

organelles. Addressing these areas of basic research, a renewed focus on characterizing the subtleties of lipid metabolism among natural and GM plants, and improved gene stacking and gene editing technologies should enable engineers of the next generation of GM plants to make better use of the lipidome for the benefit of all.

REFERENCES AND NOTES

1. K. B. Reynolds *et al.*, *Front. Plant Sci.* **6**, 164 (2015).
2. H. T. Nguyen *et al.*, *Plant Biotechnol. J.* **13**, 38–50 (2015).
3. X. Li *et al.*, *Plant Biotechnol. J.* **10**, 862–870 (2012).
4. D. Huai, Y. Zhang, C. Zhang, E. B. Cahoon, Y. Zhou, *PLoS ONE* **10**, e0131755 (2015).
5. A. R. Snapp, J. Kang, X. Qi, C. Lu, *Planta* **240**, 599–610 (2014).
6. T. A. Voelker, T. R. Hayes, A. M. Cranmer, J. C. Turner, H. M. Davies, *Plant J.* **9**, 229–241 (1996).
7. T. Vanhercke *et al.*, *Plant Biotechnol. J.* **12**, 231–239 (2014).
8. J. Zale *et al.*, *Plant Biotechnol. J.* **14**, 661–669 (2016).
9. J. A. Napier, S. Usher, R. P. Haslam, N. Ruiz-Lopez, O. Sayanova, *Eur. J. Lipid Sci. Technol.* **117**, 1317–1324 (2015).
10. J. R. Petrie *et al.*, *PLoS ONE* **9**, e85061 (2014).
11. J. Liu *et al.*, *Plant Biotechnol. J.* **13**, 858–865 (2015).
12. H. J. Kim *et al.*, *J. Exp. Bot.* **66**, 4251–4265 (2015).
13. P. D. Bates *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1204–1209 (2014).
14. Y. Guo *et al.*, *Plant Mol. Biol.* **69**, 565–575 (2009).
15. P. D. Bates, J. Browse, *Plant J.* **68**, 387–399 (2011).
16. Z. Hu, Z. Ren, C. Lu, *Plant Physiol.* **158**, 1944–1954 (2012).
17. P. J. Horn *et al.*, *Plant J.* **86**, 322–348 (2016).
18. E. Mietkiewska *et al.*, *Planta* **240**, 575–583 (2014).
19. X. H. Yu, R. R. Prakash, M. Sweet, J. Shanklin, *Plant Physiol.* **164**, 455–465 (2014).
20. K. Roessler *et al.*, *Plant Physiol.* **171**, 878–893 (2016).
21. N. Elahi, R. W. Duncan, C. Stasolla, *Plant Physiol. Biochem.* **100**, 52–63 (2016).
22. A. A. Kelly, E. Shaw, S. J. Powers, S. Kurup, P. J. Eastmond, *Plant Biotechnol. J.* **11**, 355–361 (2013).
23. A. A. Kelly *et al.*, *Plant Physiol.* **162**, 1282–1289 (2013).
24. R. Sanjaya *et al.*, *Plant Cell* **25**, 677–693 (2013).
25. Y. Yang *et al.*, *Plant Physiol.* **169**, 1836–1847 (2015).
26. S. M. Kwak, S.-K. Myung, Y. J. Lee, H. G. Seo, *Arch. Intern. Med.* **172**, 686–694 (2012).
27. W. Stonehouse *et al.*, *Am. J. Clin. Nutr.* **97**, 1134–1143 (2013).
28. N. Ruiz-Lopez, R. P. Haslam, J. A. Napier, O. Sayanova, *Plant J.* **77**, 198–208 (2014).
29. N. Tejera *et al.*, *J. Nutr.* **146**, 227–235 (2016).
30. M. B. Betancor *et al.*, *Sci. Rep.* **5**, 8104 (2015).
31. S. K. Gidda *et al.*, *Plant Physiol.* **170**, 2052–2071 (2016).
32. C. H. Tsai *et al.*, *Plant J.* **83**, 650–660 (2015).
33. P. Jolivet *et al.*, *Proteomics* **13**, 1836–1849 (2013).
34. G. J. H. van Rooijen, M. M. Moloney, *Nat. Biotechnol.* **13**, 72–77 (1995).
35. H. M. Nguyen *et al.*, *Proteomics* **11**, 4266–4273 (2011).
36. S. Yi *et al.*, *Protein Expr. Purif.* **116**, 127–132 (2015).
37. C. H. Tsai *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15833–15838 (2014).
38. B. J. Olson *et al.*, *Plant Cell* **22**, 3331–3347 (2010).
39. H. Tjellström, M. Strawsine, J. Silva, E. B. Cahoon, J. B. Ohlrogge, *FEBS Lett.* **587**, 936–942 (2013).

ACKNOWLEDGMENTS

The Benning lab is supported by the Chemical Sciences, Geosciences, and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy grants DE-FG02-91ER20021 and DE-FG02-98ER20305; NSF grants MCB-1157231 and MCB-1515169; U.S. Department of Energy–Great Lakes Bioenergy Research Center Cooperative Agreement DE-FC02-07ER64494; and Michigan State University AgBioResearch.

10.1126/science.aaf6206

REVIEW

Plant metabolism, the diverse chemistry set of the future

Eleanore T. Wurtzel^{1,2*} and Toni M. Kutchan^{3*}

New technologies are redefining how plant biology will meet societal challenges in health, nutrition, agriculture, and energy. Rapid and inexpensive genome and transcriptome sequencing is being exploited to discover biochemical pathways that provide tools needed for synthetic biology in both plant and microbial systems. Metabolite detection at the cellular and subcellular levels is complementing gene sequencing for pathway discovery and metabolic engineering. The crafting of plant and microbial metabolism for the synthetic biology platforms of tomorrow will require precise gene editing and delivery of entire complex pathways. Plants sustain life and are key to discovery and development of new medicines and agricultural resources; increased research and training in plant science will accelerate efforts to harness the chemical wealth of the plant kingdom.

From the ancient valleys of Mesopotamia to the Amazonian rainforests, mankind has sought plants for food and nutrition, biomaterials for living, and treatment of pain and disease. Traditional medicine worldwide has relied on the medicinal properties of plants, such as the opium poppy, *Papaver somniferum*, first cultivated by the Sumerians in 3400 B.C.E. Plant extracts provided ceremonial dyes, such as henna from *Lawsonia inermis*, still in use today, and the essential oils such as rose oil have provided sensual pleasure since antiquity. Cotton fibers from *Gossypium* plant species were first woven into clothing thousands of years ago, and papyrus used for writing in ancient Egypt evolved into the paper industry of today. Selected examples of this structural diversity and for which use humanity has exploited it are given in Fig. 1.

The chemical diversity of plants is enormous. Plants evolved the biosynthesis of a cornucopia of novel chemicals to survive and communicate in a complex ecological environment. Although some plant chemicals are sharp or bitter tasting (glucosinolates and pyrrolizidine alkaloids) to deter herbivory, others such as anthocyanins and carotenoids are brightly colored flower pigments that attract pollinators. Chemicals that are cytotoxic or otherwise physiologically active in mammals are used, for example, as pain-killers, chemotherapeutics, and other drugs. All of these plant chemicals are made through species-specific, specialized biochemical pathways that modify metabolites of primary metabolism. A plethora of new chemicals and metabolic pathways are likely hidden in plant genomes awaiting discovery. Although structures for 200,000 natural products are known, only 15% of the estimated 350,000 plant species have been investigated for their chemical con-

stituents (1). This estimate suggests that a relatively small percentage of the chemical space present in the plant kingdom has been discovered. With every new enzyme and underlying gene discovered, there is potential for a biochemical reaction that can improve modern medicine, human and animal health, bioenergy, and agriculture.

In this review, we present recent examples of selected exciting technologies that are transforming how we study plant metabolism and how we implement these discoveries to develop the plants and microbes of tomorrow. Although the pace of progress during the past several years has accelerated, there are still challenges to be met in order to fully harness the chemical wealth of the plant kingdom.

Accelerating pathway discovery

Inexpensive DNA sequencing, together with computational tools for genome assembly, have revolutionized pathway discovery in plants. The relatively small number of sequenced genomes revealed a surprising distribution of genes encoding specialized biochemical pathway enzymes (2). For a growing number of pathways, enzymes are encoded in clustered genomic regions (2) having shared chromatin signatures to coordinate expression (3). Pathways are also encoded by unlinked genes or a combination of linked and unlinked genes. Specialized plant chemicals function to communicate with the ecological environment, and environmental cues are likely necessary to regulate expression of certain pathways, whether organized in clusters or randomly throughout a genome. The breadth of the chemical space of plants is not known, in part because of reliance on environmental signals for the expression of some pathways. This clustered genome feature can be exploited to identify silent pathways or pathways expressed at levels so low that the metabolites that they produce accumulate below our current levels of chemical detection and have remained undiscovered. Development of computational methods that link chromatin signatures with transcript and metabolite profiles will accelerate discovery of unknown pathways encoded by gene clusters. Development of facile heterologous expression platforms will enable

¹Department of Biological Sciences, Lehman College, The City University of New York, 250 Bedford Park Boulevard West, Bronx, NY 10468, USA. ²The Graduate School and University Center–CUNY, 365 Fifth Avenue, New York, NY 10016-4309, USA. ³Donald Danforth Plant Science Center, 975 North Warsaw Road, St. Louis, MO 63132, USA.

*Corresponding author. Email: wurtzel@lehman.cuny.edu (E.T.W.); tmkutchan@danforthcenter.org (T.M.K.)

functional expression of the large number of genes (20 to 30) required to validate and exploit the biosynthetic pathways to complex plant chemicals.

Recent examples of discovery of clustered biochemical pathway genes illustrate that, given a high-quality genome sequence assembly, biochemical pathways that have long been studied but not completely elucidated at the enzyme and gene levels can be efficiently solved and exploited. Bioinformatic interrogation of genomes of the Solanaceae led to identification of genes encoding enzymes for biosynthesis of steroid alkaloids in

active monoterpene indole alkaloids, including the anticancer compounds vincristine and vinblastine (5). Genes encoding known precursors and intermediates for biosynthesis of monoterpene indole alkaloids were found coexpressed and clustered on the genome. Genic sequences interspersed among and flanking the known genes encode enzymes and transporters that may be candidates for the yet-unknown enzymes of vincristine biosynthesis and transporters that are required to move intermediates along the multicellular pathway to monoterpene indole alkaloids.

genome association analysis led to development of molecular markers for breeding high provitamin A crops needed to address global vitamin A deficiency (8). Genome-wide association study of a germplasm diversity collection led to discovery of nine clustered genes that encode enzymes that catalyze formation of the bitter triterpenoid cucurbitacin of wild cucumber. Further analysis revealed another genome location encoding two transcription factors that regulate the pathway gene cluster (9). Insights into the domestication process and the observation that certain varieties of cucumber

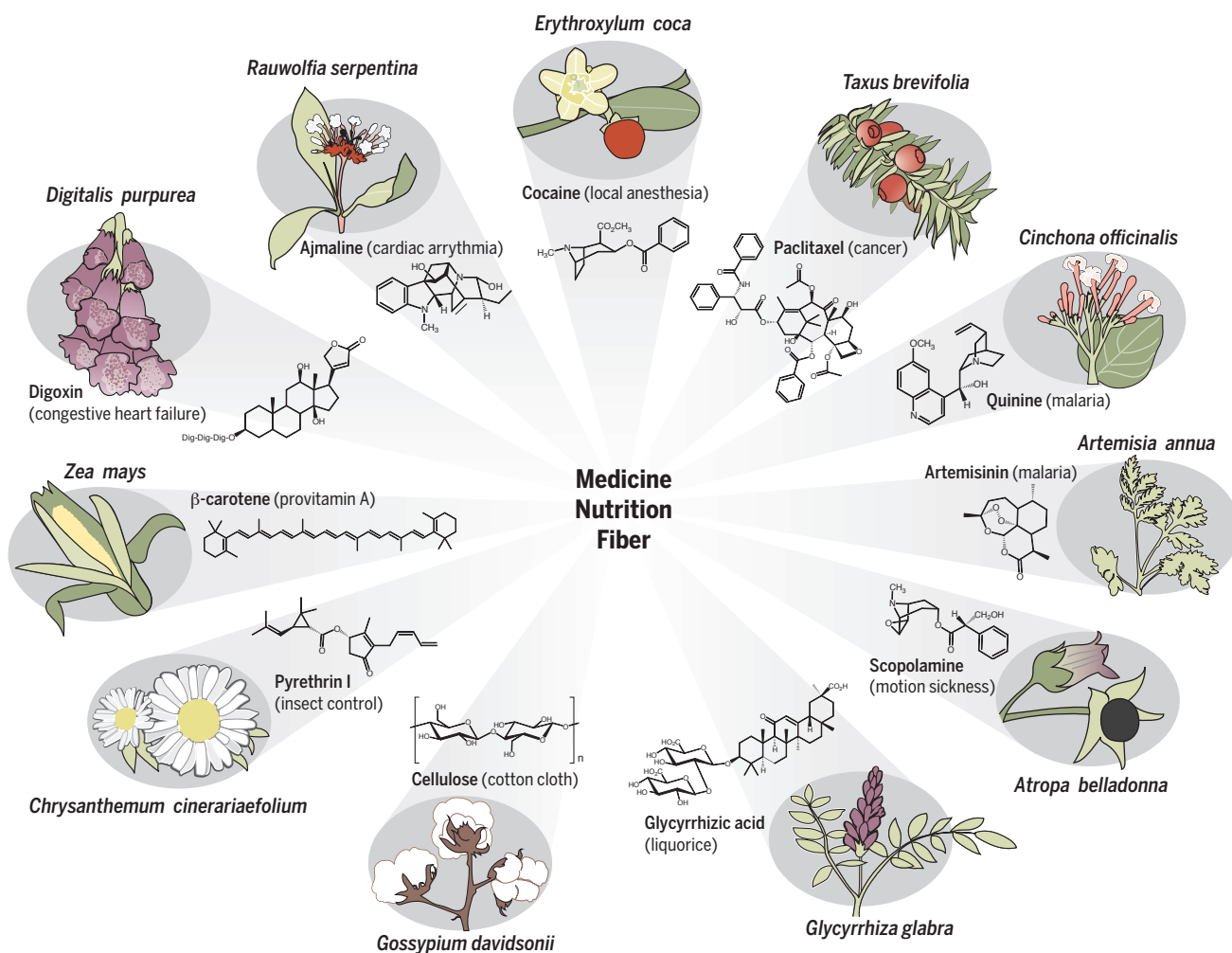


Fig. 1. Selected plants and their uses. Throughout history, plants have served as sources of a plethora of chemicals that provide humankind with medicine, fiber, and nutrition.

two clusters and coexpressed on the genomes of tomato (α -tomatine) and potato (α -solanine). One gene, *GAME4*, encoding a cytochrome P450 along the common steroid alkaloid pathway, was then used to engineer potato tubers and tomato fruit with reduced steroid alkaloids (4). When a high-quality genome is not yet available, insight into specialized metabolism can be gained from the combined use of a draft genome together with transcriptome and metabolite data, as illustrated with indole alkaloids in Madagascar periwinkle, *Catharanthus roseus*. This medicinal plant produces a suite of biologically

Genetic and biochemical variation within a plant species can be used to identify genes associated with phytochemicals. This approach was elegantly demonstrated early on with limited transcript and metabolite data sets with the identification of flavonoid biosynthetic genes and transcription factors in red and green *Perilla frutescens* and, more recently, using large omic data sets in perilla, *Arabidopsis*, and *Oryza sativa* (6, 7). Transcriptome and carotenoid analysis of a maize germplasm collection helped pinpoint key enzymes for optimizing provitamin A carotenoids. Additional ge-

come bitter when exposed to cold can now be exploited in breeding. In the search for genes related to scent production of roses, *Rosa × hybrida* rose varieties differing in scent profiles were examined for differentially expressed genes. A candidate gene was linked to a scent-related terpenoid that is synthesized by a novel pathway not found in other plants producing the identical compound (10). This gene can be used in breeding scent back into modern rose cultivars that have largely lost the aromatic properties of their predecessors. Having a repertoire of enzymes that are structurally unique

but perform the same biochemical transformation also expands the palette for metabolic engineering. The more enzymes that are discovered and characterized that have novel catalytic mechanisms and steady-state kinetic properties, the greater the potential for developing synthetic biology systems for known and novel chemicals.

Although high-quality genome sequencing and assembly are still a challenge with the large genomes of many plants, metabolite and transcriptome data are now easily obtained for any plant, changing the paradigm for biosynthetic pathway discovery. Computational interrogation of large-scale transcriptomes and metabolomes of plant tissues has revealed that many genes encoding pathway enzymes are coexpressed with each other and with accumulation of pathway metabolites. Pathway elucidation is no longer limited to a handful of model plants or to simple biochemical pathways but is a reality for the most complex pathways and for the entire plant kingdom, including rare or endangered species. As an example, the medicinal plant,

Podophyllum hexandrum (mayapple) produces the chemotherapeutic etoposide. Four of the etoposide biosynthetic steps were known, and six others were predicted, on the basis of putative chemical transformations. Investigators mined mayapple transcriptome data using bioinformatic tools for coexpression analysis, and selected candidate genes to fill pathway gaps for predicted biochemical transformations (11). Elucidating pathways to commercially relevant complex plant chemicals can now be accelerated by high-throughput nucleotide sequencing and coexpression analysis, but the challenges associated with complex genomes, low expression levels of genes, and low accumulation levels of metabolites limit the number of metabolites that can be investigated using these technologies. In addition, algorithms used to interrogate these data sets still require human curation.

Peering into the single cell

Biosynthesis of some metabolites, such as medicinal alkaloids, can be compartmentalized across

various subcellular locations in multiple, specialized cell types and can require a dedicated trafficking mechanism (12). Moreover, the induction of alkaloid biosynthesis may be linked to developmental changes signaled through interactions between plants and other organisms. A better understanding of how development and metabolism interconnect and are regulated could have far-reaching impact on the fields of chemical ecology, agriculture, and metabolic engineering. For example, association between rhizobial bacteria and legume hosts elicits formation of nitrogen-fixing nodules in roots, which in the plant *Crotalaria* is accompanied by induction of plant enzymes that catalyze formation of antiherbivore defense compounds, the pyrrolizidine alkaloids. The alkaloids synthesized in the root nodules are subsequently transported to the aboveground parts of the plant (13). Whole-plant or whole-organ metabolite analysis is insufficient to capture such compartmentalization. To optimally match genes with metabolites, instrumentation sensitivity and spatial resolution must discern

compartment-specific metabolites. Recent advances in single-cell metabolite detection and quantification have greatly improved the opportunity to dissect pathway metabolites with temporal and spatial resolution. Metabolite analysis is conducted using an array of complementary analytical instrumentation, each of which offers certain advantages [reviewed in (14, 15)]. Compounds can be identified within complex mixtures (16), and chemical structures can be elucidated from relatively small amounts of extracted samples (17). To obtain a spatial map of metabolites in situ, the current method of choice for nonvolatile compounds is matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) (Fig. 2A). Using MALDI-MSI, together with a time-of-flight analyzer, flax specialized metabolites could be analyzed at 20- μm resolution (18), which affords cellular resolution, as a typical plant cell has a diameter of 50 μm . Using this same technology allowed analysis of various lipid structures in a single cotton seed and revealed an unexpected cell-type compositional variation (19). Improved beam-delivery optics have increased spatial resolution to 5 μm (20), which approaches subcellular resolution and would facilitate an analysis of subcellular metabolite trafficking that is a component of most metabolic pathways. The promising method of laser-ablation electrospray ionization mass spectrometry utilizes fresh tissue without pretreatment, as required by MALDI-MSI, but offers only 200- μm resolution (21), which currently suffices only for tissue-level analyses. Another nondestructive method for three-dimensional metabolite

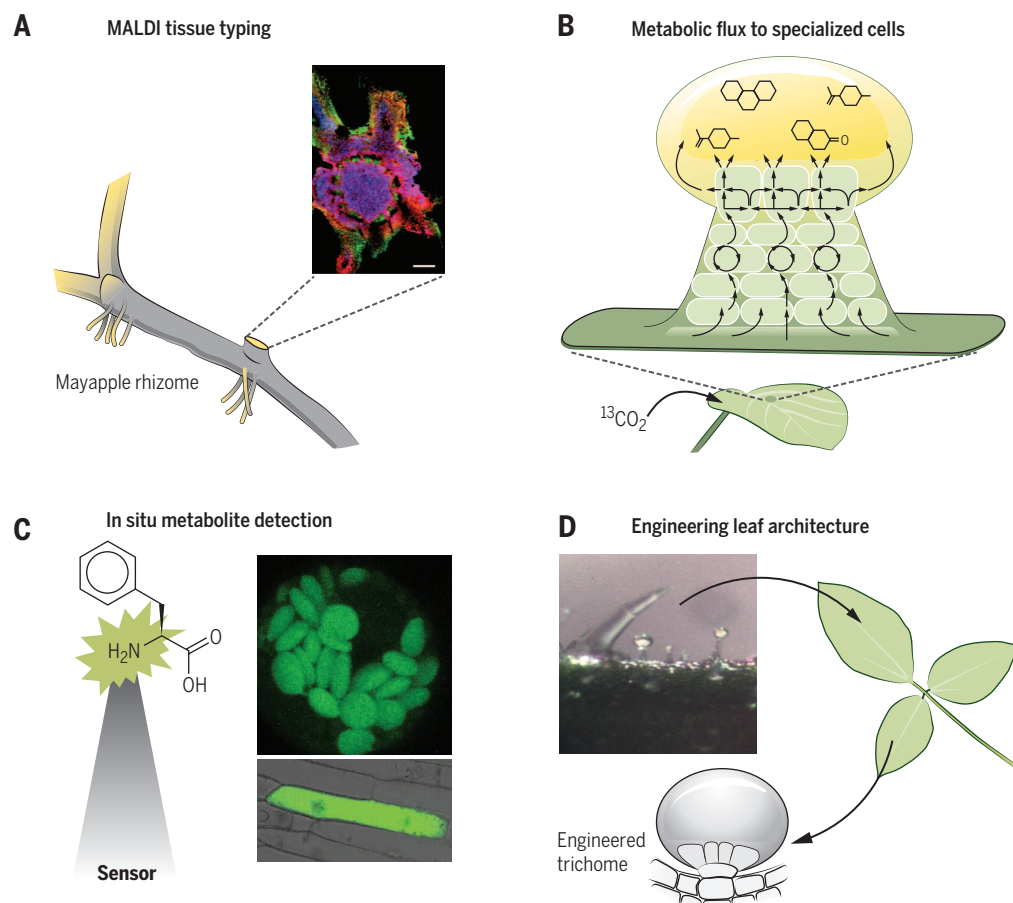


Fig. 2. Game-changing technologies and concepts under development. (A) Merged MALDI MSI of podophyllotoxin [mass/charge ratio (m/z) 453.0969 of the molecular ion plus potassium, $[M+K]^+$; the difference between the calculated and the measured mass (Δ ppm) = 3.8, red]; podophyllotoxin- β -D-glucopyranoside (m/z 615.1496 $[M+K]^+$, Δ ppm = 2.6, blue); and kaempferol-3-O-(6''-O-malonyl)-glucoside (m/z 573.0660 $[M+K]^+$, Δ ppm = 2.3, green) in *Podophyllum hexandrum* rhizome. Scale, 2 mm. (B) Schematic of ^{13}C metabolic flux to specialized metabolites in trichomes. (C) Schematic of in vivo sensors. Development for a large catalog of plant metabolites would greatly facilitate in situ quantitation of metabolites for metabolic engineering. (D) *Salvia sclarea* trichomes. Engineering trichomes to produce specialty chemicals can be achieved today with plants such as mint; the challenge of the future is to engineer trichomes into species that do not normally develop these specialized surface structures.

analysis combines confocal microscopy with Raman spectroscopy to link subcellular architecture with chemical composition at the submicron level (22, 23). To build a more complete cellular and subcellular model of biochemical pathway components, large-scale mass spectrometry-based proteomic efforts are cataloging enzyme locations at the cellular and subcellular level (24), which complements that achieved for metabolites. For rational engineering, it is important to elucidate dynamics of pathway metabolon assembly, including how scaffold proteins affect enzyme allostery in complexes and how protein modification controls enzyme function in vivo. Proteomic surveys, however, do not always detect all pathway components or distinguish among enzyme variants. Alternatively, the effect of sequence variation can be examined *in planta* by confocal microscopy using transient expression of plant genes fused to fluorescent tags to track subcellular localization (25). Together with large-scale protein and metabolite profiling, the complementary use of cell biological tools contributes to elucidating the dynamic structural organization of pathway components. Coupled with single-cell transcriptomics, single-cell metabolite analysis aids in linking metabolites to the genes that underlie biosynthesis and path-

way regulation by reducing the size of omic data sets and, thereby, simplifying computational analysis.

Building biochemical factories

The speed with which we can elucidate biosynthetic pathways to plant natural products is ever increasing, especially with implementation of the exciting new technologies just described. Synthetic biology platforms in microbes and plants are now driving production of plant natural products. With elucidation of complete biosynthetic pathways to selected commercially valuable plant natural products, production platforms in *Escherichia coli*, *Saccharomyces cerevisiae*, *Nicotiana benthamiana*, *Camelina sativa*, and *Populus* hybrids are being developed (Fig. 3). Synthetic metabolic networks are designed with biosynthetic enzymes that originate from plants and other organisms, such as microbes and mammals. Endogenous host pathways can be constructed on scaffolds to improve flux into new pathways that produce useful chemicals. For example, artemisinin, a potent antimalarial from the plant *Artemisia annua* L, is costly and difficult to obtain by chemical synthesis. Plant genes encoding key enzymes for biosynthesis of artemisinic acid, the immediate

precursor for artemisin had been previously identified. To create yeast cells that secrete artemisinic acid that can be chemically converted to the antimalarial drug, a synthetic pathway was constructed by tethering plant metabolic enzymes to a yeast scaffold (26). The artemisinic acid thus formed was secreted and retained outside the yeast cells. Secretion of metabolic products not only facilitates end-product purification but also reduces toxicity to the host cell, a feature that is certain to be further exploited by use of efflux transporters that could be endogenous (27), synthetic, or an existing metabolite transporter from plants.

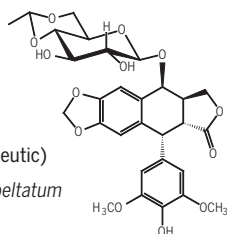
The most recent synthetic biology successes are reported for the benzyloquinoline alkaloids and lignans, a consequence of the extensive knowledge attained on the enzymes and genes and to the pharmacological importance of selected members of these classes of natural products. Although not yet optimized, thebaine and related opiate alkaloids have been produced in yeast (6.4 μg/l thebaine; 0.3 μg/l hydrocodone) (28) and bacteria (2.1 mg/l thebaine) (29) with the melding of more than 20 genes from plants, mammals, bacteria, and yeast. These examples are a first demonstration of engineering pathways to complex alkaloids from glucose or glycerol;

Plants

Tobacco
Nicotiana benthamiana



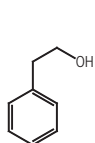
Etoposide
(chemotherapeutic)
Podophyllum peltatum



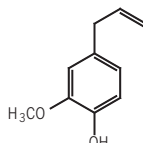
Hybrid poplar
Populus tremula × *P. alba*



2-Phenylethanol
(specialty chemical)
Rosa hybrida
Solanum lycopersicum



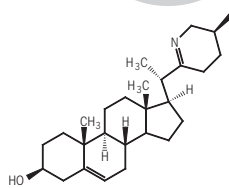
Eugenol
(specialty chemical)
Larrea tridentata



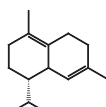
Camelina
Camelina sativa



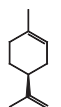
Verazine
(specialty chemical)
Veratrum californicum



(+)-δ-Cadinene
(biofuel)
Gossypium spp.



(4S)-Limonene
(solvent)
Mentha spp.

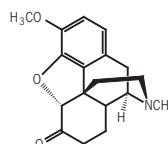


Microorganisms

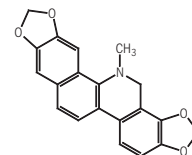
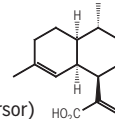
Yeast
Saccharomyces cerevisiae



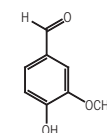
Hydrocodone
(opiate)
Papaver somniferum



Artemisinic acid
(antimalarial precursor)
Artemisia annua



Dihydrosanguinarine
(antibiotic)
Sanguinaria canadensis



Vanilla (flavor & fragrance)
Vanilla planifolia

Bacteria
Escherichia coli



Thebaine
(opiate precursor)
Papaver somniferum

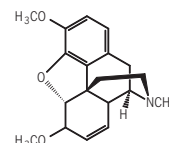


Fig. 3. Natural products produced from plants. Plant natural products currently produced in heterologous systems that range from microbial fermentation to cultivated trees.

commercial feasibility requires that at least 5 g/l be achieved. The entire pathway of 10 genes to the chemotherapeutic etoposide aglycone was successfully reconstructed and validated in tobacco (< 1 ng/mg dry weight) (11). As for the opiates, the levels of drug achieved are far from commercially viable levels, but demonstrate proof of concept. Vanillin, the most important flavor and fragrance component of vanilla extracted from cured pods of the vanilla orchid, is commercially biotechnologically produced using bacteria, fungi, and yeast (30). The synthetic pathway to yet another flavor and fragrance chemical 2-phenylethanol has been introduced into hybrid poplar with ~4% accumulation in leaf and stem of 4-month-old plants (31). Woody plants can accumulate up to 60% phenylpropanoid-derived chemicals in heartwood and this example suggests that trees may serve as platforms for production of commodity chemicals. (4S)-Limonene and (+)- δ -cadinene have been produced in seed of *C. sativa* (~14% dry weight), thereby demonstrating the utility of an oilseed as a host for chemical production (32).

Biochemical engineering of the future

For rational engineering of biosynthesis of plant chemicals, pathway flux reflective of cellular and subcellular metabolism within a metabolic network needs assessment. Now, we mostly work with an enzyme “parts list.” Even if the concentration of an enzyme in situ is known, the level of that enzyme is not indicative of enzymatic activity and does not directly inform on a rate-limiting biosynthetic step or on biochemical regulation. To better understand how carbon progresses through metabolism, ¹³C-labeling-based metabolic flux analysis has been developed to comprehensively quantify flow in metabolic networks (Fig. 2B). Mass spectrometry or nuclear magnetic resonance spectroscopy is used to measure isotope distribution in steady-state metabolite pools. In plants, steady state-labeling studies have quantitatively assessed developing seed metabolism and recently isotopically non-stationary metabolic flux analysis (INST-MFA) has been applied to leaves (33). INST-MFA can be used to study short-lived metabolic steady states and pathways that are linear or use a single carbon precursor and, therefore, require the monitoring of transient labeling to be informative concerning fluxes (e.g., autotrophic or specialized metabolism). If cellular or subcellular attributes are to be defined in a computational model, experimental measurements that serve as inputs to the model must adequately inform about spatial complexity so as to resolve the fluxes in different locations.

To achieve cellular and even subcellular resolution, genetically encoded nanosensors hold great promise as in vivo metabolite beacons for real-time quantification of metabolites without sample destruction. Förster resonance energy transfer (FRET) biosensors consist of two distal fluorescent proteins that sandwich a metabolite ligand-sensing protein domain that alternates in conformation depending on ligand binding and is visualized as

a FRET signal (Fig. 2C). For any metabolite for which there is a known receptor, the ligand-binding domain could be used to develop a FRET sensor (34). For other metabolites, structure-guided protein engineering based on a collection of scaffolds could be used to design artificial ligand-binding domains. Limitations here remain the small number of chemicals for which binding proteins are known. An alternative in vivo quantitation of metabolites is the use of RNA sensors, riboswitches. Natural riboswitches are RNA sensors that bind specific metabolites within a small aptamer region and transduce the binding signal to control gene expression through a regulatory domain (35). Artificial riboswitches that recognize a given metabolite can be created by incorporating synthetic aptamers that are identified through in vitro selection strategies to bind to an immobilized metabolite (36). Aptamers have been created for a few plant alkaloids, such as cocaine (37) and the caffeine biosynthetic intermediate, theophylline (36). Metabolite sensors developed for a large catalog of plant metabolites would facilitate the speed with which we can design metabolic pathways in native and recombinant systems.

Alongside metabolite detection and carbon flux determination, biosynthetic pathway elucidation is also limited by computational methodology. Gene annotation is based on primary sequence homology and many genes are annotated as “unknown” or are misannotated, such as a symporter concealing a new class of heme enzymes (38). More accurate annotation based on tertiary structure necessitates development of high-throughput tools for homology modeling together with expansion of available protein structure templates, the latter now closer to realization with recent advances in structure determination using high-resolution single-particle cryo-electron microscopy (39). Many gene sequences could be more efficiently linked to biological processes with computational tools that simultaneously query multiple large data sets and associate genes with pathways. Algorithms are needed that incorporate machine learning to correlate data from an expanding horizon of data sets—such as RNA, protein, cell-specific metabolites, enzyme localization, protein-protein interactions, plant-microbe interaction, chromatin signatures—for the purpose of linking genes, proteins, and pathways without a priori knowledge.

Metabolic engineering of tomorrow will target a suite of biosynthetically unrelated compounds, such as a multivitamin crop augmented with a nutraceutical, for which metabolic tailoring will also optimize plant growth in a changing climate. Engineering multiple chemical traits requires a fundamental understanding of pathway regulation including flux control, regulatory interactions between specialized and primary metabolism, and coordination with plant development, abiotic signaling, and signaling between plants and other organisms in shared ecological niches. Crafting plant metabolism for crops of tomorrow will be achieved by precise gene editing and delivery of whole pathways on synthetic chromosomes (40). Genome editing has not yet been

applied to synthetic chromosomes; however, genome-editing techniques have the potential to advance meaningfully the applications of synthetic chromosomes. Specialized plant structures, such as trichomes (Fig. 2D), could be exploited to channel compounds for alternate and novel biosynthetic transformations. Advanced tools, as those mentioned herein, are the gateway to accelerating metabolic engineering and stacking of complex traits. These technologies are the future of plant engineering and synthetic biology production of the plethora of plant chemicals useful to humanity.

REFERENCES AND NOTES

- G. M. Cragg, D. J. Newman, *Biochim. Biophys. Acta* **1830**, 3670–3695 (2013).
- H. W. Nützmann, A. Huang, A. Osbourn, *New Phytol.* **211**, 771–789 (2016).
- N. Yu et al., *Nucleic Acids Res.* **44**, 2255–2265 (2016).
- M. Itkin et al., *Science* **341**, 175–179 (2013).
- F. Kellner et al., *Plant J.* **82**, 680–692 (2015).
- K. Yonekura-Sakakibara et al., *Plant Cell* **20**, 2160–2176 (2008).
- Y. Okazaki, K. Saito, *Gigascience* **5**, 11 (2016).
- C. E. Harjes et al., *Science* **319**, 330–333 (2008).
- Y. Shang et al., *Science* **346**, 1084–1088 (2014).
- J. L. Magnard et al., *Science* **349**, 81–83 (2015).
- W. Lau, E. S. Sattely, *Science* **349**, 1224–1228 (2015).
- P. Verma, A. K. Mathur, A. Srivastava, A. Mathur, *Protoplasma* **249**, 255–268 (2012).
- S. Irmer et al., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4164–4169 (2015).
- B. B. Misra, S. M. Assmann, S. Chen, *Trends Plant Sci.* **19**, 637–646 (2014).
- T. Fujii et al., *Nat. Protoc.* **10**, 1445–1456 (2015).
- D. J. Cuthbertson et al., *Phytochemistry* **91**, 187–197 (2013).
- J.-F. Hu et al., *Phytochem. Anal.* **16**, 127–133 (2005).
- D. S. Dalisay et al., *J. Nat. Prod.* **78**, 1231–1242 (2015).
- P. J. Horn et al., *Plant Cell* **24**, 622–636 (2012).
- J. Soltwisch et al., *Science* **348**, 211–215 (2015).
- D. W. Etalo, R. C. H. De Vos, M. H. A. J. Joosten, R. D. Hall, *Plant Physiol.* **169**, 1424–1435 (2015).
- J. Kilcrease, A. M. Collins, R. D. Richins, J. A. Timlin, M. A. O’Connell, *Plant J.* **76**, 1074–1083 (2013).
- G. R. Littlejohn et al., *Plant Physiol.* **168**, 18–28 (2015).
- P. K. Lundquist et al., *Plant Physiol.* **158**, 1172–1192 (2012).
- M. Shumskaya, L. M. T. Bradbury, R. R. Monaco, E. T. Wurtzel, *Plant Cell* **24**, 3725–3741 (2012).
- D.-K. Ro et al., *Nature* **440**, 940–943 (2006).
- C. Zhang, X. Chen, G. Stephanopoulos, H.-P. Too, *Biotechnol. Bioeng.* **113**, 1755–1763 (2016).
- S. Galanie, K. Thodey, I. J. Trenchard, M. Filsinger Interrante, C. D. Smolke, *Science* **349**, 1095–1100 (2015).
- A. Nakagawa et al., *Nat. Commun.* **7**, 10390 (2016).
- N. J. Gallage, B. L. Møller, *Mol. Plant* **8**, 40–57 (2015).
- M. A. Costa et al., *PLOS ONE* **8**, e83169 (2013).
- J. M. Augustin, Y. Higashi, X. Feng, T. M. Kutchan, *Planta* **242**, 693–708 (2015).
- D. K. Allen, *Curr. Opin. Biotechnol.* **37**, 45–52 (2016).
- S. Okumoto, *Plant J.* **70**, 108–117 (2012).
- S. E. Bocobza, A. Aharoni, *Plant J.* **79**, 693–703 (2014).
- T. Herrmann, D. J. Patel, *Science* **287**, 820–825 (2000).
- M. N. Stojanovic, P. de Prada, D. W. Landry, *J. Am. Chem. Soc.* **123**, 4928–4931 (2001).
- J. Beltrán et al., *Nat. Chem. Biol.* **11**, 598–605 (2015).
- W. Kühlbrandt, *eLife* **3**, e03678 (2014).
- J. A. Birchler, *Trends Biotechnol.* **33**, 189–194 (2015).

ACKNOWLEDGMENTS

MALDI-MSI image: N. G. Lewis (Washington State University). Metabolic flux schematic: D. K. Allen, Danforth Plant Science Center. In vivo sensors: M. Shumskaya (Lehman College, CUNY) and E.T.W. (NIH GM081160 to E.T.W.). T.M.K. is inventor on the patent applications WO2013170265A1 and WO2015200831A3 submitted by the Danforth Plant Science Center on production of terpenes in camelina.

10.1126/science.aad2062

Plant metabolism, the diverse chemistry set of the future

Eleanore T. Wurtzel and Toni M. Kutchan

Science **353** (6305), 1232-1236.
DOI: 10.1126/science.aad2062

ARTICLE TOOLS

<http://science.sciencemag.org/content/353/6305/1232>

RELATED CONTENT

<http://science.sciencemag.org/content/sci/353/6305/1218.full>
<http://science.sciencemag.org/content/sci/353/6305/1220.full>
<http://science.sciencemag.org/content/sci/353/6305/1222.full>
<http://science.sciencemag.org/content/sci/353/6305/1225.full>
<http://science.sciencemag.org/content/sci/353/6305/1228.full>
<http://science.sciencemag.org/content/sci/353/6305/1237.full>
<http://science.sciencemag.org/content/sci/353/6305/1241.full>
<http://science.sciencemag.org/content/sci/353/6305/1187.full>
<http://science.sciencemag.org/content/sci/353/6305/1202.full>
<http://science.sciencemag.org/content/sci/353/6305/1211.full>
<http://science.sciencemag.org/content/sci/353/6305/1214.full>

REFERENCES

This article cites 40 articles, 15 of which you can access for free
<http://science.sciencemag.org/content/353/6305/1232#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2016, American Association for the Advancement of Science