The Chemical Logic of Plant Natural Product Biosynthesis

Gülbenk Anarat-Cappillino, **Elizabeth S. Sattely**

Department of Chemical Engineering, Stanford University, Stanford, CA 94305

Abstract

Understanding the logic of plant natural product biosynthesis is important for three reasons: it guides the search for new natural products and pathways, illuminates the function of existing pathways in the context of host biology, and builds an enabling 'parts list' for plant and microbial metabolic engineering. In this review, we highlight the chemical themes that underlie a broad range of plant pathways, dividing pathways into two parts: scaffold-generating steps that draw on a limited set of chemistries, and tailoring reactions that produce a wide range of end products from a small number of common scaffolds.

Introduction:

It is often noted that plants produce upwards of 200,000 distinct small molecule natural products (NPs). This capacity for natural product biosynthesis makes plants unique among multicellular organisms, and is comparable to the natural product diversity in bacteria and unicellular fungi [1]. The wealth of recent plant genome information, combined with decades of biochemical work, has allowed us to begin mapping how CO_2, incorporated during photosynthesis, is converted into abundant and diverse biomass. However, these metabolic networks also reveal large gaps in our understanding of biosynthetic pathways that lead to specialized plant NPs. This aspect of metabolism is critical for fitness with functions including pathogen and pest resistance and nutrient acquisition. It also includes metabolites that are invaluable as human nutrients, commodity products, and therapeutics. Knowledge of plant biosynthetic pathways provides valuable opportunities for metabolic engineering, as well as access to chemical transformations unique to plants. In addition, it allows a deeper understanding of how plant secondary metabolism impacts both plant [2••] and human health [1].

Despite tremendous effort and advances, at present there are only a handful of 'complete' secondary metabolic pathways in which all enzymes and intermediates are known. Much of the current plant secondary metabolic knowledge is a patchwork of well-characterized biosynthetic enzymes in proposed metabolic networks strengthened by our understanding of plant biochemistry, gene co-expression analysis and protein-protein interactions. From this

Corresponding author: Sattely, Elizabeth S. (sattely@stanford.edu).

admittedly incomplete picture of secondary metabolism, patterns emerge that provide insight into the logic of how plants make molecules, and possibly the evolutionary history of plant NP biosynthesis.

Notably, many known plant biosynthetic pathways appear to involve generation of a key branch point intermediate, or scaffold, that is differentially tailored or enzymatically decorated to create families of NPs. Surprisingly, only a small collection of enzymatic transformations lead to these intermediates (Figure 1); how, then, how do plants generate >200,000 different small molecules? The vast diversity of plant NPs seems to arise in two distinct (and orthogonal) ways: (1) scaffold-generating transformations that utilize distinct but closely related substrates, and (2) differential tailoring of a specific scaffold (Figure 2).

In this review, we summarize the key scaffold-generating transformations and tailoring chemistries that serve as starting points for generating the diversity of plant NPs. To close, we briefly discuss the recent progress in discovering and characterizing these enzymatic reactions. It is important to note that the examples discussed below do not cover all branches of plant metabolism, but instead highlight the general themes of diversity generation in secondary metabolite biosynthesis in plants.

Key Scaffold-Generating Chemistries

A defining feature of plant secondary metabolism is the formation of key scaffolds through the use of a small collection of chemical transformations using abundant building blocks from primary metabolism as substrates. The starting materials for these scaffold-generating chemistries are typically only a few metabolic steps from amino acids, with the exception of polyketides and terpenes (both of which are derived from acetyl-CoA) [3]. The following list highlights common scaffold-generating chemical transformations (Table 1), but is not comprehensive. For more information on the biosynthesis of individual classes of NPs (e.g. polyketides, terpenoids, and alkaloids), the interested reader is referred to additional recent reviews [4,5,6,7•].

Amine-aldehyde condensation.

Amine-aldehyde condensation is a classic example of a scaffold-generating chemistry: this transformation is efficient at physiological pH in aqueous solution, and it is an efficient method of generating chemical complexity from simple precursors. It can occur either inter- or intra-molecularly. Intermolecular amine-aldehyde coupling is promoted by Pictet-Spenglerases (PSRs), of which two notable, distinct variants - strictosidine synthase [8] (STR1) and norcoclaurine synthase [9] (NOR) - have been biochemically characterized. These reactions play a key role in generating strictosidine and norcoclaurine scaffolds, which are central biosynthetic precursors to the widely distributed indole- and isoquinoline-alkaloid NPs [10]. In both cases, the amine is derived from an aromatic amino acid, while the aldehyde is in the former case a terpene, and in the latter case an oxidized form of tyrosine. The resulting scaffolds are optimally situated for the formation of additional rings by oxidative tailoring that lead to some 5000 structurally distinct plant alkaloids [11]. Furthermore, recent efforts have shown that PSRs display some degree of plasticity, making these enzymes attractive targets for metabolic engineering. For example, wild-type and

rationally engineered STR1 can be utilized to produce unnatural alkaloids *in planta* [12,13], and most recently, a novel function of STR1 was recognized, which allows the conversion of the more common tryptoline scaffold into the less frequently observed piperazino[1,2-*a*]indole scaffold [14].

Intramolecular amine–aldehyde condensation is an equally powerful transformation for generating cyclic alkaloids, albeit less well understood. Diamine (e.g., putrescine) or polyamine (e.g., spermidine) precursors derived from basic amino acids and/or *S*-adenosylmethionine are converted by copper amine oxidases (CuAOs) to amino aldehydes that spontaneously cyclize [15]. Diamines give rise to monocyclic scaffolds, whereas polyamines are set up for additional cyclization(s). To date, the only characterized example is the *Nicotiana benthamiana* CuAOs that converts *N*-methylputrescine to *N*-methyl- [1]-pyrrolinium cation, which is the central intermediate for tropane alkaloid and nicotine biosynthesis [16].

Non-canonical oxidation of amino acids.

Since plant secondary metabolism relies heavily on oxidative transformations (including tailoring chemistries), it is no surprise that unusual oxidations of amino acid precursors lead to the key scaffolds in the biosynthesis of cruciferous phytoalexins and glucosinolates [17]. Oxidative decarboxylation of a precursor amino acid, catalyzed by dedicated cytochromes P450 (such as CYP79 family) [18], generates key scaffold aldoximes that give rise to a diverse group of heteroatom-rich phytoalexins such as camalexin, brassinin, and their derivatives [19•]. Furthermore, the resulting aldoxime scaffold sits at a metabolic branch point of several classes of plant secondary metabolites and can be converted into chemically distinct 'warheads' such as the cyanohydrin of cyanogenic glucosides [20] or the thiohydroxamate of glucosinolates [21].

Aryl-CoA acylation.

Acylation of an aryl-CoA substrate, catalyzed by type III polyketide synthases (PKSs), is a key scaffold-generating transformation that leads to the formation of a variety of diverse polyketide intermediates including chalcones, flavonoids, stilbenes, and acridone alkaloids [22]. The general mechanism of plant PKSs consists of a series of decarboxylative condensations of a starter aryl-CoA unit with multiple malonyl-CoA molecules. This generates a linear polyketide intermediate that is further cyclized to form the canonical polycylic scaffolds [4]. Typically, the identity of the aryl-CoA substrate and the number of malonyl-CoA condensations, along with the catalytic environment of the PKS, define the chemical features of the polyketide intermediate that is formed. A notable substrate of the plant PKSs in secondary metabolism is coumaryl-CoA, which is derived from the deamination of phenylalanine. This molecule is particularly abundant in plants due to its central role as the key building block for lignin biosynthesis [23]. It has been co-opted for secondary metabolism in two ways: as a starting material for making the flavonoid, stilbene and lignan scaffolds, and as a scaffold itself for coumarin and α-pyrone biosynthetic pathways, which are discussed below in the scaffold tailoring section. In the case of flavonoids, acylation of coumaryl-CoA and its derivatives by chalcone sythases leads to the central chalcone intermediate naringenin, which gives rise to several branches of flavonoid

metabolism [24]. In contrast, a similar acylation of the coumaryl-CoA substrate by another PKS, stilbene synthase (STS) generates a different scaffold - resveratrol - in stilbene phytoalexin synthesis [25]. This catalytic diversity and broad substrate specificity make plant PKSs an excellent platform for metabolic engineering, and recent efforts have shown that unnatural polyketide-alkaloid scaffolds can be synthesized using a combinatorial approach of precursor-directed biosynthesis and protein engineering [26,27••].

C–C bond formation through radical coupling.

Oxidative coupling is another important scaffold-generating transformation that yields a central dimerized intermediate in the biosynthesis of lignans, lignin, and other polyphenolic NPs. In the canonical lignan biosynthesis pathway, oxidative phenoxy coupling of coumaryl-CoA derived alcohols, promoted by designated laccases and dirigent proteins, provides entry to lignan pathways via dimerized scaffolds, such as pinoresinol [28]. More importantly, this initial reaction determines the enantiomeric configuration of the resulting lignan, as two radicals are forced to undergo regio- and stereospecific coupling that is directed by dirigent (DIR) proteins. DIRs and DIR-like proteins are found exclusively in plants, and are thought to play important roles in enantiospecific coupling of C–C or C–O bond formation in plant secondary metabolism [29]. Recent efforts have shown that various DIRs for the formation of enantiocomplementary (+)- and (−)-pinoresinol are found within the plant kingdom [30], and that the engineering of dirigent activities is possible as demonstrated by the reversal of the coupling mode of pinoresinol-forming DIRs in *Arabidopsis* and *Schizandra* species [31•].

Carbocation-mediated cyclizations.

A cascade C–C bond formation of polyolefinic carbocations via cyclization is a powerful way to generate large, polycyclic carbon-based scaffolds in terpene biosynthesis. This crucial step gives rise to more than 20,000 distinct terpene metabolites by utilizing two common isomeric C5 units: isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) [6]. While present in bacteria and fungi, terpene synthases are especially prolific in plant secondary metabolism. Recent work has shown that the genes encoding these enzymes and the tailoring enzymes that decorate the reduced scaffold are often "clustered": they are physically proximal on the chromosome [32]. Furthermore, due to their astonishing capability to synthesize complex metabolites from readily available common precursors - IPP and DMAPP -exciting progress has been made with engineering terpenoid metabolism in plants [33] and in heterologous hosts [34]. Although many terpene synthases have been biochemically characterized, recent reports continue to highlight the extraordinary diversity of scaffolds that can be generated from a very limited set of acyclic precursor pyrophosphates [35]. While this diversity is mainly derived from the carbocation intermediates, alternative connectivities by which the C5 units can be connected by terpene synthases are another key source of chemical complexity [36]. In addition, recent work by O'Connor and coworkers demonstrated that irregular cyclic terpene-scaffolds can be generated via novel mechanisms such as the reductive cyclization of 10-oxogeranial, catalyzed by iridoid synthase, which generates iridoid, another key scaffold in terpene metabolism [37••]. Yet, another source of metabolic diversity comes from the participation

of these terpene intermediates in other scaffold-generating chemistries, such as the role of the secologanin terpene moiety in strictosidine biosynthesis [8].

Variations on a Theme: Scaffold Tailoring

Once a scaffold is created using the chemistries above, it can be tailored to meet the specific molecular needs of the plant. Variations in scaffold-tailoring chemistry are arguably the most significant source of plant natural product diversity. Large families of closely related compounds are produced through either altered regiochemistry of the same scaffold modification type, catalyzed by related tailoring enzymes with different substrate specificities, or the combination of entirely different classes of tailoring chemical transformations. The examples shown in Figure 3 highlight both types of metabolic logic, and demonstrate how they can lead to collections of compounds that are significantly altered from the precursor scaffold (covered in the Kliebenstein review in this issue). The majority of scaffold modifications can be grouped into two categories: redox chemistry and group transfer including alkylation, acylation, and glycosylation.

Closely related oxidations + skeletal rearrangements.

Redox chemistry, particularly oxidation, is especially prevalent in plant secondary metabolism. In addition to hydroxylations (such as those catalyzed by canonical cytochromes P450), unusual rearrangements, ring closures, and functional group interconversions that occur through formal oxidation greatly expand the diversity of plant NP families [38].

In the Brassicaceae family of plants, sulfur-containing metabolites are critical for plant defense, including glucosinolates and indole derived phytoalexins [17]. The formation of the key C-S bonds that are characteristic of these metabolites is often enabled by cytochrome P450-catalyzed oxidation to generate a reactive intermediate that then traps a sulfur nucleophile [39,40], either through the intermolecular addition of glutathione, or through the intramolecular addition of a pendant thiol. Although current understanding of the chemistry downstream of brassinin in cruciferous plants is based on isotope feeding studies [41,42], it appears that closely related enzymes with redox capabilities, such as cytochromes P450, could be responsible for tailoring and rearranging the brassinin scaffold to produce a suite of indolic phytoalexins abundant in these edible plants.

The use of oxidation to set up subsequent bond formations is also a theme in other branches of metabolism. In many of these cases, highly similar transformations catalyzed by closely related enzymes (e.g., hydroxylating at adjacent positions) can lead to divergent structural outcomes in the product by setting up spontaneous skeletal rearrangements. Two examples are the conversion of coumaryl-CoA to either the coumarin [43] or recently discovered *a*-pyrone [44••] family of NPs, and the amazing diversity of strictosidine and norcoclaurine derivatives (see the de Luca review in this issue).

Combining oxidation with group transfer.

Plants also employ a combinatorial approach to scaffold tailoring as another way to accomplish structural diversity in NP families. Species-specific combinations of oxidation

and group transfer to C-, O-, N-, and S-atoms such as alkylation (prenylation and methylation), acylation, and glycosylation are found within the plant kingdom, resulting in diverse pools of NPs [45]. Various combinations of group transfer and/or oxidative tailoring can lead to closely related analogs of a secondary metabolite family (e.g. tropinones), or generate entirely new metabolic scaffold branch points (e.g. steroidal saponins). For example, the aliphatic sidechain of the cholesterol scaffold is modified by a series of oxidative and group transfer enzymes that results in amine insertion to a set of steroidal saponins that include tomatidine [46] and, presumably, cyclopamine [47]. Regardless of the extent of diversity created by combinations of tailoring chemistry, many NPs generated in this way have distinct biological function from the parent scaffold and other tailored derivatives.

Conclusion: What's Next for Pathway Discovery

Viewing the logic of plant pathways from a chemical perspective highlights two areas that are likely to see important advances in the near term. First, there are certain scaffold-generating chemistries about which very little is known, such as the intramolecular amine–aldehyde condensations that lead to polycyclic alkaloids [48]. Second, there are numerous pathways for which the complete set of biosynthetic genes has not yet been compiled, despite important clues about key intermediates from feeding studies and analogies to other pathways. The remarkable challenges of pathway discovery in the absence of genome sequencing data – a molecule-to-gene discovery approach – make it all the more impressive that so much is known even at this point.

Genome sequencing opens the possibility of a far more rapid molecule-to-gene approach, and potentially a gene-to-molecule approach. While assembling a complete pathway remains a significant challenge, three genome-enabled strategies for connecting genes into a pathway have shown great promise: the discovery of physical gene clustering in plants [32], co-expression analysis [49], and phylogenetic analysis [50]. In addition, the ability to express candidate genes and pathways in alternative plant hosts [51] and yeast [52,53] sets the stage for a powerful synthetic biology discovery platform. These newer technologies will be particularly enabling for exploring metabolism in non-model, genetically intractable plant hosts.

Acknowledgments

References

[1]. Dewick PM: Medicinal Natural Products: A Biosynthetic Approach, 3rd Edition. John Wiley and Sons; 2009.

[2••]. Weng J-K, Philippe RN, Noel JP: The rise of chemodiversity in plants. Science 2012, 336:1667–1670.In this review, authors provide a complementary perspective to the chemical logic of plant metabolism that is highlighted above. They discuss how permissiveness (e.g. ability to accept a broader range of substrates, tolerance to mutations, catalyzing alternative reactions) may partially explain the chemodiversity of plant natural products.

[3]. Croteau R, Kutchan TM, Lewis NG: Natural products: Secondary Metabolites In Biochemistry and Molecular Biology of Plants. Edited by Buchanan B, Gruissem W, Jones R. American Society of Plant Physiologists; 2000:1250–1318.

[4]. Flores-Sanchez IJ, Verpoorte R: Plant Polyketide Synthases: A fascinating group of enzymes. Plant Physiol and Biochem 2009, 47:167–174. [PubMed: 19071029]

[5]. Degenhardt J, Kollner TG, Gershenzon J: Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. Phytochem 2009, 70:1621–1637.

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

[7•]. Cordell GA: Fifty years of alkaloid biosynthesis in *Phytochemistry*. Phytochem 2013, 91:2951.This review provides an updated summary of alkaloid biosynthesis studies.

[8]. Maresh JJ, Giddings LA, Friedrich A, Loris EA, Panjikar S, Trout BL, Stöckigt J, Peters B, O'Connor SE: Strictosidine Synthase: Mechanism of a Pictet-Spengler Catalyzing Enzyme. J Am Chem Soc 2008, 130:710–723. [PubMed: 18081287]

[9]. Luk LY, Bunn S, Liscombe DK, Facchini PJ, Tanner ME: Mechanistic studies on norcoclaurine synthase of benzylisoquinoline alkaloid biosynthesis: an enzymatic Pictet-Spengler reaction. Biochemistry 2007, 46:10153–10161. [PubMed: 17696451]

[10]. Stockigt J, Antonchick AP, Wu F, Waldmann H: The Pictet-Spengler Reaction in Nature and in Organic Chemistry. Angew Chem IntEd 2011, 50:8538–8564.

[11]. Ziegler J, Facchini PJ: Alkaloid Biosynthesis: Metabolism and Trafficking. Annu Rev Plant Biol 2008, 59:735–769. [PubMed: 18251710]

[12]. Runguphan W, Maresh JJ, O'Connor SE: Silencing of tryptamine biosynthesis for production of non-natural alkaloids in plant culture. Proc Natl Acad Sci U S A 2009, 106:13673–13678. [PubMed: 19666570]

[13]. Runguphan W, O'Connor SE: Metabolic reprogramming of periwinkle plant culture. Nat Chem Biol 2009, 5:151–153. [PubMed: 19151732]

[14]. Wu F, Zhu H, Sun L, Rajendran C, Wang M, Ren X, Panjikar S, Cherkasov A, Zou H, Stockigt J: Scaffold Tailoring by a Newly Detected Pictet-Spenglerase Activity of Strictosidine Synthase: From the Common Tryptoline Skeleton to the Rare Piperazino-indole Framework. J Am Chem Soc 2012, 134: 1498–1500. [PubMed: 22229634]

[15]. Gupta K, Dey A, Gupta B: Plant polyamines in abiotic stress responses. Acta Physol Plant 2013, 35:2015–2036.

[16]. Katoh A, Shoji T, Hashimoto T: Molecular cloning of N-methylputrescine oxidase from tobacco. Plant Cell Physiol 2007, 48:550–554. [PubMed: 17283012]

[17]. Pedras MSC, Yaya EE, Glawischnig E: The phytoalexins from cultivated and wild crucifers: Chemistry and biology. Nat Prod Rep 2011, 28:1381–1405. [PubMed: 21681321]

[18]. Nafisi M, Sønderby IE, Hansen BG, Geu-Flores F, Nour-Eldin HH, Nørholm MHH, Jensen NB, Li J, Halkier BA: Cytochromes P450 in the biosynthesis of glucosinolates and indole alkaloids. Phytochem Rev 2006, 5:331–346.

[19•]. Bednarek P: Sulfur-Containing Secondary Metabolites from *Arabidopsis thaliana* and other Brassicaceae with Function in Plant Immunity. ChemBioChem 2012, 13:1846–1859. [PubMed: 22807086] This review summarizes the evolution of sulfur-containing secondary metabolites found in Brassicaceae and how they are interconnected on the biosynthetic level, as well as their function in plant immunity.

[20]. Møller B Lindberg: Functional diversifications of cyanogenic glucosides. Curr Opin Plant Biol 2010, 13:337–346.

[21]. Fahey JW, Zalcmann AT, Talalay P: The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. Phytochem 2001, 56:5–51.

[22]. Weng JK, Noel JP: Structure-Function analyses of Plant Type III Polyketide Synthases. Meth Enzymol 2012, 515:317–335. [PubMed: 22999180]

[23]. Vogt T: Phenylpropanoid biosynthesis. Mol Plant 2010, 3:2–20. [PubMed: 20035037]

[24]. Dao TTH, Linthorst HJM, Verpoorte R: Chalcone synthase and its functions in plant resistance. Phytochem Rev 2011, 10:397–412. [PubMed: 21909286]

[25]. Stewart C Jr, Vickery CR, Burkart MD, Noel JP: Confluence of structural and chemical biology: plant polyketide synthases as biocatalysts for a bio-based future. Curr Opin Plant Biol 2013, 16:365–372. [PubMed: 23481348]

[26]. Morita H, Yamashita M, Shi SP, Wakimoto T, Kondo S, Kato R, Sugio S, Kohno T, Abe I: Synthesis of unnatural alkaloid scaffolds by exploiting plant polyketide synthase. Proc Natl Acad Sci US A 2011, 108:13504–13509.

[27••]. Abe I: Novel applications of plant polyketide synthases. Curr Opin Chem Biol 2012, 16:179–185. [PubMed: 22245533] This review summarizes recent efforts focused on enzymatic synthesis of unnatural alkaloid scaffolds by using a combinatorial approach of precursor-directed biosynthesis and protein engineering. Authors highlight the catalytic plasticity of plant type III PKSs.

[28]. Davin LB, Wang HB, Crowell AL, Bedgar DL, Martin DM, Sarkanen S, Lewis NG: Stereoselective bimolecular phenoxy radical coupling by an auxiliary (Dirigent) protein without an active center. Science 1997, 275:362–367. [PubMed: 8994027]

[29]. Pickel B, Schaller A: Dirigent proteins: molecular characteristics and potential biotechnological applications. Appl Microbiol Biotechnol 2013, 97:8427–8438. [PubMed: 23989917]

[30]. Pickel B, Constantin MA, Pfannstiel J, Conrad J, Beifuss U, Schaller A: An enantiocomplementary dirigent protein for the enantioselective laccase-catalyzed oxidative coupling of phenols. Angew Chem IntEd 2010, 49:202–204.

[31•]. Kim KW, Moinuddin SGA, Atwell KM, Costa MA, Davin LB, Lewis NG: Opposite stereoselectivities of dirigent proteins in *Arabidopsis* and *Schizandra* species. J Biol Chem 2012, 287:33957–33972. [PubMed: 22854967] Authors engineered chimeric dirigent proteins and demonstrated that the reversal of the coupling mode in pinoresinol-forming DIRs is possible.

[32]. Nutzmann H-W, Osbourn A: Gene clustering in plant specialized metabolism. Curr Opin Biotechnol 2014, 26:91–99. [PubMed: 24679264]

[33]. Lange BM, Ahkami A: Metabolic engineering of plant monoterpenes, sesquiterpenes and diterpenes - current status and future opportunities. Plant Biotech J. 2012, 11:169–196.

[34]. Marienhagen J, Bott M: Metabolic engineering of microorganisms for the synthesis of plant natural products. J Biotech 2013, 163:166–178.

[35]. Yeo YS, Nybo SE, Chittiboyina AG, Weerasooriya AD, Wang YH, Gongora-Castillo E, Vaillancourt B, Buell CB, DellaPenna D, Celiz MD et al.: Functional identification of valerena-1,10-diene synthase, a terpene synthase catalyzing a unique chemical cascade in the biosynthesis of biologically active sesquiterpenes in *Valeriana officinalis*. J Biol Chem 2013, 288:3163–3173. [PubMed: 23243312]

[36]. Ferrer J-L, Jez JM, Bowman ME, Dixon RA, Noel JP: Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. Nat Struct Biol 1999, 6: 775–784. [PubMed: 10426957]

[37••]. Geu-Flores F, Sherden NH, Courdavault V, Burlat V, Glenn WS, Wi C, Nims E, Cui Y, O'Connor SE: An alternative route to cyclic terpenes by reductive cyclization in iridoid biosynthesis. Nature 2012, 492:138–142. [PubMed: 23172143] The authors report the discovery of iridoid synthase, which catalyzes the formation of irregular cyclic terpene scaffolds, namely iridoids via reductive cyclization of 1-oxogeranial. This paper also exemplifies the 'gene-to-molecule' strategy of pathway discovery.

[38]. Mizutani M, Sato F: Unusual P450 reactions in plant secondary metabolism. Arch Biochem Biophys 2011, 507:194–203. [PubMed: 20920462]

[39]. Hansen CH, Du L, Naur P, Olsen CE, Axelsen KB, Hick AJ, Pickett JA, Halkier BA: CYP83B1 is the oxime-metabolizing enzyme in the glucosinolate pathway in *Arabidopsis*. J Biol Chem 2001, 276:24790–24796. [PubMed: 11333274]

[40]. Klein AP, Anarat-Cappillino G, Sattely ES: Minimum set of cytochromes P450 for reconstituting the biosynthesis of camalexin, a major *Arabidopsis* antibiotic. Angew Chem Int Ed Engl 2013, 52:13625–13628. [PubMed: 24151049]

[41]. Monde K, Takasugi M, Ohnishi T: Biosynthesis of cruciferous phytoalexins. J Am Chem Soc 1994, 116:6650–6657.

[42]. Pedras MSC, Loukaci A, Okanga FI: The cruciferous phytoalexins brassinin and cyclobrassinin are intermediates in the biosynthesis of brassilexin. Bioorg Med Chem Lett 1998, 8:3037–3038. [PubMed: 9873671]

[43]. Bourgaud E, Hehn A, Larbat R, Doerper S, Gontier E, Kellner S, Matern U: Biosynthesis of coumarins in plants: a major pathway still to be unraveled for cytochrome P450 enzymes. Phytochem Rev 2006, 5:293–308.

[44••]. Weng JK, Li Y, Mo H, Chapple C: Assembly of an evolutionarily new pathway for a-pyrone biosynthesis in *Arabidopsis*. Science 2012, 337:960–964. [PubMed: 22923580] The authors report the evolution of a-pyrone class of metabolites - arabidopyrones - in Arabidopsis, as well as the discovery of a neofunctionalized cytochrome P450 enzyme (CYP84A4) that generates a catechol-substituted substrate in arabidopyrone pathway. This paper exemplifies how the introduction of a new chemical transformation to an existing catalytic framework expands chemical diversity in plants.

[45]. Veitch NC: Isoflavonoids of the Leguminosae. Nat Prod Rep 2013, 30:988–1027. [PubMed: 23736284]

[46]. Itkin M, Heinig U, Tzfadia O, Bhide AJ, Shinde B, Cardenas PD, Bocobza SE, Unger T, Malitsky S, Finkers R et al. :Biosynthesis of antinutritional alkaloids in Solanaceous crops is mediated by clustered genes. Science 2013, 341:175–179. [PubMed: 23788733]

[47]. BETH WILL PUT IN A REFERENCE FOR THIS.

[48]. Cona A, Rea G, Angelini R, Federico R, Tavladoraki P: Functions of amine oxidases in plant development and defence. Trends PlantSci 2006, 11:80–88.

[49]. Marques JV, Kim K-W, 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]

[50]. Matsuno M, Compagnon V, Schoch GA, Schmitt M, Debayle D, Bassard JE, Pollet B, Hehn A, Heintz D, Ullmann P et al.: Evolution of a novel phenolic pathway for pollen development. Science 2009, 325:1688–1692. [PubMed: 19779199]

[51]. Møldrup ME, Salomonsen B, Halkier BA: Engineering of glucosinolate biosynthesis: candidate gene identification and validation. Methods Enzymol 2012, 515:291–313. [PubMed: 22999179]

[52]. Hawkins KM, Smolke CD: Production of benzylisoquinoline alkaloids in *Saccharomyces cerevisiae*. Nat Chem Biol 2008, 4:564–573. [PubMed: 18690217]

[53]. Keasling JD: Manufacturing molecules through metabolic engineering. Science 2010, 330:1355–1358. [PubMed: 21127247]

Highlights

- Chemical diversity in plant natural products arises in two ways

- Scaffold-generating enzymes draw on a limited set of chemistries

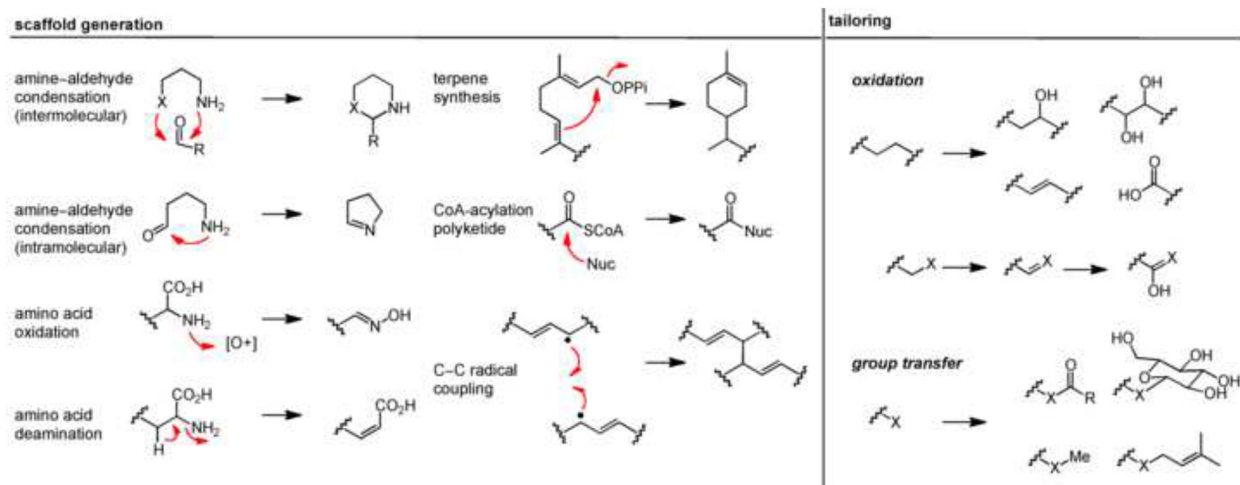- Scaffolds are chemically tailored by redox and group transfer enzymes

Figure 1.
Key chemical reactions in plant natural product biosynthesis. (a) Common scaffold-generating enzymatic steps are highlighted in the left panel, and (b) Common tailoring themes are highlighted in the right panel.
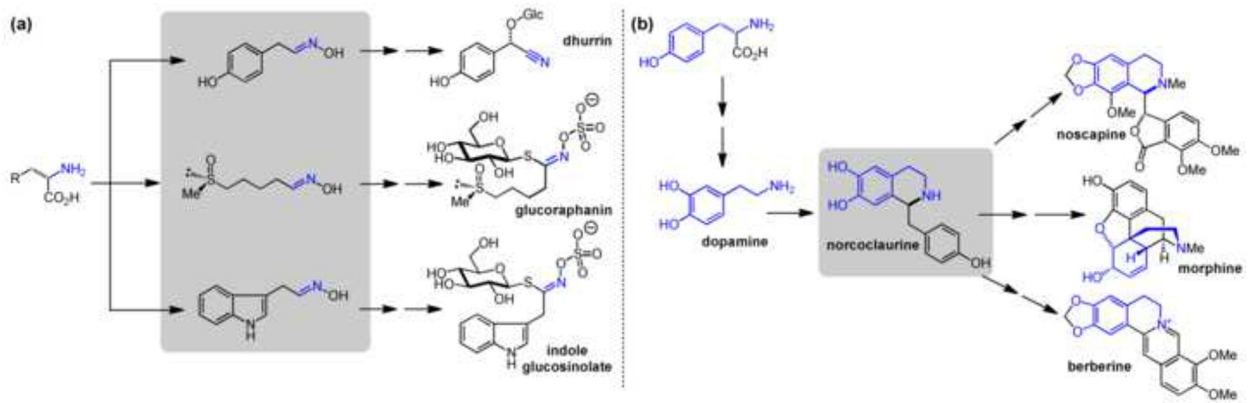
Figure 2.
Diversity in plant natural products can be established by two general chemical strategies.
Examples shown illustrate these two strategies. (a) Common scaffold-generating chemical
reactions generate diversity by utilizing closely related, but distinct substrates. (b) Key
scaffold intermediates generate diversity by differential tailoring reactions. Multiple arrows
indicate multiple chemical transformations.
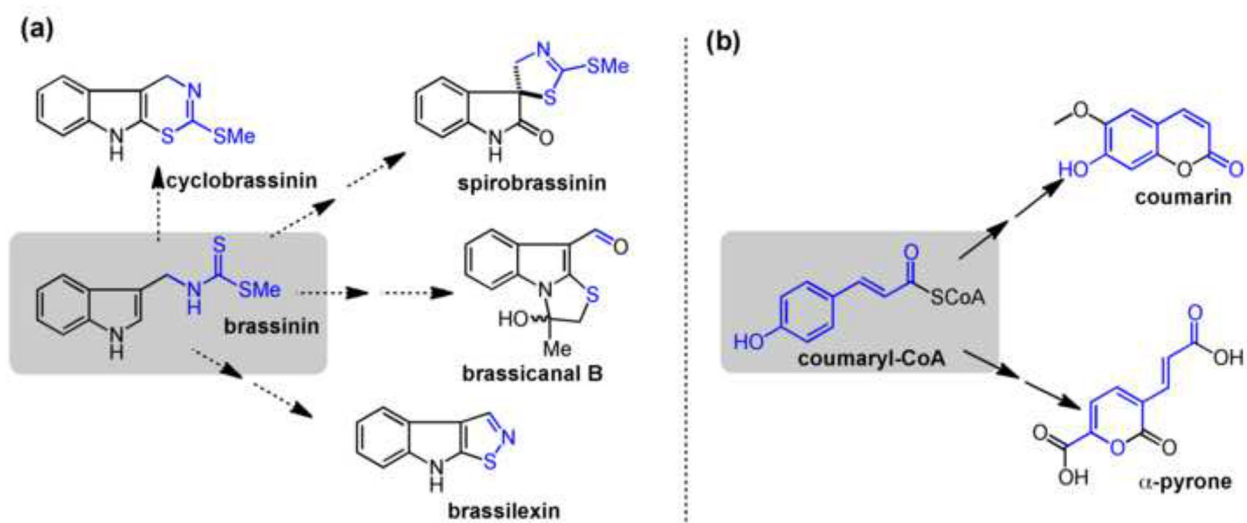
Figure 3.
Closely related oxidations and resulting skeletal rearrangements in (a) cruciferous phytoalexin biosynthesis and (b) coumaryl-CoA tailoring. Dotted arrows indicate predicted biosynthetic steps.
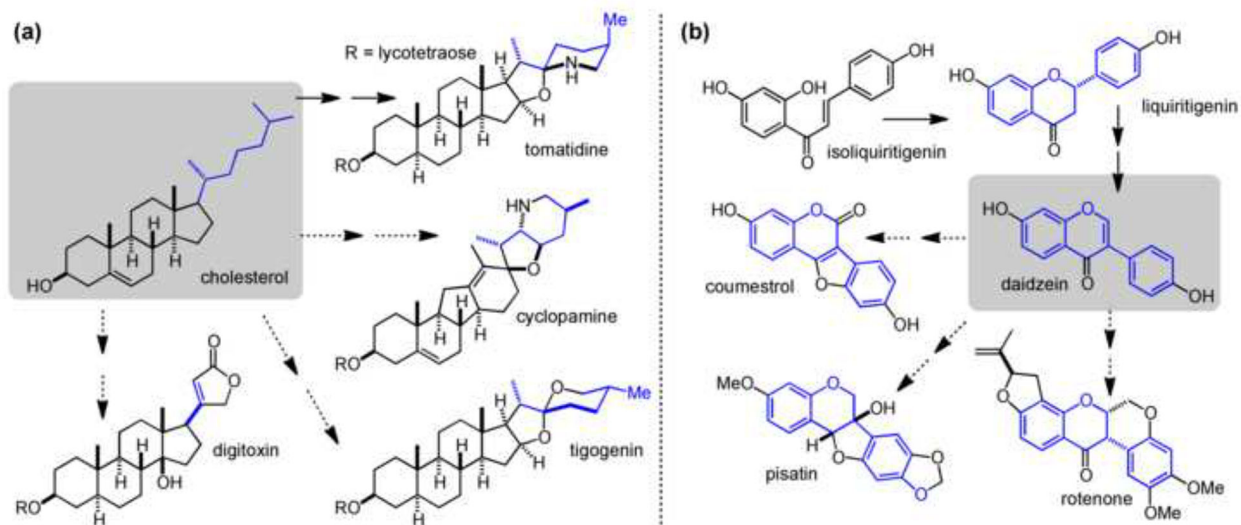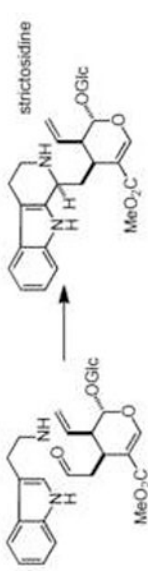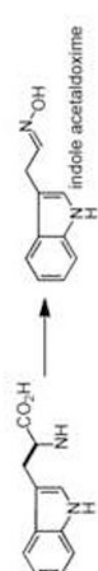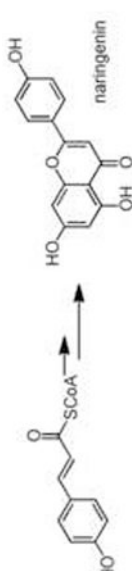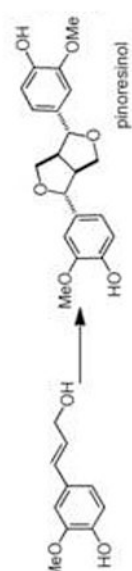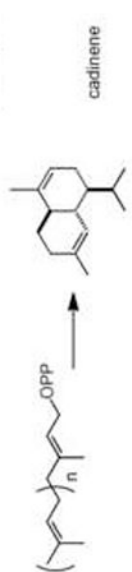
Figure 4.
Combinations of tailoring chemistries create diverse families of plant natural products. (a) Side chain tailoring of cholesterol scaffold leads to a range of steroidal saponins. (b) Tailoring of daidzein scaffold expands family of isoflavonoids.

Table 1.

Scaffold-generating chemistries and lead examples

Chemistry Type	Substrate(s)	Lead example	Enzyme(s)
amine-aldehyde condensation (intermolecular) eg Pictet-Spengler	amine, aldehyde	strictosidine	Strictosidine synthase - STR1 Norcoclaurine synthase - NOR
amine-aldehyde condensation (intramolecular)	di- or polyamine amino ketone	N-methyl pyrrolinium cation	Copper amine oxidase - CuAO N-methyl putricine oxidase - MPOI
amino acid oxidation	amino acid	indole acetaldoxime	CYP79F1/F2 - aliphatic aldoximes CYP79A1/A2 - aromatic aldoximes CYP7962/B3 - indole-3-acel aldoximes
CoA-acylation polyketide	acyl- or aryl-CoA	naringenin	Chalcone synthase - CHS Stilbene synthase - STS
C-C radical coupling	monolignol(s)	pinoresinol	Laccases and dirigent proteins -DIR *Forsythia intermedia* DIR1 *Arabdopsis thaliana* DIR6
terpene cyclization	prenyl pyrophosphate	cadinene	Cadinene synthase - COS

Chemistry Type	Substrate(s)	Lead example	Enzyme(s)
amino aod deamination	amino acid	coumaryl-CoA	Phenylalanine ammonia lyase - PAL Tyrosine ammonia lyase - TAL