biosynthetic pathway.

Forward Genetics Approach: Gene Suppression

Methods to suppress gene expression in plants have become widely available. For example, RNA interference (RNAi) can readily down-regulate gene expression in plants. Large scale RNAi screens can be used to rapidly assess the function of genes, provided that a fast screen or selection is available to interpret the phenotypes of the transformed lines (Fig. 7) [34]. Typically, cDNA fragments of approximately 500 bp in size are cloned into the appropriate vector and then transformed into the desired plant. These vector constructs can then suppress any plant gene that contains sequence corresponding to the fragment. The transformed plant material with a suppressed gene is then monitored for changes in phenotype. For example, cell cultures that exhibit decreased levels of natural products are likely to have been transformed with a fragment that suppresses a gene involved in natural product biosynthesis. If an efficient way to assess the natural product production of the transformed lines is available, large scale RNAi screens can be used to identify secondary metabolite genes.

Example

To investigate a late step in tropane alkaloid biosynthesis in *Hyoscyamus niger*, a subtracted

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

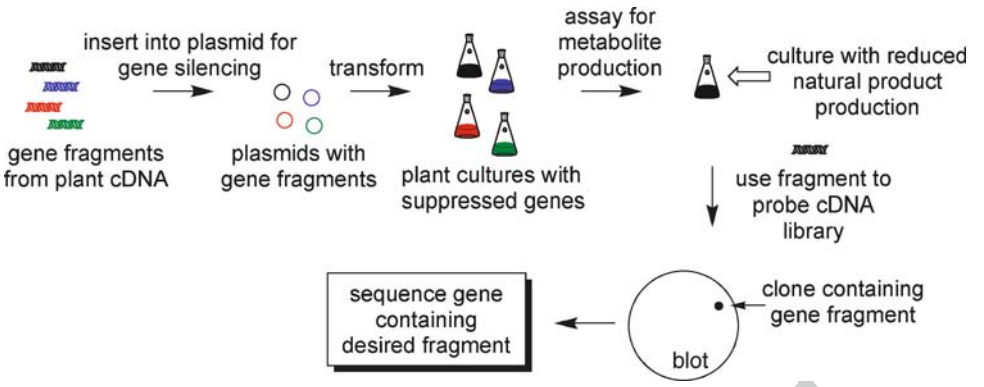

Fig. 7 Gene suppression. Gene fragments are used to silence genes in plant cell culture. If a fragment from a biosynthetic gene is used for silencing, then natural product production will be suppressed. The fragment can be used to identify the full length clone in a cDNA library

cDNA library of leaf and root was constructed to enrich the library in genes specifically expressed in the root, the site of tropane alkaloid biosynthesis [35]. The biosynthetic transformation of interest was predicted to be catalyzed by a P450 enzyme. A total of 25 ESTs from this cDNA library showed sequence homology to P450s, and the genes corresponding to these 25 ESTs were suppressed in *H. niger* using virus induced gene silencing. Virus-induced gene silencing exploits the RNA silencing pathway directed against invading viruses to silence host genes. Production of alkaloid products and alkaloid biosynthetic intermediates was then monitored by mass spectrometry in the suppressed cell lines. In one transformed line, formation of the final alkaloid product appeared to be inhibited, and the penultimate biosynthetic intermediate, littorine, was observed in large quantities. Gene function was validated by heterologous expression of littorine synthase in yeast and tobacco followed by an *in vitro* assay that validated enzymatic conversion of (*R*)-littorine to (*S*)-hyoscyamine.

Gene silencing can also be used to validate the function of a specific gene. For example, the cDNA CaMXMT1 encodes an *N*-methyltransferase involved in caffeine biosynthesis [36]. After suppression of this gene with RNAi in *Coffea Arabica*, caffeine biosynthesis was halted and the biosynthetic intermediate theobromine accumulated instead. This strongly suggested that the enzyme that CaMXMT1 encoded the *N*-methyltransferase enzyme that methylates theobromine to yield caffeine.

Scope and Limitations

Gene silencing requires that the plant be amenable to efficient transformation. Since a number of medicinal plants cannot be stably transformed by either *Agrobacterium* or particle bombardment technology, this strategy is unfortunately not universally applicable to the elucidation of all biosynthetic pathways of natural products. Additionally, suppression of a large number of genes requires a selection or screen that can rapidly assess the natural product profiles of the transformed cell lines.

Forward Genetics Approach: Expression Cloning

In expression cloning, a high quality cDNA library is constructed from the species of interest [37]. Plasmid DNA from this library is transformed

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

into a fast growing organism such as *E. coli* or *S. cerevisiae* for heterologous expression of the clones contained within the cDNA library. Transformed *E. coli* or *S. cerevisiae* cultures are then cultivated on solid media containing appropriate enzyme substrates to allow the detection of enzyme activity. If a culture that appears to catalyze the desired enzymatic reaction is detected, then the corresponding clone can be further analyzed by DNA sequence analysis and more thorough *in vitro* enzymatic assays (Fig. 8). Expression cloning can be performed in the absence of any knowledge of the enzyme to be cloned, and many different enzymes can be screened simultaneously, provided that efficient screening assays are available [38].

Example

Expression cloning has had widespread success in discovery of prokaryotic enzymes [32], but relatively few examples of expression cloning for elucidation of plant metabolism have been reported. In one example, a cDNA library constructed from genetic material of pumpkin seedlings (*Cucurbita maxima* L.) was functionally expressed in bacteriophage [39]. The plaques were screened for hybridization to a polyclonal antibody for GA 20-oxidase, an enzyme involved in the gibberellin biosynthetic pathway. This antibody was raised using enzyme that had been purified from plant material by traditional biochemical chromatography. Plaques that hybridized to the antibody also catalyzed the expected enzymatic activity *in vitro*.

In a second example, cDNA from *C. roseus* was functionally expressed in yeast in an attempt to clone secologanin synthase, an enzyme that catalyzes the conversion of loganin to secologanin [40]. Secologanin is a key precursor for the indole alkaloid natural products produced in *C. roseus*. The plant cDNA was transformed into a strain of yeast that also overexpressed two alkaloid biosynthetic enzymes that convert secologanin into a bright, readily detectable yellow pigment. Yeast transformed with the cDNA and these two known enzymes were incubated with loganin and monitored for the appearance of a yellow color. Although no putative secologanin synthase candidate gene was obtained, the study demonstrated how a well-designed colorimetric assay can be used as a high throughput screen for large numbers of clones.

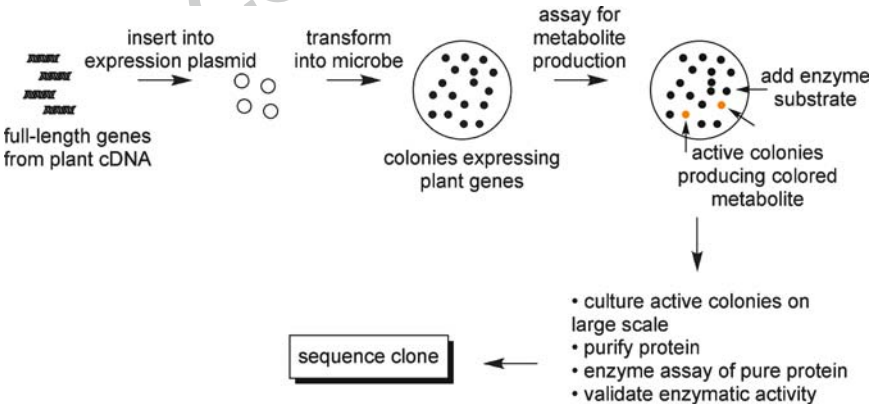

Fig. 8 Expression cloning. Plant cDNA is placed into an expression plasmid and is transformed into yeast or bacteria. Enzyme function can be directly assayed after transformation

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

Scope and Limitations

The greatest limitation of functional cloning is the design of effective enzyme assays that can rapidly assess the function of thousands of clones. Unfortunately, relatively few metabolic enzymes yield a colored product that can be easily detected. Furthermore, many biosynthetic enzymes act on unavailable, unknown or unstable biosynthetic intermediates that are not suitable for an *in vitro* enzyme assay. The plant enzymes must also be expressed in active form in *E. coli* or yeast for accurate functional assessment, and some plant proteins can only be functionally expressed in insect or plant cells, which are not amenable for this functional expression strategy. Finally, expression cloning requires exceptionally high quality cDNA that consists of full-length clones.

Forward Genetics Approach: T DNA Tagging

In the study of microbial natural product biosynthesis, analysis of mutants that express a desired phenotype is a powerful method to assess gene function. Once the desired phenotype – such as a decrease in natural product production – is observed, the corresponding genotype is analyzed. The mutated gene is presumed to be important in modulating the phenotype – i.e. natural product production - of the organism. This strategy is not widely applicable to plants since recessive mutations, in which the phenotype of the mutant is not easily visualized, are usually obtained with most methods of generating mutations. However, activation T-DNA tagging enables the dominant upregulation of genes and has been successfully used to rapidly generate and assess the phenotypes of plant mutants. In this technique, a T-DNA tag that causes transcriptional activation of flanking plant genes is randomly inserted throughout the plant genome. These transformed cells are then screened for changes in phenotype that result from the upregulation (Fig. 9). Provided that a screen or selection to detect changes in natural product production can be developed, this method provides a powerful approach to discovering the genes of secondary metabolic enzymes.

Example

Using an elegant selection approach, the ORCA3 transcriptional factor that regulates the expression of several enzymes of an alkaloid pathway in *C. roseus* was discovered [41, 42]. T-DNA tags were transformed into *C. roseus* cell suspension cultures and the T-DNA was randomly incorporated throughout the genome. The selection assessed overexpression of tryptophan decarboxylase, a known alkaloid biosynthetic enzyme that converts tryptophan into tryptamine. Cells were cultured in the presence of 4-methyltryptophan, which is toxic to *C. roseus*. Cells that grew in the presence of this compound were shown to have upregulated expression levels of tryptophan decarboxylase, which converts 4-methyltryptophan to non-toxic 4-methyltryptamine. Using this selection in combination with T-DNA tagging, the transcriptional factor ORCA3, which regulates expression of tryptophan decarboxylase as well as several other alkaloid biosynthetic enzymes, was discovered.

Scope and Limitations

The limitations that apply to gene silencing described in Section 2.6 also apply with activation T-DNA tagging. First, the plant must be ame-nable to efficient transformation. Additionally, T-DNA tagging results in the generation of many random mutants, and the natural product production levels of these mutants must be assessed by selection or a high throughput screen. Selection for production of natural products – which are not essential for the viability of the producing

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

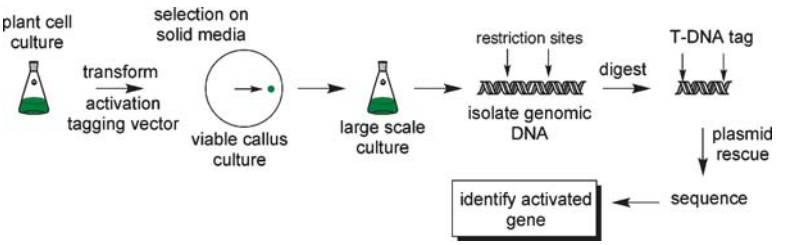

Fig. 9 T-DNA activation tagging. A T-DNA tag is randomly incorporated throughout the plant genome, where any gene adjacent to the tag is upregulated. If a biosynthetic gene is upregulated, then the natural product profile of the plant will be altered. The T-DNA tag can be recovered and the identity of the upregulated gene identified by DNA sequencing

organism – is not straightforward, and clever selection strategies such as the one described for tryptophan decarboxylase must be individually designed for each biosynthetic pathway.

Future Directions

Elucidation of metabolic pathways in plants pose significant challenges. Nevertheless, extraordinary advances have been made in the genetic elucidation of these complex metabolic pathways [6, 43]. As genomic technologies have improved, the genetic basis of the natural products chemistry catalyzed within medicinal plants has slowly – but surely – begun to emerge [44]. This genetic information allows detailed mechanistic explorations of plant pathways, and also enables metabolic engineering efforts that could improve production of medically important plant natural products [45]. This short review highlights a number of methods that have been used to elucidate the genes involved in plant biosynthetic pathways. Selected examples illustrate the successes and limitations encountered with each method. The advent of new technologies in proteomics, chemical biology and other disciplines ensure that novel strategies for elucidation of plant metabolism will continue to emerge. The increasing speed at which plant pathways are being elucidated bodes well for the future of plant derived natural products [46]. By using a variety of approaches in tandem to deconvolute plant metabolism, successes in plant natural products biosynthesis will become increasingly widespread [47].

Acknowledgments I gratefully acknowledge Dr. Xudong Qu for his helpful comments.

References

1. McCurdy, C. R. and Scully, S. S. (2005) Analgesic substances derived from natural products (natureceuticals). *Life Sci*. 78, 476–484.
2. Rainsford, K. D. History and development of the salicylates. (2004) In Aspirin and Related Drugs (K.D. Rainsford, ed.), Taylor & Francis, London/New York, pp. 1–23.
3. Altmann, K.-H. and Gertsch, J. (2007) Anticancer drugs from nature-natural products as a unique source of new microtubule-stabilizing agents. *Nat. Prod. Rep*. 24(2), 327–357.
4. Bode, H. B. and Mueller, R. (2005) The impact of bacterial genomics on natural product research. *Angew. Chem. Intl. Ed*. 44, 6828–6846.
5. Menzella, H. G. and Reeves, C. D. (2007) Combinatorial biosynthesis for drug development. *Curr. Opin. Microbiol*. 10, 238–245.
6. Dixon, R. A. (2005) Engineering of plant natural product pathways. *Curr. Opin. Plant Biol*. 8, 329–336.
7. Katsuyama, Y., Funa, N., Miyahisa, I., and Horinouchi, S. (2007) Synthesis of unnatural flavonoids and stilbenes by exploiting the plant

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

biosynthetic pathway in *Escherichia coli*. *Chem. Biol*. 14, 613–621.
8. Chang, M. C. Y., Eachus, R. A., Trieu, W., Ro, D. K., and Keasling, J. D. (2007) Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. *Nat. Chem. Biol*. 3, 274–277.
9. Allen, R. S., Millgate, A. G., Chitty, J. A., Thisleton, J., Miller, J. A. C., Fist, A. J., Gerlach, W. L., and Larkin, P. J. (2004) RNAi-mediated replacement of morphine with the nonnarcotic alkaloid reticuline in opium poppy. *Nat. Biotechnol*. 22, 1559–1566.
10. Frick, S., Kramell, R., and Kutchan, T. M. (2007) Metabolic engineering with a morphine biosynthetic P450 in opium poppy surpasses breeding. *Met. Eng*. 9, 169–176.
11. Julsing, M. K., Koulman, A., Woerdenbag, H. J., Quax, W. J., and Kayser, O. (2006) Combinatorial biosynthesis of medicinal plant secondary metabolites. *Biomol. Eng*. 23, 265–279.
12. Gantet, P. and Memelink, J. (2002) Transcription factors: tools to engineer the production of pharmacologically active plant metabolites. *Trends Pharm. Sci*. 23, 563–569.
13. Hashimoto, T. and Yamada, Y. (2003) New genes in alkaloid metabolism and transport. *Curr. Opin. Biotech*. 14, 163–168.
14. Qi, X., Bakht, S., Leggett, M., Maxwell, C., Melton, R., and Osbourn, A. (2004) A gene cluster for secondary metabolism in oat: Implications for the evolution of metabolic diversity in plants. *Proc. Natl. Acad. Sci. USA*. 101, 8233–8238.
15. http://data.kew.org/cvalues/
16. Durst, F. and O'Keefe, D. P. (1995) Plant cytochromes P450: an overview. *Drug Metabol Drug Interact*. 12, 171–187.
17. Tholl, D. (2006) Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Curr. Opin. Plant Biol*. 9, 297–304.
18. Gachon, C. M., Langlois-Meurinne, M., and Saindrenan, P. (2005) Plant secondary metabolism glycosyltransferases: the emerging functional analysis. *Trends Plant Sci*. 10, 542–549.
19. Dittrich, H. and Kutchan, T. M. (1991) Molecular cloning, expression, and induction of berberine bridge enzyme, an enzyme essential to the formation of benzophenanthridine alkaloids in the response of plants to pathogenic attack. *Proc. Natl. Acad. Sci. USA*. 88, 9969–9973.
20. Murata, J. and de Luca, V. (2005) Localization of tabersonine 16-hydroxylase and 16-OH tabersonine 16-O-methyl transferase to leaf epidermal cells defines them as a major site of precursor biosynthesis in the vindoline pathway in *Catharanthus roseus*. *Plant J*. 44, 581–594.
21. Levac, D., Murata, J., Kim, W. S., and De Luca, V. (2008) Application of carborundum abrasion for investigating the leaf epidermis: molecular cloning of *Catharanthus roseus* 16-hydroxytabersonine-16-O-methyltransferase. *Plant J*. 53, 225–236.
22. Achnine, L., Huhman, D. V., Farag, M. A., Sumner, L. W., Blount, J. W., and Dixon, R. A. (2005) Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. *Plant J*. 41, 875–887.
23. Walker, K. and Croteau, R. (2000) Taxol biosynthesis: Molecular cloning of a benzoyl-CoA:taxane 2a-O-benzoyltransferase cDNA from *Taxus* and functional expression in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*. 97, 13591–13596.
24. Kutchan, T. M. and Schroder, J. (2002) Selected cell cultures and induction methods for cloning and assaying cytochromes P450 in alkaloid pathways. *Meth. Enzymol*. 357, 370–381.
25. Pauli, H. H. and Kutchan, T. M. (1998) Molecular cloning and functional heterologous expression of two alleles encoding (S)-N- methylcoclaurine 3′-hydroxylase (CYP80B1), a new methyl jasmonate-inducible cytochrome P-450-dependent mono-oxygenase of benzylisoquinoline alkaloid biosynthesis. *Plant J*. 13, 793–801.
26. Chou, W. M. and Kutchan, T. M. (1998) Enzymatic oxidations of alkaloids. *Plant J*. 15, 289–300.
27. Jennewein, S., Wildung, M. R., Chau, M., Walker, K., and Croteau, R. (2004) Random sequencing of an induced *Taxus* cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proc. Natl. Acad. Sci. USA*. 101, 9149–9154.
28. Lange, B. M., Wildung, M. R., Stauber, E. J., Sanchez, C., Pouchnik, D., and Croteau, R. (2000) Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes. *Proc. Natl. Acad. Sci. USA*. 97, 2934–2939.
29. Hirai, M. Y., Tohge, T., and Saito, K. (2006) Systems based analysis of plant metabolism by integration of metabolomics with transcriptomics. *Biotech. Agric. Forestry*. 57, 199–209.
30. Rischer, H., Oresic, M., Seppanen-Laakso, T., Katajamaa, M., Lammertyn, F., Ardiles-Diaz, W.,

BookID 148036_ChapID 07_Proof# 1 -11/4/2009

Van Montagu, M. C. E., Inze, D., Oksman-Caldentey, K. M., and Goossens, A. (2006) Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in *Catharanthus roseus* cells. *Proc. Natl. Acad. Sci. USA*. 103, 5614–5619.

31. Martin, K. J. and Pardee, A. B. (2000) Identifying expressed genes. *Proc. Natl. Acad. Sci. USA*. 97, 3789–3791.
32. Wang, J., Sheehan, M., Brookman, H., and Timko, M. P. (2000) Characterization of cDNAs differentially expressed in roots of tobacco (*Nicotiana tabacum* cv Burley 21) during the early stages of alkaloid biosynthesis. *Plant Sci*. 158, 19–32.
33. Shukla, A. K., Shasany, A. K., Gupta, M. M., and Khanuja, S. P. S. (2006) Transcriptome analysis in *Catharanthus roseus* leaves and roots for comparative terpenoid indole alkaloid profiles. *J. Exp. Bot*. 57, 3921–3932.
34. McGinnis, K., Chandler, V., Cone, K., Kaeppler, H., Kaeppler, S., Kerschen, A., Pikaard, C., Richards, E., Sidorenko, L., Smith, T., Springer, N., and Wulan, T. (2005) Transgene-induced RNA interference as a tool for plant functional genomics. *Meth. Enzymol*. 392, 1–24.
35. Li, R., Reed, D. W., Liu, E., Nowak, J., Pelcher, L. E., Page, J. E., and Covello, P. S. (2006) Functional genomic analysis of alkaloid biosynthesis in *Hyoscyamus niger* reveals a cytochrome P450 involved in littorine rearrangement. *Chem. Biol*. 13, 513–520.
36. Ogita, S., Uefuji, H., Morimoto, M., and Sano, H. (2004) Application of RNAi to confirm theobromine as the major intermediate for caffeine biosynthesis in coffee plants with potential for construction of decaffeinated varieties. *Plant Mol. Biol*. 54, 931–941.
37. Seed, B. (1995) Developments in expression cloning. *Curr. Opin. Biotechnol*. 6, 567–573.
38. Dalboge, H. and Lange, L. (1998) Using molecular techniques to identify new microbial biocatalysts. *Trends Biotechnol*. 16, 265–272.
39. Lange, T., Hedden, P., and Graebe, J. E. (1994) Expression cloning of a gibberellin 20-oxidase, a multifunctional enzyme involved in gibberellin biosynthesis. *Proc. Natl. Acad. Sci. USA*. 91, 8552–8556.
40. Geerlings, A., Redondo, F. J., Memelink, J., Contin, A., van der Heijden, R., and Verpoorte, R. (1999) Screening method for cDNAs encoding putative enzymes converting loganin into secologanin by a transgenic yeast culture. *Biotechnol. Tech*. 13, 605–608.
41. van der Fits, L. and Memelink, J. (2000) ORCA3, a Jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. *Science*. 289, 295–297.
42. van der Fits, L., Hilliou, F., and Memelink, J. (2001) T-DNA activation tagging as a tool to isolate regulators of a metabolic pathway from a genetically non-tractable plant species. *Transgenic Res*. 10, 513–521.
43. Dixon, R. A. (1999) Plant natural products: the molecular genetic basis of biosynthetic diversity. *Curr. Opin. Biotechnol*. 10, 192–197.
44. Memelink, J. (2005) The use of genetics to dissect plant secondary pathways. *Curr. Opin. Plant Biol*. 8, 230–235.
45. Gomez-Galera, S., Pelacho, A. M., Gene, A., Capell, T., and Christou, P. (2007) The genetic manipulation of medicinal and aromatic plants. *Plant Cell Rep*. 26, 1689–1715.
46. Littleton, J. (2007) The future of plant drug discovery. *Expert Opin. Drug Discov*. 2, 673–683.
47. McChesney, J. D., Venkataraman, S. K., and Henri, J. T. (2007) Plant natural products: Back to the future or into extinction. *Phytochemistry*. 68, 2015–2022.

Author Query

Chapter No.: 7 0000950637

Query	Details Required	Author's Response
AU1	Please provide Abstract.	