Biochemistry, Biophysics and Molecular Biology Publications

Biochemistry, Biophysics and Molecular Biology

1-2015

The Application of Synthetic Biology to Elucidation of Plant Mono-, Sesqui-, and Diterpenoid Metabolism

Naoki Kitaoka
Iowa State University

Xuan Lu
Iowa State University

Bing Yang
Iowa State University, byang@iastate.edu

Reuben J. Peters
Iowa State University, rjpeters@iastate.edu

The Application of Synthetic Biology to Elucidation of Plant Mono-, Sesqui-, and Diterpenoid Metabolism

Abstract

Plants synthesize a huge variety of terpenoid natural products, including photosynthetic pigments, signaling molecules and defensive substances. These are often produced as complex mixtures, presumably shaped by selective pressure over evolutionary timescales, some of which have been found to have pharmaceutical and other industrial uses. Elucidation of the relevant biosynthetic pathways can provide increased access (e.g., via molecular breeding or metabolic engineering), and enable reverse genetic approaches towards understanding the physiological role of these natural products in plants as well. While such information can be obtained via a variety of approaches, this review describes the emerging use of synthetic biology to recombinantly reconstitute plant terpenoid biosynthetic pathways in heterologous host organisms as a functional discovery tool, with a particular focus on incorporation of the historically problematic cytochrome P450 mono-oxygenases. Also falling under the synthetic biology rubric and discussed here is the nascent application of genome-editing tools to probe physiological function.

Keywords

metabolic engineering, cytochromes P450, terpene synthase, reverse genetics

Disciplines

Biochemistry | Cell and Developmental Biology | Genetics and Genomics | Molecular Biology | Plant Breeding and Genetics

Comments

The application of synthetic biology to elucidation of plant mono-, sesqui- and di-terpenoid metabolism

Naoki Kitaoka[a,1], **Xuan Lu**[a], **Bing Yang**[b], and **Reuben J. Peters**[a,2]

[a]Department of Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames, IA 50011, U.S.A

[b]Department of Genetics, Developmental & Cell Biology, Iowa State University, Ames, IA 50011, U.S.A

Abstract

Plants synthesize a huge variety of terpenoid natural products, including photosynthetic pigments, signaling molecules and defensive substances. These are often produced as complex mixtures, presumably shaped by selective pressure over evolutionary timescales, some of which have been found to have pharmaceutical and other industrial uses. Elucidation of the relevant biosynthetic pathways can provide increased access (e.g., via molecular breeding or metabolic engineering), and enable reverse genetic approaches towards understanding the physiological role of these natural products in plants as well. While such information can be obtained via a variety of approaches, this review describes the emerging use of synthetic biology to recombinantly reconstitute plant terpenoid biosynthetic pathways in heterologous host organisms as a functional discovery tool, with a particular focus on incorporation of the historically problematic cytochrome P450 mono-oxygenases. Also falling under the synthetic biology rubric and discussed here is the nascent application of genome-editing tools to probe physiological function.

Keywords

metabolic engineering; cytochromes P450; terpene synthase; reverse genetics

Introduction

Plants are a prolific source of natural products, with their complex chemical mixtures reflecting adaptation to previous and current selective pressures over evolutionary timescales. From origins in central metabolism, often via initial derivation from hormone biosynthetic pathways, natural products have diversified into the astounding array of extant compounds, only a fraction of which are currently known. The exact physiological roles played by these small molecules are then presumably similarly diverse. However, our

[2]To whom correspondence should be addressed: rjpeters@iastate.edu, tel. 1-515-294-8580, fax 1-515-294-0453.
[1]Current address: Tohoku University, Sendai 980-8577, JAPAN

understanding of physiological function lags well behind even our efforts to catalog chemical diversity, almost invariably relying on the biological activity exhibited in laboratory tests by isolated compounds. Nevertheless, plant natural products have been utilized for millennia as pharmaceuticals, flavors and fragrances, along with other industrial uses.

Not surprisingly, investigation of plant natural products biosynthesis is largely focused on those compounds with economic significance (i.e., industrial use) or of obvious physiological importance (e.g., hormones). These studies have progressed from biochemical assays with cell-free extracts and isolation of the relevant enzymes to molecular biology based recombinant expression and analysis of plants genetically modified to over or under express the encoding genes. However, outside of hormones, the biosynthesis of very few plant natural products has been fully elucidated.

Identification of the relevant biosynthetic pathway enables access to the resulting natural product, via metabolic engineering in the native plant or in microbial hosts. The use of microbes as production platforms provided one the early examples of the use of the term synthetic biology due in large part to the substantial efforts required to optimize production levels. Such application of synthetic biology has been extensively reviewed elsewhere (Keasling, 2012; Keasling et al., 2012). Here is reviewed the use of this synthetic biology approach to investigate plant metabolism (i.e., pathways/ networks), as well as the use of synthetic biology-enabled genome-editing tools to investigate physiological function.

Terpenoid metabolism

More specifically, this review focuses on terpenoid metabolism, which has been extensively diversified in plants, to the point that these represent the largest class of natural products with over 45,000 known examples (Renault et al., 2014). Terpenoids are derived from 5-carbon building blocks, namely isopentenyl diphosphate (IPP) and dimethyl allyl diphosphate (DMAPP). These building blocks are produced by two distinct pathways, the mevalonate (MVA) dependent pathway that operates in the cytosol of plants, and the 2-C-methyl-D-erythritol-4-phosphate (MEP) dependent pathway found in the plastids (Vranova et al., 2013).

DMAPP and IPP are coupled together in a head-to-tail fashion to form the precursor to the 10-carbon monoterpenoids, in turn this is coupled to another molecule of IPP to form the precursor to the 15-carbon sesquiterpenoids, with coupling of yet another molecule of IPP yielding the precursor to the 20-carbon diterpenoids. These precursors take the form of allylic diphosphate containing acyclic chains, generally with internal double bonds of *trans* configuration, although the use of analogous *cisoid* precursors also has been reported (Bohlmann and Gershenzon, 2009). Formation of these elongated precursors is catalyzed by isoprenyl diphosphate synthases. Two of the *transoid* C_{15} farnesyl diphosphate can be condensed together in a head-to-head manner to form the squalene precursor to the triterpenoids, and two of the C_{20} geranylgeranyl diphosphate (GGPP) similarly condensed to form the phytoene precursor to the tetraterpenoids, although neither of these types of

terpenoids will be further discussed here as at least the triterpenoids have been recently reviewed (Moses et al., 2013).

The allylic diphosphate ester bond of the acyclic prenyl diphosphate precursors for mono-, sesqui- and di-terpenoids is lysed, with the resulting carbocation used to drive cyclization and/or rearrangement of the hydrocarbon chain in reactions catalyzed by (class I) terpene synthases (TPSs). In the case of diterpenoids, this can be preceded by an initial (bi)cyclization of the general precursor GGPP, typically forming copalyl diphosphate (CPP), in reactions catalyzed by CPP synthases (CPSs). In the course of the catalyzed reaction TPSs and CPSs can add water to produce oxygenated terpenes. However, the resulting compounds are most often olefins (Gao et al., 2012).

Given the hydrophobic nature of the hydrocarbon terpene products of TPSs, the addition of oxygen is almost invariably catalyzed by endoplasmic reticulum membrane-associated (microsomal) cytochromes P450 (CYPs). These heme-thiolate mono-oxygenases often catalyze the insertion of oxygen into carbon-hydrogen bonds to produce hydroxyl groups, although they can mediate more complex reactions as well (Guengerich and Munro, 2013). In any case, the addition of oxygen both increases polarity/solubility and imparts hydrogen-bonding capacity, enabling specific binding interactions, as well as providing functional groups for further modification.

The extent of terpenoid natural product diversity in plants perhaps can be best appreciated by noting the number of TPSs found in the known genomes, with all spermatophytes containing at least 40 such genes, with over 100 observed in some species (Chen et al., 2011). Although all of genes these may not encode functional TPSs, and there is some overlap in catalyzed product outcome (and/or outright genetic redundancy), this is balanced by the propensity for these enzymes to yield a range of products. In addition, it has been noted that the CYPs form the largest family of plant metabolic enzymes, with over 200 found in all spermatophytes (Nelson and Werck-Reichhart, 2011), and many of these CYPs participate in terpenoid metabolism (Hamberger and Bak, 2013). Thus, even without consideration of the other classes of enzymes that participate in terpenoid biosynthesis, it can be appreciated that terpenoid metabolism forms a complex network in any given plant! However, very few pathways are known, and there is no example of a completely elucidated network, even at just the level of the TPSs – e.g., not even for the model plant *Arabidopsis thaliana*. Accordingly, it also has not been possible to fully determine the full range of physiological roles played by the various terpenoid natural products made by any one species.

Our understanding of plant terpenoid metabolism has been limited by a number of factors. These include the perceived recalcitrance of eukaryotic enzymes to functional recombinant expression in the usual hosts – particularly for membrane-associated CYPs – as well as the difficulty of identifying and obtaining the correct intermediate/substrate for the enzyme in question. In addition, the typical lack of functional gene clustering in plants dictates a general need to laboriously individually identify each enzyme, unlike the situation in microbes where such clustering often translates to identification of any relevant enzyme leading to the complete biosynthetic pathway, which is mediated by the enzymes encoded by

the neighboring genes. The difficulty in identification of plant biosynthetic enzymes is increased by the relative lack of plant sequence information – e.g., high-quality genomes and/or transcriptomes – although the amount of such information is currently rapidly increasing.

Elucidation of menthol biosynthesis: A monoterpenoid case study

With regards to pathway elucidation it is perhaps most instructive to examine the case of the monoterpene menthol, whose biosynthesis was worked out over the course of the stellar career of Prof. Rodney Croteau at Washington State University (Figure 1). Enabled by the localization of methol production to protruding glandular trichomes that could be physically isolated (Gershenzon et al., 1992), this work began with enzymatic assays with plant cell-free extracts, progressed to molecular genetic identification of the relevant biosynthetic enzymes (for both on- and off-pathway reactions), leading to genetic/metabolic engineering in peppermint (*Mentha* x *piperita*) plants (Croteau et al., 2005). The full elucidation of the relevant pathway and (relatively few) branch points enabled a systems biology approach to understanding of biochemical regulation (Rios-Estepa et al., 2008), and comprehensive mathematical modeling guided evaluation of the pathway (Rios-Estepa et al., 2010). Notably, in part based on these advances, mint plants have been engineered for higher and more reliable yield of menthol, as demonstrated in multi-year field trials (Lange et al., 2011).

As part of this work, the Croteau group was able to easily recombinantly express the relevant mint TPS, 4*S*-limonene synthase, in *Escherichia coli* (Colby et al., 1993), with particularly strong activity observed upon truncation of the N-terminal plastid targeting sequence (Williams et al., 1998). However, functional recombinant expression for identification of the initial acting peppermint CYP, limonene-3-hydroxylase (CYP71D13), was carried out using a cumbersome insect cell culture system (Lupien et al., 1999), in large part due to the presumed similarity in membrane composition and presence of the requisite CYP reductase (CPR), which is not present in bacteria. Later work then demonstrated that it was possible to also functionally express N-terminally modified versions of CYP71D13 in *Saccharomyces cerevisiae* (yeast) or *E. coli*, with activity observed upon reconstitution with a plant CPR (Haudenschild et al., 2000). Identification of these enzymes offered the possibility for pathway reconstitution in a heterologous host (micro)organism. Such metabolic engineering was attempted in *E. coli*, with some production of limonene observed, but flux through even the first CYP step was not observed (Carter et al., 2003). Notably, feeding studies indicated that this lack of flux was due to insufficient production of the upstream terpenoids (and possibly also the volatile nature of limonene), rather than lack of CYP activity per se. Nevertheless, functional expression of eukaryotic microsomal CYPs in *E. coli* is traditionally considered problematic, and most studies use yeast if not insect cells or, in the case of plant CYPs, plant expression systems.

Increasing flux: Production of artemisinic acid

Even as this work was taking place, the issue of flux towards terpenoid metabolism was being addressed by the group of Prof. Jay Keasling at the Univ. of California at Berkeley.

Although the complete biosynthetic process underlying production of the important antimalarial sesquiterpene artemisinin was and still remains to this day unknown (Brown, 2010), the Keasling group had received substantial funding to metabolically engineer microbial production of the intermediate artemisinic acid, from which artemisinin can be efficiently made (Paddon and Keasling, 2014). Focusing on production of the olefin intermediate, amorphadiene, for which the relevant TPS had already been identified, the Keasling group imported the entire MEV pathway from *S. cerevisiae* into *E. coli* (Figure 2), dramatically improving terpenoid production levels (Martin et al., 2003). Notably, this work also featured one of the first examples of the use of a synthetic gene, recoded to optimize codon usage for expression in *E. coli*. Later reported work led to even more dramatic increases in yield, with production levels reported to be over 27 g/L of amorphadiene, and application of the term synthetic biology to this work (Keasling, 2012). The development of such highly engineered microbial production systems obviously provides sufficient substrate for subsequently acting enzymes, either via isolation and re-feeding or additional incorporation of the enzyme into the metabolically engineered microorganism.

Of particular interest here, the latter approach – extended metabolic engineering, albeit carried out in yeast rather than *E. coli* – was used for functional identification of the CYP71AV1 relevant to artemisinin biosynthesis, which is capable of catalyzing the multiple reactions necessary to convert amorphadiene to artemisinic acid (Ro et al., 2006). Indeed, yeast was chosen as the microbial host platform for further development by Amyris, the company founded by Keasling and co-workers to commercialize such engineering of high-level terpene production. In addition to extensive modifications of the endogenous metabolic network of yeast, it is of interest here to note that commercial production of artemisinic acid (reported to be 25 g/L) also was dependent on incorporation of the additional alcohol and aldehyde dehydrogenases involved in the plant biosynthetic pathway (Paddon et al., 2013), which were identified from work carried out in the group of Dr. Pat Covello at the National Research Council of Canada, Saskatoon (Teoh et al., 2009; Teoh et al., 2006).

Engineering yeast to elucidate CYP roles in terpenoid metabolism

The use of metabolic engineering of yeast for functional characterization of CYPs involved in plant terpenoid biosynthesis was pioneered by Dr. Dae-Kyun Ro. First, working with Prof. Joerg Bohlmann at the Univ. of British Columbia, where this approach was used to at least confirm the functional identification of CYP724B1 as a promiscuous diterpene oxidase involved in diterpene resin acid biosynthesis in loblolly pine, which is capable of catalyzing the multiple reactions necessary to convert a methyl to carboxylic acid (Ro et al., 2005). He then went on to the Keasling group, and was a key player in their identification of amorphadiene oxidase (Ro et al., 2006). Now a professor at the Univ. of Calgary, his group has used this approach to functionally identify the CYPs involved in sesquiterpene lactone biosynthesis from plant species in the Asteraceae family, again multiply-reactive, methyl to carboxylic acid converting oxidases that also are members of the CYP71AV sub-family (Nguyen et al., 2010), and subsequently acting hydroxylases from the CYP71BL sub-family, with presumably spontaneous formation of the lactone ring occurring upon hydroxylation at appropriately spaced locations (Ikezawa et al., 2011). Prof. Bohlmann's group also has applied this approach to functional identification of the spruce CYP720B4 involved in

diterpene resin acid biosynthesis (Hamberger et al., 2011), and the sandalwood CYP71F sub-family members involved in production of the essential oil sesquiterpene components santalols and bergamotol (Diaz-Chavez et al., 2013).

Emerging alternative expression systems

Part of the rationale for the use of yeast in such work is the ability to functionally express unmodified plant microsomal CYPs. Another system attracting use for this reason is transient transformation in tobacco, *Nicotiana benthamiana* (Bach et al., 2014a). In addition, a particularly promising system for investigation of terpenoid biosynthesis is the moss *Physcomitrella patens* (Bach et al., 2014b). However, there do not appear to be any CYPs from terpenoid biosynthesis identified using these systems as of yet. Moreover, at least with the tobacco system production of terpenoids can be plagued by interference from endogenous metabolism – e.g., the introduction of artemisinic acid biosynthesis led to glycosylated compound (van Herpen et al., 2010) – which presumably would complicate analysis of novel enzymatic activity.

Data deluge: Increasing availability of plant sequence information

The use of metabolic engineering to investigate biosynthetic pathways is of particular interest when genes encoding enzymes of potential relevance are readily available. Such information is being provided by next-generation sequencing based approaches, such as the RNA-Seq methodology that has been applied over the last few years by several consortia to medicinal plants (Giddings et al., 2011; Gongora-Castillo et al., 2012; Marques et al., 2013; Xiao et al., 2013; Zerbe et al., 2013), with explicit incorporation of the use of metabolic engineering to at least functionally characterize the encoded enzymes by the PhytoMedSyn group in Canada (Facchini et al., 2012).

Even beyond these medicinal plant transcriptomic sequence based investigations, which were largely targeted at natural products of pharmaceutical interest, with the availability of plant genome sequences we now have essentially complete gene inventories that can be used to probe the encoded natural product metabolic networks. Of particular interest here, rice (*Oryza sativa*) provides a long-standing example that couples a high-quality genome with extensive cDNA sequence information (International Rice Genome Sequencing, 2005; Kikuchi et al., 2003).

Rice as a model system for investigating diterpenoid metabolism

Intrigued by the presence of several families of diterpene phytoalexins in rice, we have undertaken a functional genomics approach towards investigating the relevant biosynthetic network. Although a diterpene phytoalexin derived from casbene, which is formed by direct TPS catalyzed cyclization of GGPP, has been recently reported (Inoue et al., 2013), the remainder require initial cyclization by a CPS, with subsequent cyclization and/or rearrangement catalyzed by TPSs. These fall into the TPS-e sub-family related to the *ent*-kaurene synthase (KS) required for gibberellin hormone biosynthesis (Chen et al., 2011),

which we refer to as KS-like (KSLs). The resulting phytoalexins then fall into the labdane-related diterpenoid (LRD) super-family (Peters, 2010).

Since only the LRD phytoalexins were originally known, we focused on biochemical characterization of the corresponding rice diterpene synthases – i.e., the OsCPSs and OsKSLs. Our work was carried out at the same time as similar studies by groups led by Profs. Hisakazu Yamane at the Univ. of Tokyo and Tomonobu Toyomasu at Yamagata Univ., all of which relied on biochemical in vitro assays and the use of authentic standards for the expected diterpene olefin intermediates, almost all of which were provided by Prof. Robert M. Coates from the Univ. of Illinois. The results of these studies have been extensively reviewed elsewhere (Peters, 2006; Schmelz et al., 2014; Toyomasu, 2008; Yamane, 2013), and here we simply note that all the functional OsCPSs and OsKSLs have been characterized, with rice found to encode a diterpene metabolic network even more complex than originally expected (Figure 3).

The complex nature of rice diterpene metabolism was evident to some degree simply from the number of OsKLS, which exceeded that required for the known families of phytoalexins – each of which is defined on the basis of their derivation from a common hydrocarbon backbone formed by the combined activity of a CPS and KSL. Specifically, seven functional OsKS(L) are found in the rice genome, although only five were expected (this includes the OsKS required for gibberellin biosynthesis; see Figure 3). In our own work, we found an OsKSL whose product did not match any of the olefin intermediates for the known LRDs, forcing an almost year-long search for corresponding authentic standard to verify its production of *syn*-stemodene (Morrone et al., 2006).

Development of a modular metabolic engineering system in *E. coli*

Both daunted by this experience and inspired by the complexity of rice diterpene biosynthesis, we developed a modular metabolic engineering system to co-express any pairing of a CPS and KS(L), along with the necessary upstream GGPP synthase in *E. coli* (Cyr et al., 2007). In order to produce enough of the resulting diterpene(s) for de novo structural analysis (e.g., by NMR) we further extended this system with either increased flux to terpenoids via either overexpression of key enzymes from the endogenous MEP pathway, or importation of the yeast MEV pathway using vectors kindly shared with us by Prof. Keasling (Morrone et al., 2010b). This system has proven utility for functional characterization of novel diterpene synthases from various plants (Gao et al., 2009; Jackson et al., 2014; Wu et al., 2012; Zhou et al., 2012), as well as microbes (Hershey et al., 2014; Lu et al., submitted; Morrone et al., 2009; Xu et al., submitted), along with mutational analysis (Criswell et al., 2012; Morrone et al., 2008; Potter et al., 2014).

Even beyond such de novo characterization of functionally novel diterpene synthases, this plug-and-play approach enables examination of enzymatic substrate specificity, which revealed further complexity/range for rice diterpene metabolism. One of the OsKSLs uniquely reacts with both stereoisomers of CPP (*ent* and *syn*) produced by rice, with both resulting products found in planta. Moreover, two other OsKSLs that act on *syn*-CPP in rice also will react with CPP of normal stereochemistry (Morrone et al., 2011). This promiscuity

may be related to the evolution of alternative stereoisomers of CPP (i.e., other than the *ent*-CPP required for gibberellin biosynthesis) in the cereals, as the wheat ortholog of the rice *syn*-CPP producing CPS produces normal CPP instead (Wu et al., 2012), and several wheat KSLs similarly react with both the endogenous (normal) CPP as well as *syn*-CPP, which does not seem to be produced by wheat (Zhou et al., 2012). Accordingly, the ability to easily interrogate diterpene synthase activity via this modular metabolic engineering system offers insights not only into existing metabolism, but also (through examination of related plants) the evolution of diterpene biosynthesis as well.

Investigation of rice CYPs in *E. coli*: Enabled by gene synthesis

Fortuitously, in the course of characterizing the rice CPSs and KSLs, we noted that some of the genes encoding sequentially acting CPS and KSL were close together in the rice genome, one pair on chromosome 4 and others grouped together on chromosome 2, which seemed to define diterpenoid biosynthesis gene clusters (Prisic et al., 2004; Wilderman et al., 2004). Indeed, in a report on rice gibberellin metabolism, these regions had been previously noted to also contain genes encoding a number of CYPs, as well as short-chain alcohol dehydrogenases (Sakamoto et al., 2004). Many of these genes were found to be co-regulated, specifically inducible by the fungal cell wall component chitin, as well (Okada et al., 2007). Subsequently, genetic (RNAi) evidence was reported indicating a role for the CYPs from the chromosome 4 cluster in momilactone diterpenoid metabolism (Shimura et al., 2007), consistent with the role of the co-clustered OsCPS4 and OsKSL4 [Figure 3 (Wilderman et al., 2004)]. This information was particularly useful in enabling focused investigation of CYPs – i.e., rather than investigating all >350 CYP found in the rice genome, we then simply targeted those from the diterpenoid biosynthesis clusters, along with the few others that exhibited inducible transcription (15 total).

Based on our ability to functionally express the CYP kaurene oxidase required for gibberellin biosynthesis from *A. thaliana* (CYP701A3) in *E. coli* (Morrone et al., 2010a), we attempted to incorporate rice CYPs into our metabolic engineering system. In an initial study, we were able to do so using an N-terminally modified version of the native gene for one of these, CYP76M7, which acts on the *ent*-cassadiene product of the co-clustered OsCPS2 and OsKSL7, carrying out C11α-hydroxylation that is consistent with a role in phytocassane biosynthesis (Swaminathan et al., 2009). However, no activity was observed for the other CYPs with this approach. Knowing that at least one of the two closely related CYPs found in the cluster on chromosome 4 (i.e., CYP99A2 and 3) played a role in momilactone biosynthesis (Shimura et al., 2007), we turned to use of synthetic genes optimized for expression in *E.* coli, along with N-terminal modification. This led to functional expression and successful incorporation into our metabolic engineering system of at least one of these, CYP99A3, which acts on the *syn*-pimaradiene product of the co-clustered OsCPS4 and OsKSL4, carrying out conversion of C19 from a methyl to carboxylic acid (Wang et al., 2011). Such use of synthetic gene constructs has since led to functional characterization of the ability of a number of rice CYPS to act on the diterpene olefin products of the OsCPSs and OsKSLs (Wang et al., 2012a; Wang et al., 2012b; Wu et al., 2011), with the further optimization of this synthetic biology approach reported in the

accompanying paper enabling functional expression of even previously recalcitrant CYPs (Kitaoka et al., submitted).

Based on the observed activity with olefins, it is clear that several CYPs will react with the same diterpene olefin intermediate (Figure 4). Thus, it is clear that some CYPs will act downstream (i.e., sequentially). This can be verified in part by incorporation of multiple CYPs into the metabolic engineering system, which we have demonstrated for at least two consecutively acting CYPs from rice oryzalexin metabolism (Figure 4C). However, such analysis leaves the order of reaction undefined. Defining pathway then requires isolation and refeeding experiments, which we have used not only in rice (Wu et al., 2013), but also in our studies of diterpenoid metabolism from other plants as well (Zi and Peters, 2013). Notably, use of this synthetic biology approach also enables facile access to such putative biosynthetic intermediates as well as functional recombinant expression of multiple plant CYPs.

Other groups have independently demonstrated the ability to functionally incorporate plant microsomal CYPs for the purpose of reconstituting terpenoid biosynthetic pathways in *E. coli*, most notably for the production of early intermediates from the biosynthetic pathway leading to the anti-cancer drug taxol (Ajikumar et al., 2010). Although this study was not directed at identification of novel CYPs, instead building on previous work from the Croteau group (Jennewein et al., 2004), it emphasizes the continuing exploration of synthetic biology approaches in *E. coli* for production of high-value terpenoid natural products.

Investigation of physiological function via synthetic biology

Given the paucity of information, investigation of the physiological role played by most natural products seems well worth pursuit. With molecular identification of biosynthetic pathways one approach might involve transplantation into heterologous host plants, with subsequent investigation of physiological effect (i.e., phenotype). Such an approach might be useful for investigation of evolution as well. For example, the change in CPP stereochemistry (and that of the derived diterpenoid natural products) between rice and wheat might be probed by expression of the CPS producing the alternative stereoisomer and subsequent analysis of the array of resulting diterpenoids (to probe metabolic promiscuity) as well as physiological effect.

Particularly given that the biosynthesis of natural products is often embedded in more complex metabolic networks, definitive investigations rely on reverse genetic analysis. For example, this can be used to verify enzymatic roles in biosynthetic pathways assigned on the basis of biochemical activity observed with recombinantly expressed genes. Arguably more importantly, this further enables investigation of the physiological function of the resulting natural products, which is particularly critical given the complex mixtures of compounds generally produced by plants that confound functions ascribed solely on the basis of in vitro analyses of natural product biological activity. For example, despite being the first rice diterpenoids suggested to act as phytoalexins against the devastating fungal blast pathogen *Magneportha oryzae* (Cartwright et al., 1977), the momilactones seem to be more important as allelochemicals that are constitutively secreted from the roots and suppress the growth of

other plant species (Xu et al., 2012). This is consistent with the original isolation of the momilactones as plant growth inhibitors (Kato et al., 1973), and more recent work on rice allelopathy (Kato-Noguchi and Peters, 2013), although it should be noted that some evidence has been presented indicating a role for the momilactones (or related OsCPS4-dependent diterpenoids) as phytoalexins (Toyomasu et al., 2013).

Although RNAi can be used to knock-down expression of the relevant genes, this approach often leads to incomplete suppression of natural product biosynthesis and can be somewhat non-selective (i.e., affecting more than the targeted gene), confounding interpretation of studies with the resulting plants. However, identification of gene knockout mutants has relied on laborious screening of large-scale mutagenesis projects (generally insertional), and only for Arabidopsis have such efforts led to a (almost) complete set of mutants. Even in rice there are many genes for which no mutant is yet available, and those that are available are found in a variety of different genetic backgrounds, complicating comparative analysis. Accordingly, the recent development of targetable genome-editing tools, particularly the modular nature of the recently developed transcription activator-like nucleases (TALENs) and clustered regularly-interspaced short palindromic repeats (CRISPR)/Cas9 based technologies (Christian et al., 2010; Jinek et al., 2012; Li et al., 2011), enables interrogation of physiological functional in native biological context. Notably, judicious construction of a construct targeting any given gene (or multiple genes or even gene cluster) allows both construction of all related mutants in same genetic background (e.g., cultivar), and identical mutants in different backgrounds (depending on conservation of the targeted sequence, perhaps even in closely related species), enabling relatively sophisticated analyses. Such genome editing has been successfully carried out with regeneration of whole plants that stably carry the modified genes over multiple generations (Feng et al., 2013; Miao et al., 2013; Zhang et al., 2014; Zhou et al., 2014), and we anticipate that such an approach will be applied to investigation of the physiological roles of plant natural products in the near future.

CONCLUSIONS

Traditional approaches to plant natural products generally involved isolation on the basis of targeted biological activity with subsequent investigation of the underlying biosynthetic pathway via purification of the relevant enzymes from the native producing plant. With the rapidly increasing availability of plant sequence information, both genomic and transcriptomic, it is possible to more systematically study both biosynthesis and physiological function. Here we have reviewed the application of synthetic biology approaches to such investigation of plant terpenoid natural products. Key developments for study of plant terpenoid biosynthesis are derived from the work of Prof. Keasling on production of artemisinin via semi-synthesis from artemisinic acid, demonstrating the use of synthetic biology for commercial purposes. As reviewed here, this has further led to the use of synthetic biology as a functional discovery tool for elucidation of plant terpenoid metabolism more generally, including investigation of biosynthetic networks and their evolution. In addition, we suggest the potential for application of synthetic biology-enabled genome-editing tools to systematic reverse genetic studies of physiological function.

Acknowledgments

References

Ajikumar PK, Xiao WH, Tyo KE, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G. Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. Science. 2010; 330:70–74. [PubMed: 20929806]

Bach SS, Bassard JE, Andersen-Ranberg J, Moldrup ME, Simonsen HT, Hamberger B. High-throughput testing of terpenoid biosynthesis candidate genes using transient expression in Nicotiana benthamiana. Methods Mol Biol. 2014a; 1153:245–255. [PubMed: 24777803]

Bach SS, King BC, Zhan X, Simonsen HT, Hamberger B. Heterologous stable expression of terpenoid biosynthetic genes using the moss Physcomitrella patens. Methods Mol Biol. 2014b; 1153:257–271. [PubMed: 24777804]

Bohlmann J, Gershenzon J. Old substrates for new enzymes of terpenoid biosynthesis. Proc Natl Acad Sci U S A. 2009; 106:10402–10403. [PubMed: 19553206]

Brown GD. The biosynthesis of artemisinin (Qinghaosu) and the phytochemistry of Artemisia annua L. (Qinghao). Molecules. 2010; 15:7603–7698. [PubMed: 21030913]

Carter OA, Peters RJ, Croteau R. Monoterpene biosynthesis pathway construction in *Escherichia coli*. Phytochemistry. 2003; 64:425–433. [PubMed: 12943759]

Cartwright DW, Langcake P, Pryce RJ, Leworthy DP, Ride JP. Chemical activation of host defence mechanisms as a basis for crop protection. Nature. 1977; 267:511–513.

Chen F, Tholl D, Bohlmann J, Pichersky E. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. Plant J. 2011; 66:212–229. [PubMed: 21443633]

Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ, Voytas DF. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics. 2010; 186:757–761. [PubMed: 20660643]

Colby SM, Alonso WR, Katahira EJ, McGarvey DJ, Croteau R. 4S-limonene synthase from the oil glands of spearmint (Mentha spicata). cDNA isolation, characterization, and bacterial expression of the catalytically active monoterpene cyclase. J Biol Chem. 1993; 268:23016–23024. [PubMed: 8226816]

Criswell J, Potter K, Shephard F, Beale MB, Peters RJ. A single residue change leads to a hydroxylated product from the class II diterpene cyclization catalyzed by abietadiene synthase. Org Lett. 2012; 14:5828–5831. [PubMed: 23167845]

Croteau RB, Davis EM, Ringer KL, Wildung MR. (−)-Menthol biosynthesis and molecular genetics. Die Naturwissenschaften. 2005; 92:562–577. [PubMed: 16292524]

Cyr A, Wilderman PR, Determan M, Peters RJ. A Modular Approach for Facile Biosynthesis of Labdane-Related Diterpenes. J Am Chem Soc. 2007; 129:6684–6685. [PubMed: 17480080]

Diaz-Chavez ML, Moniodis J, Madilao LL, Jancsik S, Keeling CI, Barbour EL, Ghisalberti EL, Plummer JA, Jones CG, Bohlmann J. Biosynthesis of Sandalwood Oil: Santalum album CYP76F cytochromes P450 produce santalols and bergamotol. PLoS One. 2013; 8:e75053. [PubMed: 24324844]

Facchini PJ, Bohlmann J, Covello PS, De Luca V, Mahadevan R, Page JE, Ro DK, Sensen CW, Storms R, Martin VJ. Synthetic biosystems for the production of high-value plant metabolites. Trends in biotechnology. 2012; 30:127–131. [PubMed: 22209518]

Feng Z, Zhang B, Ding W, Liu X, Yang DL, Wei P, Cao F, Zhu S, Zhang F, Mao Y, et al. Efficient genome editing in plants using a CRISPR/Cas system. Cell research. 2013; 23:1229–1232. [PubMed: 23958582]

Gao W, Hillwig ML, Huang L, Cui G, Wang X, Kong J, Yang B, Peters RJ. A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. Org Lett. 2009; 11:5170–5173. [PubMed: 19905026]

Gao Y, Honzatko RB, Peters RJ. Terpenoid synthase structures: a so far incomplete view of complex catalysis. Nat Prod Rep. 2012; 29:1153–1175. [PubMed: 22907771]

Gershenzon J, McCaskill D, Rajaonarivory JIM, Mihaliak C, Karp R, Croteau R. Isolation of secretory cells from plant glandular trichomes and their use in biosynthetic studies of monoterpenes and other glanular products. Anal Biochem. 1992; 200:130–138. [PubMed: 1595887]

Giddings LA, Liscombe DK, Hamilton JP, Childs KL, DellaPenna D, Buell CR, O'Connor SE. A stereoselective hydroxylation step of alkaloid biosynthesis by a unique cytochrome P450 in Catharanthus roseus. J Biol Chem. 2011; 286:16751–16757. [PubMed: 21454651]

Gongora-Castillo E, Fedewa G, Yeo Y, Chappell J, DellaPenna D, Buell CR. Genomic approaches for interrogating the biochemistry of medicinal plant species. Methods in enzymology. 2012; 517:139–159. [PubMed: 23084937]

Guengerich FP, Munro AW. Unusual cytochrome p450 enzymes and reactions. J Biol Chem. 2013; 288:17065–17073. [PubMed: 23632016]

Hamberger B, Bak S. Plant P450s as versatile drivers for evolution of species-specific chemical diversity. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2013; 368:20120426. [PubMed: 23297350]

Hamberger B, Ohnishi T, Seguin A, Bohlmann J. Evolution of diterpene metabolism: Sitka spruce CYP720B4 catalyzes multiple oxidations in resin acid biosynthesis of conifer defense against insects. Plant Physiology. 2011; 157:1677–1695. [PubMed: 21994349]

Haudenschild C, Schalk M, Karp R, Croteau R. Functional expression of regiospecific cytochrome P450 limonene hydroxylases from mint (*Mentha* spp.) in *Escherichia coli* and *Saccharomyces cervisiae*. Arch Biochem Biophys. 2000; 279:127–136.

Hershey DM, Lu X, Zi J, Peters RJ. Functional conservation of the capacity for *ent*-kaurene biosynthesis and an associated operon in certain rhizobia. J Bact. 2014; 196:100–106. [PubMed: 24142247]

Ikezawa N, Gopfert JC, Nguyen DT, Kim SU, O'Maille PE, Spring O, Ro DK. Lettuce costunolide synthase (CYP71BL2) and its homolog (CYP71BL1) from sunflower catalyze distinct regio- and stereoselective hydroxylations in sesquiterpene lactone metabolism. J Biol Chem. 2011; 286:21601–21611. [PubMed: 21515683]

Inoue Y, Sakai M, Yao Q, Tanimoto Y, Toshima H, Hasegawa M. Identification of a novel casbane-type diterpene phytoalexin, ent-10-oxodepressin, from rice leaves. Biosci Biotechnol Biochem. 2013; 77:760–765. [PubMed: 23563545]

International Rice Genome Sequencing Project. The map-based sequence of the rice genome. Nature. 2005; 436:793–800. [PubMed: 16100779]

Jackson AJ, Hershey DM, Chesnut T, Xu M, Peters RJ. Biochemical characterization of the castor bean *ent*-kaurene synthase(-like) family supports quantum chemical view of diterpene cyclization. Phytochemistry. 2014 in press.

Jennewein S, Long RM, Williams RM, Croteau RB. Cytochrome p450 taxadiene 5alpha-hydroxylase, a mechanistically unusual monooxygenase catalyzing the first oxygenation step of taxol biosynthesis. Chem Biol. 2004; 11:379–387. [PubMed: 15123267]

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337:816–821. [PubMed: 22745249]

Kato T, Kabuto C, Sasaki N, Tsunagawa M, Aizawa H, Fujita K, Kato Y, Kitahara Y, Takahashi N. Momilactones, growth inhibitors from rice, *Oryza sativa* L. Tetrahedron Lett. 1973; 14:3861–3864.

Kato-Noguchi H, Peters RJ. The role of momilactones in rice allelopathy. J Chem Ecol. 2013; 39:175–185. [PubMed: 23385366]

Keasling JD. Synthetic biology and the development of tools for metabolic engineering. Metabolic engineering. 2012; 14:189–195. [PubMed: 22314049]

Keasling JD, Mendoza A, Baran PS. Synthesis: A constructive debate. Nature. 2012; 492:188–189. [PubMed: 23235869]

Kikuchi S, Satoh K, Nagata T, Kawagashira N, Doi K, Kishimoto N, Yazaki J, Ishikawa M, Yamada H, Ooka H, et al. Collection, mapping, and annotation of over 28,000 cDNA clones from *japonica* rice. Science. 2003; 301:376–379. [PubMed: 12869764]

Kitaoka N, Wu Y, Xu M, Peters RJ. Optimization of a synthetic biology approach enables discovery of novel cytochrome P450 activity in rice diterpenoid biosynthesis. Mol Plant. submitted.

Lange BM, Mahmoud SS, Wildung MR, Turner GW, Davis EM, Lange I, Baker RC, Boydston RA, Croteau RB. Improving peppermint essential oil yield and composition by metabolic engineering. Proc Natl Acad Sci U S A. 2011; 108:16944–16949. [PubMed: 21963983]

Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, Yang B. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic Acids Res. 2011; 39:359–372. [PubMed: 20699274]

Lu X, Hershey DM, Wang L, Bogdanove AJ, Peters RJ. An *ent*-kaurene derived diterpenoid virulence factor from *Xanthomonas oryzae* pv. oryzicola. New Phytol. submitted.

Lupien S, Karp F, Wildung M, Croteau R. Regiospecific cytochrome P450 limonene hydroxylases from mint (Mentha) species. cDNA isolation, characterization, and functional expression of (−)-4S-limonene-3-hydroxylase and (−)-4S-limonene-6-hydroxylase. Arch Biochem Biophys. 1999; 368:181–192. [PubMed: 10415126]

Marques JV, Kim KW, Lee C, Costa MA, May GD, Crow JA, Davin LB, Lewis NG. Next generation sequencing in predicting gene function in podophyllotoxin biosynthesis. J Biol Chem. 2013; 288:466–479. [PubMed: 23161544]

Martin VJJ, Pitera DJ, Withers ST, Newman JD, Keasling JD. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. Nature biotechnology. 2003; 21:796–802.

Miao J, Guo D, Zhang J, Huang Q, Qin G, Zhang X, Wan J, Gu H, Qu LJ. Targeted mutagenesis in rice using CRISPR-Cas system. Cell research. 2013; 23:1233–1236. [PubMed: 23999856]

Morrone D, Chambers J, Lowry L, Kim G, Anterola A, Bender K, Peters RJ. Gibberellin biosynthesis in bacteria: Separate *ent*-copalyl diphosphate and *ent*-kaurene synthases in *Bradyrhizobium japonicum*. FEBS Lett. 2009; 583:475–480. [PubMed: 19121310]

Morrone D, Chen X, Coates RM, Peters RJ. Characterization of the kaurene oxidase CYP701A3, a multifunctional cytochrome P450 from gibberellin biosynthesis. Biochem J. 2010a; 431:337–344. [PubMed: 20698828]

Morrone D, Hillwig ML, Mead ME, Lowry L, Fulton DB, Peters RJ. Evident and latent plasticity across the rice diterpene synthase family with potential implications for the evolution of diterpenoid metabolism in the cereals. Biochem J. 2011; 435:589–595. [PubMed: 21323642]

Morrone D, Jin Y, Xu M, Choi S-Y, Coates RM, Peters RJ. An unexpected diterpene cyclase from rice: Functional identification of a stemodene synthase. Arch Biochem Biophys. 2006; 448:133–140. [PubMed: 16256063]

Morrone D, Lowry L, Determan MK, Hershey DM, Xu M, Peters RJ. Increasing diterpene yield with a modular metabolic engineering system in E. coli: comparison of MEV and MEP isoprenoid precursor pathway engineering. Appl Microbiol Biotechnol. 2010b; 85:1893–1906. [PubMed: 19777230]

Morrone D, Xu M, Fulton DB, Determan MK, Peters RJ. Increasing complexity of a diterpene synthase reaction with a single residue switch. J Am Chem Soc. 2008; 130:5400–5401. [PubMed: 18366162]

Moses T, Pollier J, Thevelein JM, Goossens A. Bioengineering of plant (tri)terpenoids: from metabolic engineering of plants to synthetic biology in vivo and in vitro. New Phytol. 2013; 200:27–43. [PubMed: 23668256]

Nelson D, Werck-Reichhart D. A P450-centric view of plant evolution. Plant J. 2011; 66:194–211. [PubMed: 21443632]

Nguyen DT, Gopfert JC, Ikezawa N, Macnevin G, Kathiresan M, Conrad J, Spring O, Ro DK. Biochemical conservation and evolution of germacrene A oxidase in asteraceae. J Biol Chem. 2010; 285:16588–16598. [PubMed: 20351109]

Okada A, Shimizu T, Okada K, Kuzuyama T, Koga J, Shibuya N, Nojiri H, Yamane H. Elicitor induced activation of the methylerythritol phosphate pathway towards phytoalexin biosynthesis in rice. Plant Mol Biol. 2007; 65:177–187. [PubMed: 17634747]

Paddon CJ, Keasling JD. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. Nature reviews Microbiology. 2014; 12:355–367. [PubMed: 24686413]

Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell MD, Tai A, Main A, Eng D, et al. High-level semi-synthetic production of the potent antimalarial artemisinin. Nature. 2013; 496:528–532. [PubMed: 23575629]

Peters RJ. Uncovering the complex metabolic network underlying diterpenoid phytoalexin biosynthesis in rice and other cereal crop plants. Phytochemistry. 2006; 67:2307–2317. [PubMed: 16956633]

Peters RJ. Two rings in them all: The labdane-related diterpenoids. Nat Prod Rep. 2010; 27:1521–1530. [PubMed: 20890488]

Potter K, Criswell J, Peters RJ. Mechanistic analysis of the ent-copalyl diphosphate synthases required for plant gibberellin hormone biosynthesis leads to novel product chemistry. Angew Chem. 2014 in press.

Prisic S, Xu M, Wilderman PR, Peters RJ. Rice contains two disparate *ent*-copalyl diphosphate synthases with distinct metabolic functions. Plant Physiol. 2004; 136:4228–4236. [PubMed: 15542489]

Renault H, Bassard JE, Hamberger B, Werck-Reichhart D. Cytochrome P450-mediated metabolic engineering: current progress and future challenges. Current opinion in plant biology. 2014; 19:27–34. [PubMed: 24709279]

Rios-Estepa R, Lange I, Lee JM, Lange BM. Mathematical modeling-guided evaluation of biochemical, developmental, environmental, and genotypic determinants of essential oil composition and yield in peppermint leaves. Plant Physiol. 2010; 152:2105–2119. [PubMed: 20147490]

Rios-Estepa R, Turner GW, Lee JM, Croteau RB, Lange BM. A systems biology approach identifies the biochemical mechanisms regulating monoterpenoid essential oil composition in peppermint. Proc Natl Acad Sci U S A. 2008; 105:2818–2823. [PubMed: 18287058]

Ro DK, Paradise E, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature. 2006; 440:940–943. [PubMed: 16612385]

Ro DK, Arimura G, Lau SY, Piers E, Bohlmann J. Loblolly pine abietadienol/abietadienal oxidase PtAO (CYP720B1) is a multifunctional, multisubstrate cytochrome P450 monooxygenase. Proc Natl Acad Sci U S A. 2005; 102:8060–8065. [PubMed: 15911762]

Sakamoto T, Miura K, Itoh H, Tatsumi T, Ueguchi-Tanaka M, Ishiyama K, Kobayashi M, Agrawal GK, Takeda S, Abe K, et al. An overview of gibberellin metabolism enzyme genes and their related mutants in rice. Plant Physiol. 2004; 134:1642–1653. [PubMed: 15075394]

Schmelz EA, Huffaker A, Sims JW, Christensen SA, Lu X, Okada K, Peters RJ. Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins. Plant J. 2014; 79:659–678. [PubMed: 24450747]

Shimura K, Okada A, Okada K, Jikumaru Y, Ko K-W, Toyomasu T, Sassa T, Hasegawa M, Kodama O, Shibuya N, et al. Identification of a biosynthetic gene cluster in rice for momilactones. J Biol Chem. 2007; 282:34013–34018. [PubMed: 17872948]

Swaminathan S, Morrone D, Wang Q, Fulton DB, Peters RJ. CYP76M7 is an *ent*-cassadiene C11α-hydroxylase defining a second multifunctional diterpenoid biosynthetic gene cluster in rice. Plant Cell. 2009; 21:3315–3325. [PubMed: 19825834]

Teoh KH, Polichuk DR, Reed DW, Covello PS. Molecular cloning of an aldehyde dehydrogenase implicated in artemisinin biosynthesis in *Artemisia annua*. Botany. 2009; 87:635–642.

Teoh KH, Polichuk DR, Reed DW, Nowak G, Covello PS. Artemisia annua L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone aremisinin. FEBS Lett. 2006; 580:1411–1416. [PubMed: 16458889]

Toyomasu T. Recent Advances Regarding Diterpene Cyclase Genes in Higher Plants and Fungi. Biosci Biotechnol Biochem. 2008; 72:1168–1175. [PubMed: 18460786]

Toyomasu T, Usui M, Sugawara C, Otomo K, Hirose Y, Miyao A, Hirochika H, Okada K, Shimizu T, Koga J, et al. Reverse-genetic approach to verify physiological roles of rice phytoalexins: characterization of a knockdown mutant of *OsCPS4*phytoalexin biosynthetic gene in rice. Physiologia plantarum. 2013

van Herpen TW, Cankar K, Nogueira M, Bosch D, Bouwmeester HJ, Beekwilder J. Nicotiana benthamiana as a production platform for artemisinin precursors. PLoS One. 2010; 5:e14222. [PubMed: 21151979]

Vranova E, Coman D, Gruissem W. Network analysis of the MVA and MEP pathways for isoprenoid synthesis. Annu Rev Plant Biol. 2013; 64:665–700. [PubMed: 23451776]

Wang Q, Hillwig ML, Okada K, Yamazaki K, Wu Y, Swaminathan S, Yamane H, Peters RJ. Characterization of CYP76M5-8 indicates metabolic plasticity within a plant biosynthetic gene cluster. J Biol Chem. 2012a; 287:6159–6168. [PubMed: 22215681]

Wang Q, Hillwig ML, Peters RJ. CYP99A3: Functional identification of a diterpene oxidase from the momilactone biosynthetic gene cluster in rice. Plant J. 2011; 65:87–95. [PubMed: 21175892]

Wang Q, Hillwig ML, Wu Y, Peters RJ. CYP701A8: A rice *ent*-kaurene oxidase paralog diverted to more specialized diterpenoid metabolism. Plant Physiol. 2012b; 158:1418–1425. [PubMed: 22247270]

Wilderman PR, Xu M, Jin Y, Coates RM, Peters RJ. Identification of *syn*-pimara-7,15-diene synthase reveals functional clustering of terpene synthases involved in rice phytoalexin/allelochemical biosynthesis. Plant Physiol. 2004; 135:2098–2105. [PubMed: 15299118]

Williams DC, McGarvey DJ, Katahira EJ, Croteau R. Truncation of limonene synthase preprotein provides a fully active 'pseudomature' form of this monoterpene cyclase and reveals the function of the amino-terminal arginine pair. Biochemistry. 1998; 37:12213–12220. [PubMed: 9724535]

Wu Y, Hillwig ML, Wang Q, Peters RJ. Parsing a multifunctional biosynthetic gene cluster from rice: Biochemical characterization of CYP71Z6 & 7. FEBS Lett. 2011; 585:3446–3451. [PubMed: 21985968]

Wu Y, Wang Q, Hillwig ML, Peters RJ. Picking sides: Distinct roles for CYP76M6 and -8 in rice oryzalexin biosynthesis. Biochem J. 2013; 454:209–216. [PubMed: 23795884]

Wu Y, Zhou K, Toyomasu T, Sugawara C, Oku M, Abe S, Usui M, Mitsuhashi W, Chono M, Chandler PM, et al. Functional characterization of wheat copalyl diphosphate synthases elucidates the early evolution of labdane-related diterpenoid metabolism in the cereals. Phytochemistry. 2012; 84:40–46. [PubMed: 23009878]

Xiao M, Zhang Y, Chen X, Lee EJ, Barber CJ, Chakrabarty R, Desgagne-Penix I, Haslam TM, Kim YB, Liu E, et al. Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. Journal of biotechnology. 2013; 166:122–134. [PubMed: 23602801]

Xu M, Galhano R, Wiemann P, Bueno E, Tiernan M, Wu W, Chung IM, Gershenzon J, Tudzynski B, Sesma A, et al. Genetic evidence for natural product-mediated plant-plant allelopathy in rice (*Oryza sativa*). New Phytol. 2012; 193:570–575. [PubMed: 22150231]

Xu M, Hillwig ML, Lane AL, Tiernan MS, Moore BS, Peters RJ. Characterization of an orphan diterpenoid biosynthetic operon from *Salinispora arenicola*. J Nat Prod. submitted.

Yamane H. Biosynthesis of phytoalexins and regulatory mechanisms of it in rice. Biosci Biotechnol Biochem. 2013; 77:1141–1148. [PubMed: 23748776]

Zerbe P, Hamberger B, Yuen MM, Chiang A, Sandhu HK, Madilao LL, Nguyen A, Hamberger B, Bach SS, Bohlmann J. Gene discovery of modular diterpene metabolism in nonmodel systems. Plant Physiol. 2013; 162:1073–1091. [PubMed: 23613273]

Zhang H, Zhang J, Wei P, Zhang B, Gou F, Feng Z, Mao Y, Yang L, Zhang H, Xu N, et al. The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. Plant biotechnology journal. 2014; 12:797–807. [PubMed: 24854982]

Zhou H, Liu B, Weeks DP, Spalding MH, Yang B. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. Nucleic Acids Res. 2014; doi: 10.1093/nar/gku806

Zhou K, Xu M, Tiernan MS, Xie Q, Toyomasu T, Sugawara C, Oku M, Usui M, Mitsuhashi W, Chono M, et al. Functional characterization of wheat ent-kaurene(-like) synthases indicates continuing evolution of labdane-related diterpenoid metabolism in the cereals. Phytochemistry. 2012; 84:47–55. [PubMed: 23009879]

Zi J, Peters RJ. Characterization of CYP76AH4 clarifies phenolic diterpenoid biosynthesis in the Lamiaceae. Org Biomol Chem. 2013; 11:7650–7652. [PubMed: 24108414]

Summary

Synthetic biology offers a means to explore plant metabolism, such as the smaller terpenoid natural products reviewed here. This includes both investigation of biosynthetic pathways via reconstruction in heterologous host (micro)organisms and the nascent application of genome-editing tools to enable reverse genetic studies of physiological function.
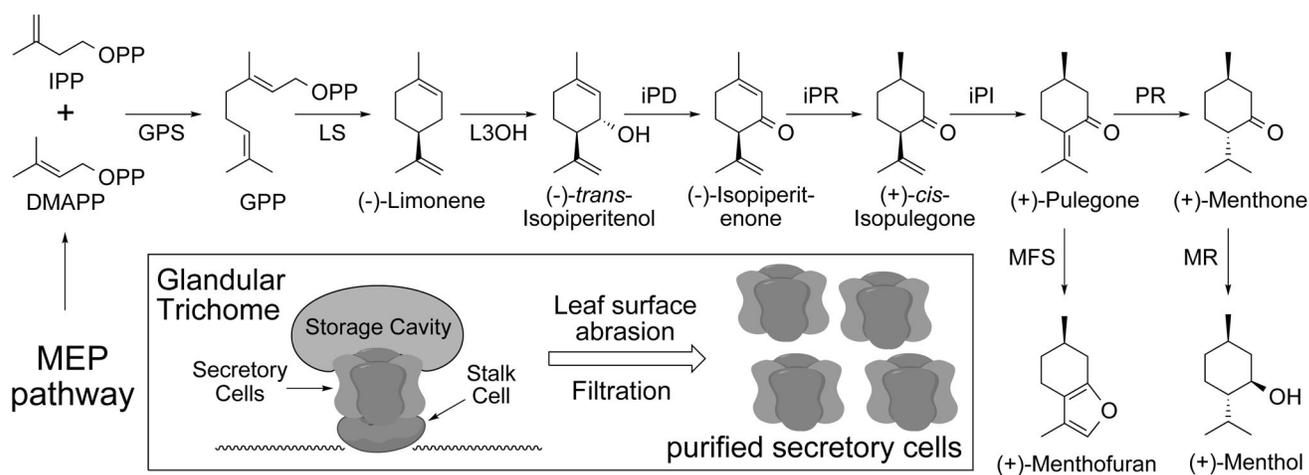
Figure 1. Menthol biosynthetic pathway from glandular trichomes of peppermint
The inset depicts a glandular trichome on the leaf surface, and schematic of the secretory
cell isolation process. The isoprenyl diphosphate precursors IPP and DMAPP are produced
via the plastidial MEP pathway, with subsequent more specific enzymatic steps individually
shown. DMAPP is condensed with IPP to form geranyl diphosphate (GPP) by a GPP
synthase (GPS), which is then cyclized by a TPS, limonene synthase (LS). The resulting (−)-
limonene is transformed to (−)-*trans*-isopiperitenol by limonene-3-hydroxylase (L3OH),
with the 3α-hydroxyl group oxidized to the ketone of (−)-isopiperitenone by isopiperitenol
dehydrogenase (iPD). The endo-cyclic double bond is stereospecifically reduced to (+)-*cis*-
isopulegone by isopiperitenone reductase (iPR), with subsequent isomerization of the exo-
cyclic double-bond to (+)-pulegone by isopulegone isomerase (iPI). Pulegone is a
branchpoint metabolite, which can either be transformed to the undesirable by-product (+)-
menthofuran by menthofuran synthase (MFS) or reduced to the intermediate (+)-menthone
by pulegone reductase (PR), which is then reduced to the desired (+)-menthol end-product
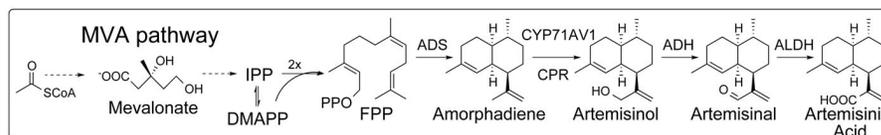by menthone reductase (MR).

Figure 2. Schematic for production of artemisinic acid by *E. coli*
Genes for the MVA pathway, to produce IPP from acetyl-CoA via mevalonate, were imported from *S. cerevisiae*, along with those encoding an isopentenyl diphosphate isomerase and farnesyl diphosphate synthase. The resulting farnesy diphosphate (FPP) is then converted to artemisinic acid using enzymatic genes from the native producing plant *Artemisia annua*. First FPP is cyclized by a TPS, amorphadiene synthase (AS), and then amorphadiene is hydroxylated by CYP71AV1 (in conjunction with a CPR), which can further oxidize the resulting artemisinol to arteminsinal and then artemisinic acid, although this is much more efficiently catalyzed by artemisinol and artemisinal dehydrogenases (ADH and ALDH, respectively). Isolation of the artemisinic acid "end" product is followed by chemical conversion to the anti-malarial drug artemisinin.
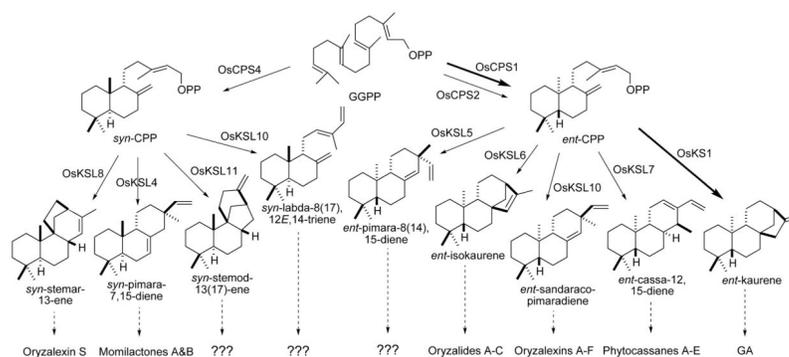
Figure 3. Metabolic map of the rice diterpene synthases

Shown are the functional diterpene synthases, both CPSs and KS(L)s, from rice (*Oryza sativa*), alongside the reaction catalyzed by each of these OsCPS and OsKS(L) (thicker arrows indicate the reactions/enzymes involved in gibberellin (GA) phytohormone biosynthesis rather than more specialized secondary metabolism). Also shown are the derived natural products, where known. Notably, the metabolic fate of several of the resulting diterpene olefins, which can be found in planta, is not yet known (as indicated by "???").
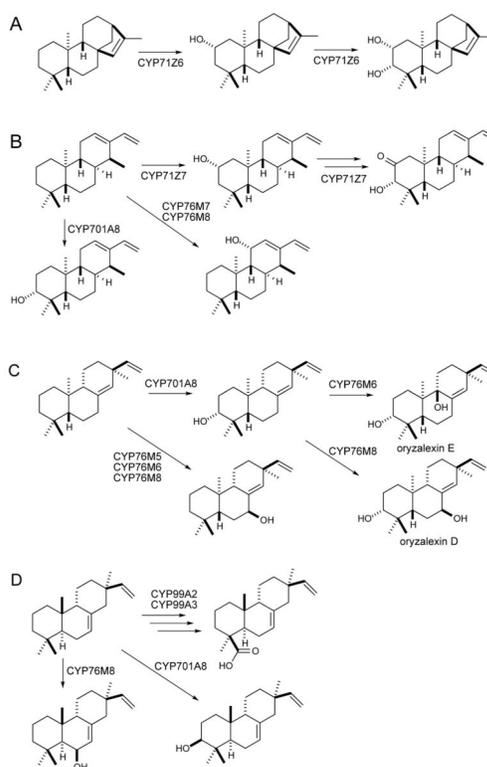
Figure 4. Characterized CYPs from rice diterpenoid biosynthesis

Depicted are reactions catalyzed by the noted rice CYPs with endogenous diterpene olefins. A) CYP71Z6 catalyzes C2α- and then C3α-hydroxylation of *ent*-isokaurene. These reactions may be relevant to the production of oryzadione and related diterpenoids (Kitaoka et al., submitted). B) With *ent*-cassadiene alternative reactions are catalyzed by CYP701A8 [C3α-hydroxylation (Wang et al., 2012b)], CYP76M7 & 8 [both catalyze C11α-hydroxylation (Wang et al., 2012a)], or CYP71Z7 [C2α-hydroxylation followed by further oxidation to a C2-keto and additional C3α-hydroxylation (Kitaoka et al., submitted)]. The C11α-hydroxylation reaction catalyzed by CYP76M7 & 8 is relevant to phytocassane biosynthesis (Wang et al., 2012a), while those catalyzed by CYP71Z7 and (potentially) CYP701A8 may be relevant as well. C) While CYP76M5, 6 & 8 all catalyze C7β-hydroxylation of *ent*-sandaracopimaradiene, it seems likely that this olefin first undergoes C3α-hydroxylation catalyzed by CYP701A8, with subsequent hydroxylation at C7β catalyzed by CYP76M8 forming oryzalexin D and alternative C9β-hydroxylation catalyzed by CYP76M6 to form oryzalexin E (Wu et al., 2013). D) Both CYP99A2 & 3 can convert C19 of *syn*-pimaradiene to a carboxylic acid, although this olefin can be alternatively hydroxylated at C3β by CYP701A8 or C7β by CYP76M8. The C19 oxidation catalyzed by CYP99A2 & 3 is relevant to momilactone biosynthesis (Shimura et al., 2007), while those catalyzed by CYP701A8 and CYP76M8 may be relevant as well (Kitaoka et al., submitted).