


Genomics-enabled analysis of specialized metabolism in bioenergy crops: current progress and challenges

Kira Tiedge [†], Andrew Muchlinski[†], and Philipp Zerbe*

Department of Plant Biology, University of California-Davis, Davis, CA 95616, USA

*Corresponding author: E-mail: pzerbe@ucdavis.edu

[†]These authors contributed equally to this work.

Abstract

Plants produce a staggering diversity of specialized small molecule metabolites that play vital roles in mediating environmental interactions and stress adaptation. This chemical diversity derives from dynamic biosynthetic pathway networks that are often species-specific and operate under tight spatiotemporal and environmental control. A growing divide between demand and environmental challenges in food and bioenergy crop production has intensified research on these complex metabolite networks and their contribution to crop fitness. High-throughput omics technologies provide access to ever-increasing data resources for investigating plant metabolism. However, the efficiency of using such system-wide data to decode the gene and enzyme functions controlling specialized metabolism has remained limited; due largely to the recalcitrance of many plants to genetic approaches and the lack of ‘user-friendly’ biochemical tools for studying the diverse enzyme classes involved in specialized metabolism. With emphasis on terpenoid metabolism in the bioenergy crop switchgrass as an example, this review aims to illustrate current advances and challenges in the application of DNA synthesis and synthetic biology tools for accelerating the functional discovery of genes, enzymes and pathways in plant specialized metabolism. These technologies have accelerated knowledge development on the biosynthesis and physiological roles of diverse metabolite networks across many ecologically and economically important plant species and can provide resources for application to precision breeding and natural product metabolic engineering.

Key words: plant specialized metabolism; DNA synthesis; gene discovery; enzyme functional analysis; terpenoid metabolism.

1. Introduction

Meeting the needs of a growing world population while addressing urgent sustainability challenges has spurred intensive efforts in renewable energy production including biofuels (1, 2). In 2018, global biofuel production from plant-derived biomass reached an all-time high of over ~150 billion liters, and the International Energy Agency projects production to increase by 25% over the next 5 years (3). Furthermore, U.S. Department of Energy assessments predict that 990–1150 M dry tons of plant biomass feedstock could be sustainably produced in the USA by 2030 (4), a level sufficient to supply more than a

quarter of the current biofuel consumption, while maintaining food and feed production demands. First-generation agricultural feedstock crops such as maize, rapeseed, *Sorghum* and wheat, continue to serve as major resources for biofuel production (5) (Figure 1). In addition, cultivation of second-generation non-food model crops for biofuel production, including switchgrass, energy cane, poplar and *Miscanthus*, is becoming tractable to contribute to the needed production increase (5–10). But this development must address several key challenges of economic feasibility, balanced arable land use and a long-term sustainable environmental footprint (2). In the face of shifting environmental pressures that hamper plant fitness and cause

Submitted: 7 February 2020; Received (in revised form): 3 May 2020; Accepted: 25 May 2020

© The Author(s) 2020. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

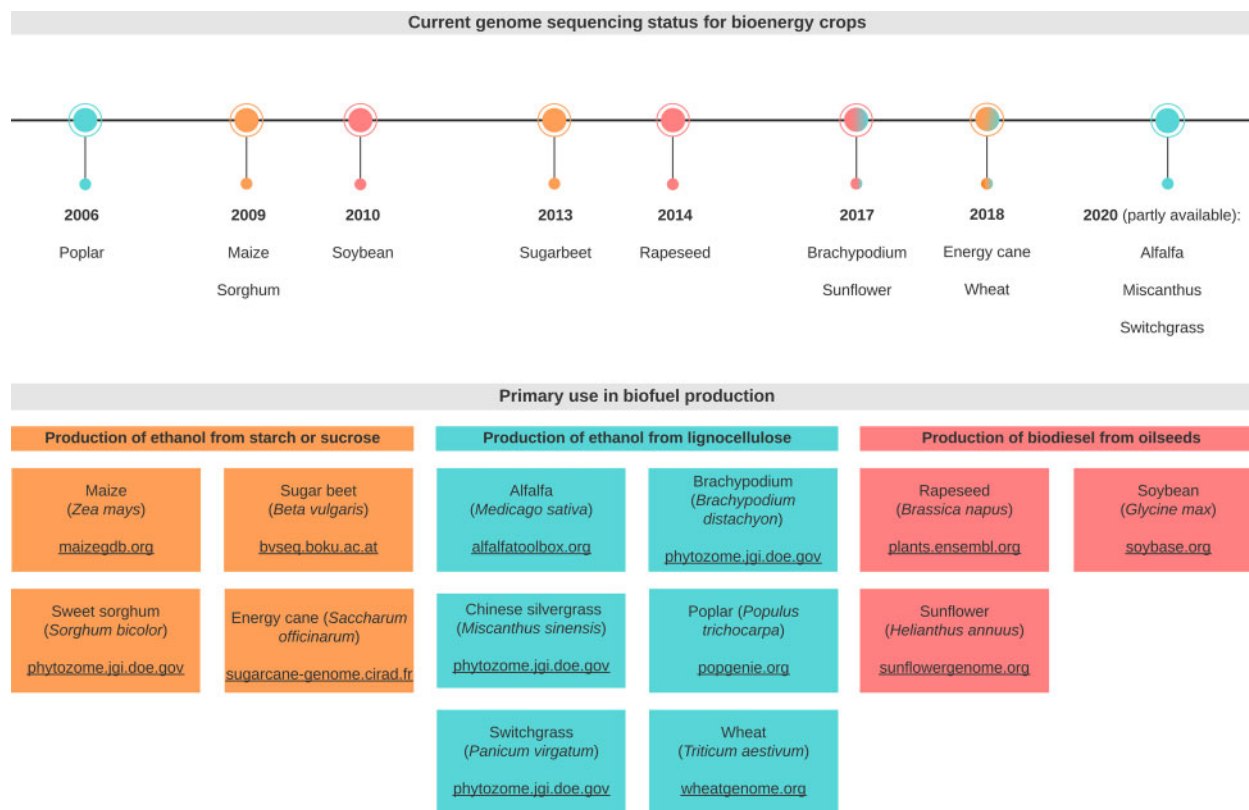


Figure 1. Overview of currently available genome resources for key bioenergy crops.

agricultural yield losses (11), an important factor in advancing biofuel crop production is a deeper understanding of the genetic and biochemical mechanisms by which crops adapt to environmental perturbations. Such knowledge can provide key resources for agricultural optimization through precision breeding and targeted gene editing to generate crops that can withstand biotic and abiotic stressors, and enable scalable crop cultivation on marginal lands that do not compete with food production (12, 13).

Extremely diverse mixtures of small molecule natural products (termed secondary or specialized metabolites), such as terpenoids, phenylpropanoids, alkaloids, cyanogenic glycosides and glucosinolates, function as core mediators of plant-environment interactions (14–17). Among these metabolites, terpenoids are the largest class, comprising ~40% of presently known plant natural products (17). Terpenoids are of essential importance to plant fitness due to their various functions as plant hormones, chemical defenses and signaling compounds in ecological interactions (14, 18). In addition, terpenoid bioproducts are of substantial economic value as food additives, therapeutics, cosmetics and biofuels (19–21). The latter includes several terpene hydrocarbons of high energy density, including farnesene, bisabolene or pinene, that are microbially manufactured as value-added alternatives to D2-biodiesel and high-performance aviation fuels (21, 22).

Their agricultural importance and broad use as bioproducts have driven a long-standing interest in developing robust strategies for the discovery and engineering of the genes, enzymes and pathways involved in plant specialized metabolism including terpenoids (12). Yet, progress in this field is hindered by the vast diversity of plant metabolism, limitations in pathway discovery tools and the often narrow taxonomic distribution of

individual genes, enzymes and products that require analysis of the natural producer rather than established model systems (23). The post-genomic era has given rise to ever-advancing tools for generating genome, transcriptome and metabolome data that, in turn, have revolutionized the investigation of the biosynthetic and regulatory machineries controlling plant metabolism and how it impacts crop phenotypes and traits (12, 23, 24). In addition, developments in DNA synthesis and synthetic biology have made great strides toward establishing advanced approaches for the functional analysis of genes and enzymes involved in plant specialized metabolism (12, 25–28). Despite these advances, throughput in genetic and biochemical techniques for gene and enzyme functional studies lag behind the rapid increase in available omics data. As exemplified by the enzyme families of terpene synthases (TPSs) and cytochrome P450 monooxygenases (P450s), key terpenoid-diversifying enzymes, the functional diversity, dynamic regulation and taxonomic restriction of terpenoid and other specialized metabolite biosynthetic genes constrain reliable computational and functional annotation and necessitate empirical biochemical analyses (29–31). The latter then rely on the development of efficient platforms for plant transformation and/or the heterologous or cell-free expression and activity analysis of different enzyme classes.

With emphasis on specialized terpenoid metabolism in the bioenergy crop switchgrass (*Panicum virgatum* L.), this article highlights current progress and challenges in gene-to-function approaches. These approaches help to elucidate the molecular and biochemical foundations that govern the biosynthesis, diversity and function of specialized metabolic networks as an important foundation to agricultural crop optimization.

2. Omics-enabled pathway discovery in plant specialized metabolism

With sequencing and assembly platforms becoming faster and less expensive, the past decade has seen the release of numerous bioenergy crop genomes, including flagship biofuel feedstock and model crops such as maize, soybean, switchgrass, *Brachypodium distachyon* and poplar (32–41) (Figure 1). These genomic repositories form a broad foundation for pathway discovery, yet the road to unambiguously annotating gene functions can be rocky. As exemplified by the family of P450s that can comprise several hundred members in an individual species (30), gene families involved in specialized metabolism are often large and diverse as a result of repeated duplication of genes with ancestral functions in general metabolism, followed by sub- or neo-functionalization of duplicated genes (42, 43). Functional redundancy, catalytic promiscuity and emergence of new enzyme functions with as little as a single active site residue substitution as demonstrated, for example, for the TPS family (42, 44), curtail the reliability of sequence-based predictions of specialized gene and enzyme activities, thus requiring genetic and/or biochemical functional verification. In addition, computational gene annotation largely relies on comparison to well-established reference genomes, including *Arabidopsis thaliana*, rice and maize, which complicates the correct annotation of specialized gene functions that may often occur in only a single plant family or species. As a consequence, intensive efforts have focused on developing modular tools and strategies for reducing the pool of candidate genes and enzymes targeted for genetic and biochemical studies (Figure 2).

Where knowledge of the biosynthetic chemical reactions and enzyme classes relevant for different specialized metabolites exists and/or can be generated using chemoinformatic algorithms (26, 45, 46), such insights can be applied to querying genome and transcriptome inventories in a pathway/product-guided fashion using BLAST-enabled homology searches against gene- or protein-specific databases and/or signature protein motifs and domains to collect an initial pool of candidates (47–51). Numerous public gene, protein and motif databases are available for such approaches, including PhytoMine (<https://phytozome.jgi.doe.gov/phytozome/begin.do>), the Plant Metabolic Network (www.plantcyc.org), the Kyoto Encyclopedia of Genes and Genomes (KEGG, www.genome.jp/kegg) or Pfam (<https://pfam.xfam.org/>). The output of database searches should further be carefully curated to triage sequences with high likelihood of representing pseudogenes or misassembled and chimeric transcripts, a common occurrence especially among gene families with high degrees of sequence similarity such as the TPS gene family (31). In addition, although less widely distributed as compared to microbial natural product pathways, genomic co-localization of biosynthetic genes can provide important insights for candidate selection where specialized metabolic pathways can be organized as gene clusters. Such biosynthetic gene clusters have been demonstrated for several compound classes, including select pathways of diterpenoid, triterpenoid, benzoxazinoid, cyanogenic glycoside and alkaloid biosynthesis (52–55). The abundance and composition of such biosynthetic gene clusters are diverse across different species, ranging from 3 to 10 or more such clusters that may span genomic regions of 35 kb up to several hundred kilobytes containing three to six or more metabolic enzymes (52). Predominantly, these clustered genes will encode for enzymes catalyzing the

first committed reactions channeling metabolites into specific specialized pathways as well as additional scaffolding or tailoring enzymes. For example, gene clusters containing multiple pairings of TPS and P450 genes have been identified in diterpenoid and triterpenoid biosynthesis of several eudicot and monocot species, including biofuel model crops such as *Sorghum* and *B. distachyon* (53, 56). Likewise, in switchgrass co-localization of specialized class II and class I diterpene synthases occurs on chromosomes 3K and 3N, potentially indicating the presence of a diterpenoid-biosynthetic gene cluster (57). The mechanisms underlying the evolution of TPS/P450 gene pairs appear to differ between eudicots and monocots and are not fully understood (53). Yet, discovery of operon-like triterpenoid gene clusters in *A. thaliana* demonstrated that, unlike horizontal gene transfer in bacteria, selection pressures likely drive clustering of defensive pathway genes in plants via gene duplications, neofunctionalization and genomic reorganization (58). Indeed, TPS/P450 pairings occur more frequently across available sequenced plant genomes than would be expected from a random distribution, indicating the potential for co-regulation and co-expression that can be exploited for gene candidate triage (53). Continued advances in genetics, map-based cloning and bioinformatics approaches are expected to reveal biosynthetic gene clusters of specialized metabolisms across a wider range of plant genomes. Development of databases representing the defining parameters of plant biosynthetic gene clusters and bioinformatics tools that enable the mining of plant genomes for physically co-localized genes encoding enzymes of specialized metabolism (59), will provide broadly applicable tools for pathway discovery for species with available high-quality genomic data. In addition, experimental approaches that use operon-like, co-regulated regions of specific chromatin signatures associated with the repression or activation of biosynthetic gene clusters, more commonly known in animals and some fungi, have been demonstrated for screening plant genomes for co-localized, co-regulated gene clusters (60).

Complementary to patterns of co-regulation, and especially where no biosynthetic gene clusters are observed, functional insight can be further inferred from comparative phylogenetic analyses. Such approaches have become increasingly useful as more enzyme functions in plant specialized metabolism are discovered. Yet, these strategies have to be viewed with caution, since activity predictions can be complicated by the functional diversity and taxonomic rather than functional relatedness of specialized enzymes such as it is frequently observed in terpenoid metabolism (24, 29, 31). Combined sequence- and phylogeny-based gene discovery have been successfully employed to identify terpenoid-metabolic gene families in several monocot crops, including rice, wheat and maize (14, 61, 62). With the increasing availability of genomics resources, these methods can now be readily applied to a broader range of crops. For example, in the allotetraploid bioenergy crop switchgrass (cv. AP13), mining of genome and transcriptome inventories against a TPS-specific protein database (31) identified an expansive TPS family of more than 70 members (57, 63). Additional comparative sequence and phylogenetic analyses suggested a substantial divergence of the switchgrass TPS family, preventing functional annotations beyond a few more common TPS functions in gibberellin metabolism or those occurring in related monocot crops (57).

Considering the typically organ- or tissue-specific localization and often tight regulation of specialized metabolic pathways in plants, correlation of such sequence-based functional

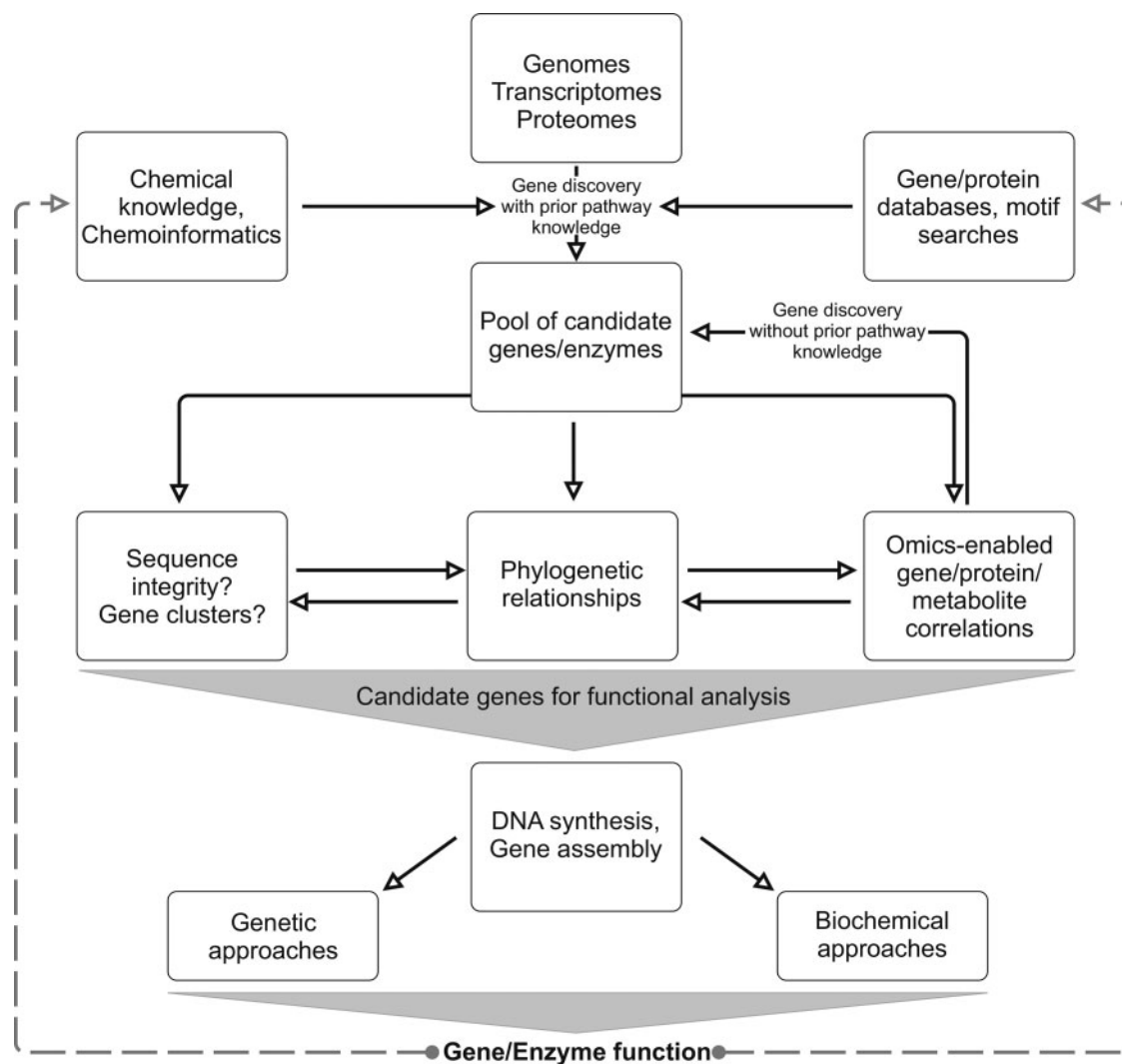


Figure 2. Scheme of possible customizable workflows for the omics-enabled discovery of gene candidates for subsequent functional characterization using genetic and/or biochemical approaches.

prediction approaches with genomics-, transcriptomics- and proteomics-enabled insights into transcript and protein abundance has become a powerful tool to identify pathway components that likely function together in individual tissues and under specific environmental stimuli (64, 65). Such approaches can further be integrated with advances in metabolomics and metabolite imaging technologies for identifying specific pathway components and their abundance in *planta* (49, 63, 66–68). For example, metabolite-guided pathway analysis recently led to the discovery of key TPS and P450 genes for the biosynthesis of potent antimicrobial kauralexin diterpenoids in maize (62). Drawing on these pathway insights, gene expression correlation studies revealed additional gene functions, the products of which were later verified by metabolite profiling of maize tissues as previously unknown dolabralexin diterpenoids (69). Beyond co-expression analysis, where suitable diversity panels and plant populations are available, Genome-Wide Association Studies (GWAS) and Quantitative Trait Locus (QTL) analyses can guide the identification of genomic regions or individual genes with targeted physiological traits (70). For example, QTL studies

led to the discovery of a steroid 5α -reductase that catalyzes desaturation reactions in kauralexin biosynthesis that were not predicted based on *a priori* pathway knowledge (62). Thus, such gene-metabolite and gene-trait correlation studies can be a facile tool for gene-guided pathway discovery where no or limited knowledge on the involved chemical reactions and corresponding biosynthetic enzyme classes exists.

Beyond pathway discovery comparative omics studies can guide the investigation of the physiological relevance of specialized metabolites and the corresponding biosynthetic pathways (23). For example, transcriptional analysis of switchgrass (cv. AP13) suggested a role of stress-inducible TPS genes in above- and below-ground herbivory defenses (63, 71). Metabolomics analyses further demonstrated the accumulation of steroidal saponin triterpenoids in several switchgrass varieties with possible functions in biotic stress defenses (72, 73). In addition, expression of specific TPSs, and corresponding accumulation terpene products, in response to drought and oxidative or ultraviolet stress, indicated possible roles of switchgrass terpenoids in abiotic stress tolerance (57).

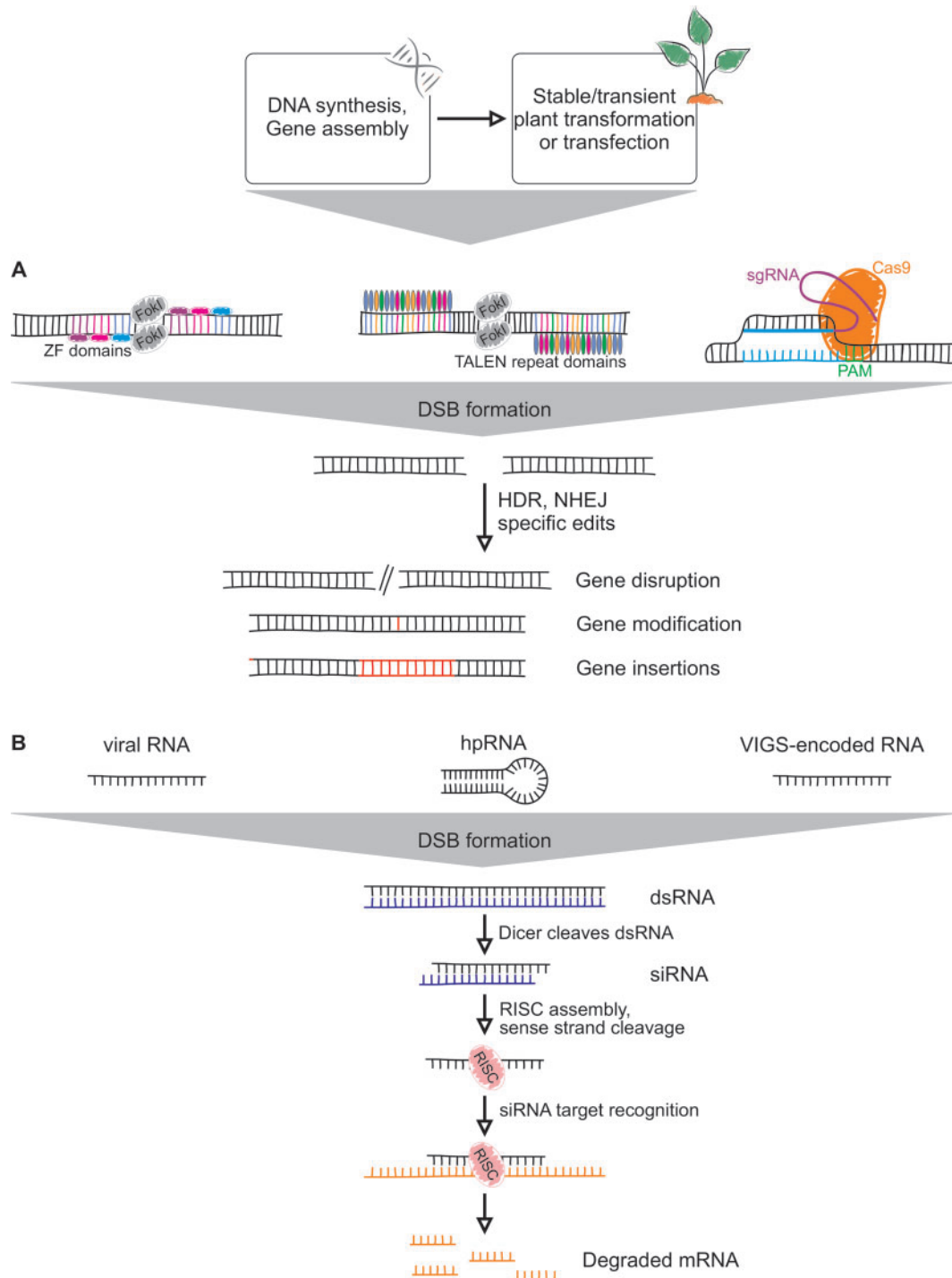


Figure 3. Schematic overview of possible genetic workflows for gene functional studies. (A) ZFN, TALEN, or CRISPR-Cas genome-editing nuclease approaches employ engineered non-specific nucleases fused to gene-specific DNA-binding domains to bind to a desired gene with high target specificity and induce double-strand breaks (DSB), followed by DNA repair via either non-homologous end joining (NHEJ) or homology-directed repair (HDR) to generate gene modifications that allow the analysis of the underlying gene function in *planta*. FokI, *Flavobacterium okeanoikoites* type IIS restriction endonuclease; PAM, protospacer adjacent motif. (B) Gene silencing can exploit the formation of small interfering RNAs (siRNAs) by dicer-mediated cleavage of double-stranded RNA (dsRNA) generated from viral RNA, VIGS-derived RNA or hairpin RNA (hpRNA) that was transcribed from a transgene of interest. The antisense strand of the resulting siRNA guides an endonuclease (RNAi-silencing complex, RISC) to enable mRNA degradation in a gene-specific manner.

3. Next-generation genetic and biochemical approaches accelerate functional studies

Downstream of candidate discovery, the verification of gene and enzyme functions requires robust genetic and biochemical

approaches (74) (Figures 3 and 4). However, often low abundance of focal transcripts in *planta*, requirement for assay optimization for different enzyme classes and plant species and limited availability and knowledge of relevant substrates remain major factors constraining the throughput of these

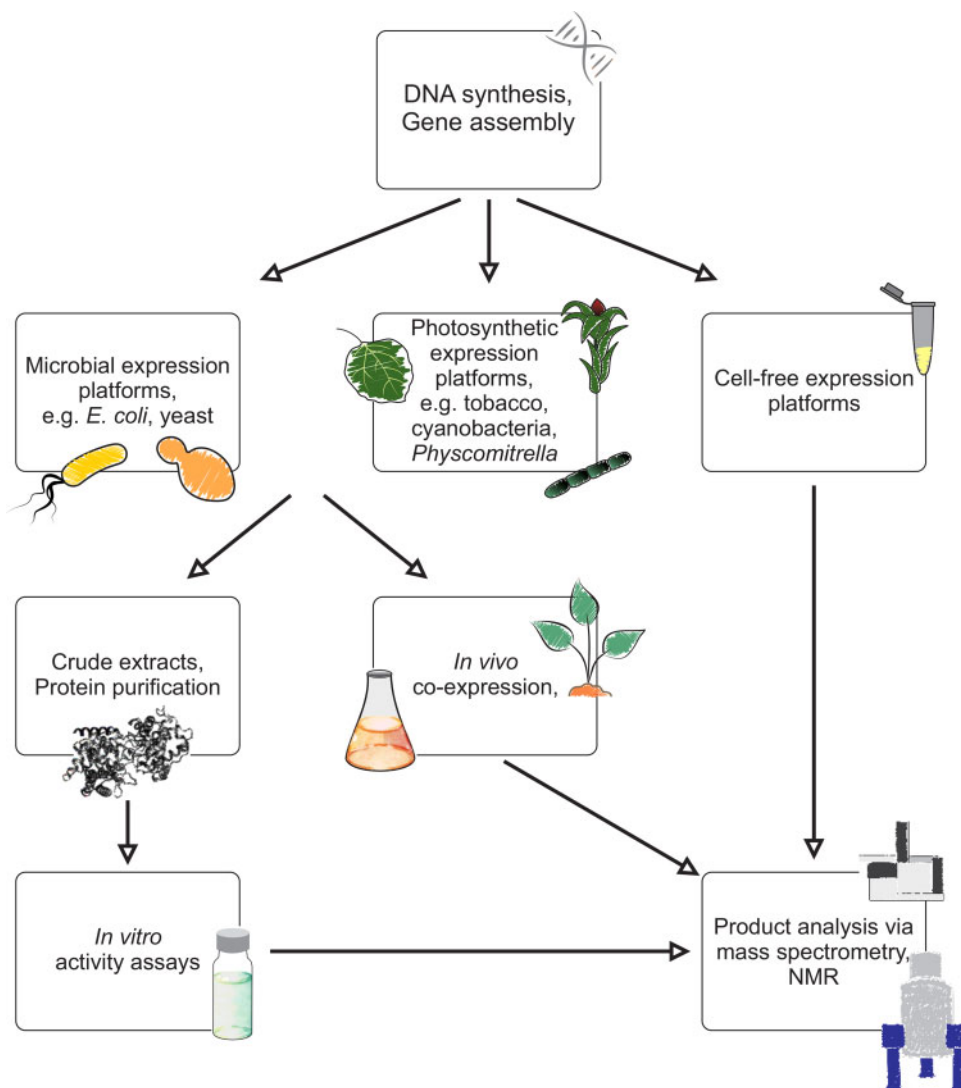


Figure 4. Schematic overview of possible biochemical workflows for enzyme functional studies. Following gene synthesis or cloning and assembly of gene constructs, numerous microbial, plant and cell-free platforms can be used for the heterologous expression of target enzymes. Enzyme functional studies can be conducted via protein purification and *in vitro* activity assays or using *in vivo* combinatorial enzyme expression strategies. Regardless of the chosen assay system, enzyme product identification typically involves liquid or gas chromatography-mass spectrometry (LC- or GC-MS) and/or nuclear magnetic resonance (NMR) analysis.

techniques. Alongside the enhanced quality and completeness of genome and transcriptome assemblies, DNA synthesis has become increasingly cost-effective for the rapid synthesis and modification of genes for functional analysis (75). Following the first gene synthesis of a yeast alanine tRNA in the early 1970s (76), today's DNA synthesis platforms allow for the rapid, large-scale generation of even complex sequences at costs below 0.1 USD per base pair. In addition, DNA synthesis readily enables access to transcripts that are difficult to obtain via amplification, as well as gene modifications including codon-optimization, sequence modification and gene assembly, making genes substantially more amenable to a broader range of genetic and biochemical approaches (75, 77, 78).

Taking advantage of these DNA synthesis capacities, targeted gene modification/editing techniques, including nuclease-enabled gene modifications such as Transcription Activator-Like Effector Nucleases (TALEN), Zinc-Finger Nucleases (ZFN) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) approaches, as well as RNA-guided gene silencing techniques such as RNAi and virus-induced gene silencing (VIGS),

can be used to identify gene functions in *planta* and without the need for *a priori* knowledge of pathway substrates and products (79–81) (Figure 3). Continued advances in the applicability of the CRISPR/Cas9 system, for example, by improving CRISPR/Cas9 catalytic specificity and efficiency by applying new generations of viral vectors (82), offers unprecedented opportunities for gene and genome editing. Furthermore, targeting of different genomic sites allows for simultaneously upregulating genes within a pathway while repressing others, resulting in modified genomes as well as transcriptomes (83). However, the complexity and high ploidy of many plant genomes, along with limitations in plant transformation and regeneration, continue to be a key bottleneck for the routine application of genetic gene-to-function studies; although, they are expected to continue to advance in coming years. In the case of switchgrass, variable typically high ploidy and self-incompatibility among diverse ecotypes poses a substantial impediment to genetic studies and crop optimization (84, 85). However, genetic approaches have been successfully employed in several instances, e.g. Xu et al. (86) employed an RNAi strategy to identify a 4-coumarate: CoA

ligase relevant for reducing lignin formation. CRISPR/Cas9-enabled gene editing also has been successfully applied to switchgrass as exemplified by knock-down of a UDP-Arabinomutase gene that resulted in the accumulation of cell wall lignin (87). Likewise, CRISPR/Cas9-mediated loss of function of switchgrass 4-coumarate: coenzyme A ligase 1 (Pv4CL1) resulted in a reduced lignin content (88). Along with such gene-specific studies, the development of mutant collections, such as the *B. distachyon* T-DNA insertion collection and expansive maize mutant lines and diversity panels provide valuable resources expedite gene and pathway discovery (89–92). Continued optimization of plant transformation protocols will undoubtedly improve the applicability of genetic tools in switchgrass and other bioenergy crops (27, 93–95). For example, advances in *de novo* meristem induction techniques as demonstrated in select dicotyledonous species will help to overcome limitations in regenerating transgenic plants from tissue culture (27, 94). Complementary to stable plant transformation approaches, transient gene modifications can be applied to examine gene functions *in planta*. VIGS has been established as a powerful tool for transient gene function studies, since it is independent of plant regeneration protocols and defined genetic backgrounds (96–99). Availability of different VIGS vectors with a broad host range and high silencing efficiency and duration have proven useful for gene function studies associated with stress resistance in several crop plants (99–101).

Advances in synthetic biology and natural product metabolic engineering tools now also enable larger-scale protein biochemical analyses via improved protein expression and combinatorial pathway reconstruction in heterologous host systems, including yeast (*Saccharomyces cerevisiae*), *Escherichia coli*, *Nicotiana benthamiana*, *Physcomitrella patens* and cyanobacteria (21, 25, 102, 103) (Figure 4). Simplified gene modifications through DNA synthesis, such as codon optimization and altered signal peptides and compartmentalization, as well as advanced chassis for custom gene and promoter assemblies, have helped to overcome limitations in protein expression due to, for example, misfolding and feedback inhibition caused by lack of regulatory control in the heterologous host (103–106). However, while numerous robust protocols for heterologous expression, protein purification and *in vitro* assays have been established for many enzymes, insufficient soluble protein expression, low activity after translation and lack of relevant substrates remain key challenges. Advances in gene stacking strategies have become an invaluable tool for elucidating both known and novel biochemical pathways by means of combinatorial protein expression assays. This process has traditionally been reliant on introducing combinations of individual plasmids containing different pathway enzymes, which can present challenges due to high antibiotic loads and variable expression levels (107, 108). However, tools for simultaneous assembly of multiple target DNAs such as GoldenGate and related cloning methods have enabled more efficient reconstruction of pathways and enzyme combinations with additionally flexibly for tailoring expression levels using specific promoter, terminator and untranslated region (UTR) combinations (109, 110). A multiple chassis approach can also be used to facilitate discovery of natural products using non-conventional strains (103), in addition to a wider range of microbial and plant host platforms can be explored to optimize functional analyses (Figure 4). Microbial expression platforms offer the benefit of scalability and a wealth of molecular tools for gene and host optimization (25, 106, 111), but sufficient

protein expression often remains a limiting factor for some enzymes such as the family of P450s. Additionally, plant-based systems have seen substantial advancement. In particular, *Agrobacterium*-mediated transient expression in *N. benthamiana* has become a routine system for enzyme functional studies, as it does not require feeding of carbon precursors, offers the natural compartmentalization for expressing plant proteins and naturally produces many key substrates for specialized metabolite biosynthesis (25, 112, 113). Current limitations of this plant expression system include an often higher metabolic background as compared to microbial systems, lack of endogenous precursor pathways for some natural product classes, and presence of endogenous enzymes and pathways that can interfere with introduced pathways through substrate competition and/or undesired metabolization of target compounds such as degradation or glycosylation (114). Bryophytes, especially *P. patens*, are also emerging as chassis for enzyme functional studies and pathway engineering, due to the availability of homologous recombination for precise genome-engineering and scalability in bioreactor platforms, yet require continued development of efficient synthetic biology tools along with a deeper metabolic understanding of these systems to enable more efficient and robust usage (115).

Next to *in vivo* expression systems, great strides have been made in developing cell-free expression systems for enzyme and pathway analysis and engineering, especially where heterologous expression is a limiting factor (116, 117). While membrane-associated proteins like P450s have been successfully produced in cell-free expression systems (118), misfolding and lack of activity after translation remain major challenges especially for membrane-bound proteins. The above approaches allow for the co-expression of genes of interest in different combinations to enable the efficient analysis of complex, modular pathways of plant specialized metabolism (25, 119–121). For example, combinatorial expression of identified TPS and P450 candidates in *E. coli* and *N. benthamiana* enabled the rapid discovery of modular terpenoid pathway networks in rice, wheat, maize and switchgrass (14, 57, 62, 69). Especially, where previously unknown enzyme functions are identified, the resulting products can, in turn, guide the discovery of metabolites and additional pathway genes *in planta* (69). In addition, pathway knowledge and new gene/protein functions deriving from the approaches outlined above, continuously expand the databases for improving the precision and efficiency of gene discovery and computational functional annotation (Figure 2).

4. Future directions

The biodiversity of specialized metabolism forms the foundation of the ability of plants to interact with and adapt to their environment with direct impact on plant fitness and agriculture. With an estimated number of more than 1 million plant compounds (122), this chemical repository offers an immense and largely untapped natural source to devise innovative strategies for crop trait discovery and improvement, as well as biomaterial manufacturing plant-derived biofuels, medicines and other high-value natural products (12, 13, 21, 123). With our ever-increasing and cost-effective capacity for reading and writing nucleic acid sequences, unparalleled opportunities arise to unlock this biochemical potential. The examples discussed above highlight the power of integrating system-wide omics analyses with DNA synthesis and next-generation gene functional tools to decode

the complex biosynthetic machinery controlling specialized metabolic networks in almost any organism. In this context, the continued development of higher throughput and more broadly applicable biochemical methods will be critical to accelerate the functional analysis of gene and protein candidates that result from genomics-enabled pathway discovery approaches. Advances in gene and enzyme modification through DNA synthesis and gene assembly can be integrated with new and improved microbial, plant and cell-free expression platforms to ultimately establish ‘user-friendly’ functional assay systems that can be readily applied to a broader range of enzymes and pathways. Further optimization of metabolomics, nuclear magnetic resonance and imaging tools can also be anticipated to continue to improve. This will simplify the often still cost- and time-intensive identification of enzyme products, and to precisely analyze the correlation of *in vitro* enzyme products with plant-, tissue- and cell-specific metabolite profiles (66, 67, 124). Likewise, gene editing and other targeted genetic tools must be further advanced to streamline the empirical verification of gene functions even in complex plant systems such as switchgrass (12, 123, 125). Alongside such approaches, facile techniques for transforming a broader range of plant species are urgently needed to apply genetic gene-function studies directly to the species where metabolites of interest are produced, and to improve the economic feasibility of optimizing crop traits (27). Driven by these transformative technologies, the rapid discovery of specialized genes, enzymes and pathways provides essential target genes for precision breeding and crop engineering. Furthermore, a comprehensive gene and enzyme catalog that can be translated into microbial and plant metabolic engineering platforms to produce high-value biofuels and other natural products, would be of societal benefit (21, 126, 127).

Acknowledgments

We appreciate the discussion of the manuscript with members of the Zerbe Lab at the University of California, Davis. The references below represent a large body of literature on the topics discussed in this article. We apologize for any relevant articles not cited in this work due to space constraints.

Funding

The authors’ work has been supported with funds from U.S. Department of Energy (DOE) Early Career Research Program (DE-SC0019178, to P.Z.); DOE Joint Genome Institute (JGI) DNA Synthesis Science Program (grant #2568, to P.Z.); German Research Foundation (DFG) Research Fellowship (TI 1075/1-1, to K.T.).

Conflict of interest statement. None declared.

References

- Perin, G. and Jones, P.R. (2019) Economic feasibility and long-term sustainability criteria on the path to enable a transition from fossil fuels to biofuels. *Curr. Opin. Biotechnol.*, 57, 175–182.
- Dunn, J.B., Bidy, M., Jones, S., Cai, H., Benavides, P.T., Markham, J., Tao, L., Tan, E., Kinchin, C., Davis, R. et al. (2018) Environmental, economic, and scalability considerations and trends of selected fuel economy-enhancing biomass-derived blendstocks. *ACS Sustain. Chem. Eng.*, 6, 561–569.
- Renewables. (2019). *Market Analysis and Forecast from 2019 to 2024*, International Energy Agency, Paris, 2019. <https://www.iea.org/reports/renewables-2019> (29 April 2020, date last accessed).
- U.S. Department of Energy. (2016) 2016 Billion-Ton Report: Advancing Domestic Resources for a Thriving Bioeconomy, Volume 1: Economic Availability of Feedstocks. M. H. Langholtz, B. J. Stokes, and L. M. Eaton (Leads), ORNL/TM-2016/160. Oak Ridge National Laboratory, Oak Ridge, TN, p. 448. <http://energy.gov/eere/bioenergy/2016-billion-ton-report> (29 April 2020, date last accessed).
- Kluts, I.N., Brinkman, M.L.J., de Jong, S.A. and Junginger, H.M. (2019) Biomass resources: agriculture. *Adv. Biochem. Eng. Biotechnol.*, 166, 13–26.
- Dahlberg, J. (2019) The role of sorghum in renewables and biofuels. *Methods Mol. Biol.*, 1931, 269–277.
- Tubeileh, A., Rennie, T.J. and Goss, M.J. (2016) A review on biomass production from C4 grasses: yield and quality for end-use. *Curr. Opin. Plant Biol.*, 31, 172–180.
- Biswal, A.K., Atmodjo, M.A., Li, M., Baxter, H.L., Yoo, C.G., Pu, Y., Lee, Y.-C., Mazarei, M., Black, I.M., Zhang, J.-Y. et al. (2018) Sugar release and growth of biofuel crops are improved by downregulation of pectin biosynthesis. *Nat. Biotechnol.*, 36, 249–257.
- Bouton, J.H. (2007) Molecular breeding of switchgrass for use as a biofuel crop. *Curr. Opin. Genet. Dev.*, 17, 553–558.
- Keshwani, D.R. and Cheng, J.J. (2009) Switchgrass for bioethanol and other value-added applications: a review. *Bioresour. Technol.*, 100, 1515–1523.
- Ray, D.K., Gerber, J.S., MacDonald, G.K. and West, P.C. (2015) Climate variation explains a third of global crop yield variability. *Nat. Commun.*, 6, 5989.
- Wurtzel, E.T. and Kutchan, T.M. (2016) Plant metabolism, the diverse chemistry set of the future. *Science*, 353, 1232–1236.
- Jez, J.M., Lee, S.G. and Sherr, A.M. (2016) The next green movement: plant biology for the environment and sustainability. *Science*, 353, 1241–1244.
- Schmelz, E.A., Huffaker, A., Sims, J.W., Christensen, S.A., Lu, X., Okada, K. and Peters, R.J. (2014) Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins. *Plant J*, 79, 659–678.
- Blažević, I., Montaut, S., Burčul, F., Olsen, C.E., Burow, M., Rollin, P. and Agerbirk, N. (2020) Glucosinolate structural diversity, identification, chemical synthesis and metabolism in plants. *Phytochemistry*, 169, 112100.
- Weng, J.-K., Philippe, R.N. and Noel, J.P. (2012) The rise of chemodiversity in plants. *Science*, 336, 1667–1670.
- Chassagne, F., Cabanac, G., Hubert, G., David, B. and Marti, G. (2019) The landscape of natural product diversity and their pharmacological relevance from a focus on the Dictionary of Natural Products. *Phytochemistry Rev*, 18, 601–622.
- Tholl, D. (2015) Biosynthesis and biological functions of terpenoids in plants. *Adv. Biochem. Eng. Biotechnol.*, 148, 63–106.
- Mafu, S. and Zerbe, P. (2018) Plant diterpenoid metabolism for manufacturing the biopharmaceuticals of tomorrow: prospects and challenges. *Phytochem. Rev.*, 17, 113–130.
- Philippe, R.N., De Mey, M., Anderson, J. and Ajikumar, P.K. (2014) Biotechnological production of natural zero-calorie sweeteners. *Curr. Opin. Biotechnol.*, 26, 155–161.
- d’Espaux, L., Mendez-Perez, D., Li, R. and Keasling, J.D. (2015) Synthetic biology for microbial production of lipid-based biofuels. *Curr. Opin. Chem. Biol.*, 29, 58–65.

22. Sarria,S., Wong,B., Martín,H.G., Keasling,J.D. and Peralta-Yahya,P. (2014) Microbial synthesis of pinene. *ACS Syn. Biol.*, 3, 466–475.
23. Schillmiller,A.L., Pichersky,E. and Last,R.L. (2012) Taming the hydra of specialized metabolism: how systems biology and comparative approaches are revolutionizing plant biochemistry. *Curr. Opin. Plant. Biol.*, 15, 338–344.
24. Zerbe,P. and Bohlmann,J. (2015) Plant diterpene synthases: exploring modularity and metabolic diversity for bioengineering. *Trends Biotechnol.*, 33, 419–428.
25. Kitaoka,N., Lu,X., Yang,B. and Peters,R.J. (2015) The application of synthetic biology to elucidation of plant mono-, sesqui-, and diterpenoid metabolism. *Mol. Plant*, 8, 6–16.
26. Jeffries,J.G., Seaver,S.M.D., Faria,J.P. and Henry,C.S. (2018) A pathway for every product? Tools to discover and design plant metabolism. *Plant Sci.*, 273, 61–70.
27. Altpeter,F., Springer,N.M., Bartley,L.E., Blechl,A.E., Brutnell,T.P., Citovsky,V., Conrad,L.J., Gelvin,S.B., Jackson,D.P., Kausch,A.P. et al. (2016) Advancing crop transformation in the era of genome editing. *Plant Cell*, 28, 1510–1520.
28. O'Connor,S.E. (2015) Engineering of secondary metabolism. *Annu. Rev. Genet.*, 49, 71–94.
29. Chen,F., Tholl,D., Bohlmann,J. and Pichersky,E. (2011) 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.*, 66, 212–229.
30. Nelson,D. and Werck-Reichhart,D. (2011) A P450-centric view of plant evolution. *Plant J.*, 66, 194–211.
31. Zerbe,P., Hamberger,B., Yuen,M.M.S., Chiang,A., Sandhu,H.K., Madilao,L.L., Nguyen,A., Hamberger,B., Bach,S.S., Bohlmann,J. et al. (2013) Gene discovery of modular diterpene metabolism in nonmodel systems. *Plant Physiol.*, 162, 1073–1091.
32. Paterson,A.H., Bowers,J.E., Bruggmann,R., Dubchak,I., Grimwood,J., Gundlach,H., Haberer,G., Hellsten,U., Mitros,T., Poliakov,A. et al. (2009) The *Sorghum bicolor* genome and the diversification of grasses. *Nature*, 457, 551–556.
33. Schnable,P.S., Ware,D., Fulton,R.S., Stein,J.C., Wei,F., Pasternak,S., Liang,C., Zhang,J., Fulton,L., Graves,T.A. et al. (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science*, 326, 1112–1115.
34. Bennetzen,J.L., Schmutz,J., Wang,H., Percifield,R., Hawkins,J., Pontaroli,A.C., Estep,M., Feng,L., Vaughn,J.N., Grimwood,J. et al. (2012) Reference genome sequence of the model plant *Setaria*. *Nat. Biotechnol.*, 30, 555–561.
35. Schmutz,J., Cannon,S.B., Schlueter,J., Ma,J., Mitros,T., Nelson,W., Hyten,D.L., Song,Q., Thelen,J.J., Cheng,J. et al. (2010) Genome sequence of the palaeopolyploid soybean. *Nature*, 463, 178–183.
36. Chalhoub,B., Denoed,F., Liu,S., Parkin,I.A.P., Tang,H., Wang,X., Chiquet,J., Belcram,H., Tong,C., Samans,B. et al. (2014) Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science*, 345, 950–953.
37. Garsmeur,O., Droc,G., Antonise,R., Grimwood,J., Potier,B., Aitken,K., Jenkins,J., Martin,G., Charron,C., Hervouet,C. et al. (2018) A mosaic monoploid reference sequence for the highly complex genome of sugarcane. *Nature Commun.*, 9, 2638.
38. The International Wheat Genome Sequencing Consortium (IWGSC); Appels,R., Eversole,K., Stein,N., Feuillet,C., Keller,B., Rogers,J., Pozniak,C.J., Choulet,F., Distelfeld,A., Poland,J. et al. (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, 361, eaar7191.
39. Ma,X.-F., Jensen,E., Alexandrov,N., Troukhan,M., Zhang,L., Thomas-Jones,S., Farrar,K., Clifton-Brown,J., Donnison,I., Swaller,T. et al. (2012) High resolution genetic mapping by genome sequencing reveals genome duplication and tetraploid genetic structure of the diploid *Miscanthus sinensis*. *PLoS One*, 7, e33821.
40. International Brachypodium Initiative (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature*, 463, 763–768.
41. Tuskan,G.A., Difazio,S., Jansson,S., Bohlmann,J., Grigoriev,I., Hellsten,U., Putnam,N., Ralph,S., Rombauts,S., Salamov,A. et al. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, 313, 1596–1604.
42. Zi,J., Mafu,S. and Peters,R.J. (2014) To gibberellins and beyond! Surveying the evolution of (di)terpenoid metabolism. *Annu. Rev. Plant Biol.*, 65, 259–286.
43. Banerjee,A. and Hamberger,B. (2018) P450s controlling metabolic bifurcations in plant terpene specialized metabolism. *Phytochem. Rev.*, 17, 81–111.
44. Karunanithi,P.S. and Zerbe,P. (2019) Terpene synthases as metabolic gatekeepers in the evolution of plant terpenoid chemical diversity. *Front. Plant Sci.*, 10, 1166.
45. Anarat-Cappillino,G. and Sattely,E.S. (2014) The chemical logic of plant natural product biosynthesis. *Curr. Opin. Plant Biol.*, 19, 51–58.
46. Johnson,S.R., Bhat,W.W., Bibik,J., Turmo,A., Hamberger,B. and Hamberger,B.; Evolutionary Mint Genomics Consortium (2019) A database-driven approach identifies additional diterpene synthase activities in the mint family (Lamiaceae). *J. Biol. Chem.*, 294, 1349–1362.
47. Xiao,M., Zhang,Y., Chen,X., Lee,E.-J., Barber,C.J.S., Chakrabarty,R., Desgagné-Penix,I., Haslam,T.M., Kim,Y.-B., Liu,E. et al. (2013) Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. *J. Biotechnol.*, 166, 122–134.
48. Sato,H., Saito,K. and Yamazaki,M. (2019) Acceleration of mechanistic investigation of plant secondary metabolism based on computational chemistry. *Front. Plant Sci.*, 10, 802.
49. Saito,K. (2013) Phytochemical genomics—a new trend. *Curr. Opin. Plant Biol.*, 16, 373–380.
50. Weber,T. and Kim,H.U. (2016) The secondary metabolite bioinformatics portal: computational tools to facilitate synthetic biology of secondary metabolite production. *Synth. Syst. Biotechnol.*, 1, 69–79.
51. Arkin,A.P., Cottingham,R.W., Henry,C.S., Harris,N.L., Stevens,R.L., Maslov,S., Dehal,P., Ware,D., Perez,F., Canon,S. et al. (2018) KBase: the United States department of energy systems biology knowledgebase. *Nat. Biotechnol.*, 36, 566–569.
52. Nützmann,H.-W., Huang,A. and Osbourn,A. (2016) Plant metabolic clusters—from genetics to genomics. *New Phytol.*, 211, 771–789.
53. Boutanaev,A.M., Moses,T., Zi,J., Nelson,D.R., Mugford,S.T., Peters,R.J. and Osbourn,A. (2015) Investigation of terpene diversification across multiple sequenced plant genomes. *Proc. Natl. Acad. Sci. U S A*, 112, E81–E88.
54. Guo,L., Winzer,T., Yang,X., Li,Y., Ning,Z., He,Z., Teodor,R., Lu,Y., Bowser,T.A., Graham,I.A. et al. (2018) The opium poppy genome and morphinan production. *Science*, 362, 343–347.
55. Takos,A.M., Knudsen,C., Lai,D., Kannangara,R., Mikkelsen,L., Motawia,M.S., Olsen,C.E., Sato,S., Tabata,S.,

- Jørgensen, K. et al. (2011) Genomic clustering of cyanogenic glucoside biosynthetic genes aids their identification in *Lotus japonicus* and suggests the repeated evolution of this chemical defence pathway. *Plant J.*, 68, 273–286.
56. Wu, Y., Hillwig, M.L., Wang, Q. and Peters, R.J. (2011) Parsing a multifunctional biosynthetic gene cluster from rice: biochemical characterization of CYP71Z6 & 7. *FEBS Lett.*, 585, 3446–3451.
57. Pelot, K.A., Chen, R., Hagelthorn, D.M., Young, C.A., Addison, J.B., Muchlinski, A., Tholl, D. and Zerbe, P. (2018) Functional diversity of diterpene synthases in the biofuel crop switchgrass. *Plant Physiol.*, 178, 54–71.
58. Field, B. and Osbourn, A.E. (2008) Metabolic diversification-independent assembly of operon-like gene clusters in different plants. *Science*, 320, 543–547.
59. Töpfer, N., Fuchs, L.-M. and Aharoni, A. (2017) The PhytoClust tool for metabolic gene clusters discovery in plant genomes. *Nucl. Acids Res.*, 45, 7049–7063.
60. Yu, N., Nützmann, H.-W., MacDonald, J.T., Moore, B., Field, B., Berriri, S., Trick, M., Rosser, S.J., Kumar, S.V., Freemont, P.S. et al. (2016) Delineation of metabolic gene clusters in plant genomes by chromatin signatures. *Nucl. Acids Res.*, 44, 2255–2265.
61. Xu, M., Ross Wilderman, P., Morrone, D., Xu, J., Roy, A., Margis-Pinheiro, M., Upadhyaya, N.M., Coates, R.M. and Peters, R.J. (2007) Functional characterization of the rice kaurene synthase-like gene family. *Phytochemistry*, 68, 312–326.
62. Ding, Y., Murphy, K.M., Poretsky, E., Mafu, S., Yang, B., Char, S.N., Christensen, S.A., Saldivar, E., Wu, M., Wang, Q. et al. (2019) Multiple genes recruited from hormone pathways partition maize diterpenoid defences. *Nat. Plants*, 5, 1043–1056.
63. Muchlinski, A., Chen, X., Lovell, J.T., Köllner, T.G., Pelot, K.A., Zerbe, P., Ruggiero, M., Callaway, L., Laliberte, S., Chen, F. et al. Biosynthesis and emission of stress-induced volatile terpenes in roots and leaves of switchgrass (*Panicum virgatum* L.). *Front. Plant Sci.*, 10, 20191144.
64. Schaefer, R.J., Michno, J.-M. and Myers, C.L. (2017) Unraveling gene function in agricultural species using gene co-expression networks. *Biochim. Biophys. Acta Gene Regul. Mech.*, 1860, 53–63.
65. Higashi, Y. and Saito, K. (2013) Network analysis for gene discovery in plant-specialized metabolism. *Plant Cell Environ.*, 36, 1597–1606.
66. Kind, T., Tsugawa, H., Cajka, T., Ma, Y., Lai, Z., Mehta, S.S., Wohlgemuth, G., Barupal, D.K., Showalter, M.R., Arita, M. et al. (2018) Identification of small molecules using accurate mass MS/MS search. *Mass Spectrom. Rev.*, 37, 513–532.
67. Sumner, L.W., Lei, Z., Nikolau, B.J. and Saito, K. (2015) Modern plant metabolomics: advanced natural product gene discoveries, improved technologies, and future prospects. *Nat. Prod. Rep.*, 32, 212–229.
68. de Raad, M., Fischer, C.R. and Northen, T.R. (2016) High-throughput platforms for metabolomics. *Curr. Opin. Chem. Biol.*, 30, 7–13.
69. Mafu, S., Ding, Y., Murphy, K.M., Yaacoobi, O., Addison, J.B., Wang, Q., Shen, Z., Briggs, S.P., Bohlmann, J., Castro-Falcon, G. et al. (2018) Discovery, biosynthesis and stress-related accumulation of dolabradiene-derived defenses in maize. *Plant Physiol.*, 176, 2677–2690.
70. Korte, A. and Farlow, A. (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*, 9, 29.
71. Donze-Reiner, T., Palmer, N.A., Scully, E.D., Prochaska, T.J., Koch, K.G., Heng-Moss, T., Bradshaw, J.D., Twigg, P., Amundsen, K., Sattler, S.E. et al. (2017) Transcriptional analysis of defense mechanisms in upland tetraploid switchgrass to greenbugs. *BMC Plant Biol.*, 17, 46.
72. Huhman, D.V., Berhow, M.A. and Sumner, L.W. (2005) Quantification of saponins in aerial and subterranean tissues of *Medicago truncatula*. *J. Agri. Food Chem.*, 53, 1914–1920.
73. Lee, S.T., Mitchell, R.B., Wang, Z., Heiss, C., Gardner, D.R. and Azadi, P. (2009) Isolation, characterization, and quantification of steroidal saponins in switchgrass (*Panicum virgatum* L.). *J. Agri. Food Chem.*, 57, 2599–2604.
74. Pickar-Oliver, A. and Gersbach, C.A. (2019) The next generation of CRISPR-Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.*, 20, 490–507.
75. Hughes, R.A. and Ellington, A.D. (2017) Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. *Cold Spring Harb. Perspect. Biol.*, 9, a023812.
76. Khorana, H.G., Agarwal, K.L., Büchi, H., Caruthers, M.H., Gupta, N.K., Klbppe, K., Kumar, A., Ohtsuka, E., Rajbhandary, U.L., van de Sande, J.H. et al. (1972) Total synthesis of the structural gene for an alanine transfer ribonucleic acid from yeast. *J. Mol. Biol.*, 72, 209–217.
77. Kosuri, S. and Church, G.M. (2014) Large-scale de novo DNA synthesis: technologies and applications. *Nat. Methods*, 11, 499–507.
78. Kunjapur, A.M., Pfingst, P., Thompson, N.C. (2018) Gene synthesis allows biologists to source genes from farther away in the tree of life. *Nat. Commun.*, 9, 1–11.
79. Doudna, J.A. and Charpentier, E. (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*, 346, 1258096.
80. Shan, Q. and Voytas, D.F. (2018) Editing plant genes one base at a time. *Nat. Plants*, 4, 412–413.
81. Baltés, N.J., Gil-Humanes, J. and Voytas, D.F. (2017) Genome engineering and agriculture: opportunities and challenges. *Prog. Mol. Biol. Transl. Sci.*, 149, 1–26.
82. Cody, W.B. and Scholthof, H.B. (2019) Plant virus vectors 3.0: transitioning into synthetic genomics. *Annu. Rev. Phytopathol.*, 57, 211–230.
83. Puchta, H. (2016) Using CRISPR/Cas in three dimensions: towards synthetic plant genomes, transcriptomes and epigenomes. *Plant J.*, 87, 5–15.
84. Parrish, D.J. and Fike, J.H. (2005) The biology and agronomy of switchgrass for biofuels. *Crit. Rev. Plant Sci.*, 24, 423–459.
85. McLaughlin, S.B. and Adams Kszos, L. (2005) Development of switchgrass (*Panicum virgatum*) as a bioenergy feedstock in the United States. *Biomass Bioenergy*, 28, 515–535.
86. Xu, B., Escamilla-Treviño, L.L., Sathitsuksanoh, N., Shen, Z., Shen, H., Percival Zhang, Y.-H., Dixon, R.A. and Zhao, B. (2011) Silencing of 4-coumarate: coenzyme A ligase in switchgrass leads to reduced lignin content and improved fermentable sugar yields for biofuel production. *New Phytol.*, 192, 611–625.
87. Willis, J.D., Smith, J.A., Mazarei, M., Zhang, J.-Y., Turner, G.B., Decker, S.R., Sykes, R.W., Poovaiah, C.R., Baxter, H.L., Mann, D.G.J. et al. (2016) Downregulation of the UDP-arabinomutase gene in switchgrass (*Panicum virgatum* L.) results in increased cell wall lignin while reducing arabinose-glycans. *Front. Plant Sci.*, 7, 1580.
88. Park, J.-J., Yoo, C.G., Flanagan, A., Pu, Y., Debnath, S., Ge, Y., Ragauskas, A.J. and Wang, Z.-Y. (2017) Defined tetra-allelic

- gene disruption of the 4-coumarate: coenzyme A ligase 1 (Pv4CL1) gene by CRISPR/Cas9 in switchgrass results in lignin reduction and improved sugar release. *Biotechnol. Biofuels*, 10, 284.
89. Nannas,N.J. and Dawe,R.K. (2015) Genetic and genomic toolbox of *Zea mays*. *Genetics*, 199, 655–669.
 90. Hsia,M.M., O'Malley,R., Cartwright,A., Nieu,R., Gordon,S.P., Kelly,S., Williams,T.G., Wood,D.F., Zhao,Y., Bragg,J. et al. (2017) Sequencing and functional validation of the JGI *Brachypodium distachyon* T-DNA collection. *Plant J.*, 91, 361–370.
 91. Yu,J., Holland,J.B., McMullen,M.D. and Buckler,E.S. (2008) Genetic design and statistical power of nested association mapping in maize. *Genetics*, 178, 539–551.
 92. Kremling,K.A.G., Chen,S.-Y., Su,M.-H., Lepak,N.K., Romay,M.C., Swarts,K.L., Lu,F., Lorant,A., Bradbury,P.J., Buckler,E.S. et al. (2018) Dysregulation of expression correlates with rare-allele burden and fitness loss in maize. *Nature*, 555, 520–523.
 93. Ali,Z., Abul-Faraj,A., Li,L., Ghosh,N., Piatek,M., Mahjoub,A., Aouida,M., Piatek,A., Baltus,N.J., Voytas,D.F. et al. (2015) Efficient virus-mediated genome editing in plants using the CRISPR/Cas9 system. *Mol. Plant*, 8, 1288–1291.
 94. Maher,M.F., Nasti,R.A., Vollbrecht,M., Starker,C.G., Clark,M.D. and Voytas,D.F. (2020) Plant gene editing through *de novo* induction of meristems. *Nat. Biotechnol.*, 38, 84–89.
 95. Burris,K.P., Dlugosz,E.M., Collins,A.G., Stewart,C.N. and Lenaghan,S.C. (2016) Development of a rapid, low-cost proplast transfection system for switchgrass (*Panicum virgatum* L.). *Plant Cell Rep.*, 35, 693–704.
 96. Burch-Smith,T.M., Anderson,J.C., Martin,G.B. and Dinesh-Kumar,S.P. (2004) Applications and advantages of virus-induced gene silencing for gene function studies in plants. *Plant J.*, 39, 734–746.
 97. Mei,Y., Zhang,C., Kernodle,B.M., Hill,J.H. and Whitham,S.A. (2016) A Foxtail mosaic virus vector for virus-induced gene silencing in maize. *Plant Physiol.*, 171, 760–772.
 98. Wang,R., Yang,X., Wang,N., Liu,X., Nelson,R.S., Li,W., Fan,Z. and Zhou,T. (2016) An efficient virus-induced gene silencing vector for maize functional genomics research. *Plant J.*, 86, 102–115.
 99. Liu,N., Xie,K., Jia,Q., Zhao,J., Chen,T., Li,H., Wei,X., Diao,X., Hong,Y., Liu,Y. et al. (2016) Foxtail mosaic virus-induced gene silencing in monocot plants. *Plant Physiol.*, 171, 1801–1807.
 100. Ramegowda,V., Mysore,K.S., Senthil-Kumar,M. (2014) Virus-induced gene silencing is a versatile tool for unraveling the functional relevance of multiple abiotic-stress-responsive genes in crop plants. *Front Plant Sci.*, 5, 323.
 101. Zhang,H., Hong,Y., Huang,L., Liu,S., Tian,L., Dai,Y., Cao,Z., Huang,L., Li,D., Song,F. et al. (2016) Virus-induced gene silencing-based functional analyses revealed the involvement of several putative trehalose-6-phosphate synthase/-phosphatase genes in disease resistance against *Botrytis cinerea* and *Pseudomonas syringae* pv. tomato DC3000 in tomato. *Front. Plant Sci.*, 7, 1176.
 102. Reski,R., Parsons,J. and Decker,E.L. (2015) Moss-made pharmaceuticals: from bench to bedside. *Plant Biotechnol. J.*, 13, 1191–1198.
 103. Ke,J. and Yoshikuni,Y. (2020) Multi-chassis engineering for heterologous production of microbial natural products. *Curr. Opin. Biotechnol.*, 62, 88–97.
 104. Ignea,C., Trikka,F.A., Nikolaidis,A.K., Georgantea,P., Ioannou,E., Loupassaki,S., Kefalas,P., Kanellis,A.K., Roussis,V., Makris,A.M. et al. (2015) Efficient diterpene production in yeast by engineering Erg20p into a geranylgeranyl diphosphate synthase. *Metab. Eng.*, 27, 65–75.
 105. Wang,G., Zhao,Z., Ke,J., Engel,Y., Shi,Y.-M., Robinson,D., Bingol,K., Zhang,Z., Bowen,B., Louie,K. et al. (2019) CRAGE enables rapid activation of biosynthetic gene clusters in undomesticated bacteria. *Nat. Microbiol.*, 4, 2498–2510.
 106. Ajikumar,P.K., Xiao,W.-H., Tyo,K.E.J., Wang,Y., Simeon,F., Leonard,E., Mucha,O., Phon,T.H., Pfeifer,B., Stephanopoulos,G. et al. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science*, 330, 70–74.
 107. Reed,J. and Osbourn,A. (2018) Engineering terpenoid production through transient expression in *Nicotiana benthamiana*. *Plant Cell Rep.*, 37, 1431–1441.
 108. Montague,N.P., Thuenemann,E.C., Saxena,P., Saunders,K., Lenzi,P. and Lomonosoff,G.P. (2011) Recent advances of cowpea mosaic virus-based particle technology. *Hum. Vaccin.*, 7, 383–390.
 109. Engler,C., Youles,M., Gruetzner,R., Ehnert,T.-M., Werner,S., Jones,J.D.G., Patron,N.J. and Marillonnet,S. (2014) A golden gate modular cloning toolbox for plants. *ACS Syn. Biol.*, 3, 839–843.
 110. Weber,E., Engler,C., Gruetzner,R., Werner,S. and Marillonnet,S. (2011) A modular cloning system for standardized assembly of multigene constructs. *PLoS One*, 6, e16765.
 111. Keasling,J.D. (2012) Synthetic biology and the development of tools for metabolic engineering. *Metab. Eng.*, 14, 189–195.
 112. Brückner,K. and Tissier,A. (2013) High-level diterpene production by transient expression in *Nicotiana benthamiana*. *Plant Methods*, 9, 46.
 113. Bach,S.S., Bassard,J.-É., Andersen-Ranberg,J., Møldrup,M.E., Simonsen,H.T. and Hamberger,B. (2014) High-throughput testing of terpenoid biosynthesis candidate genes using transient expression in *Nicotiana benthamiana*. *Methods Mol. Biol.*, 1153, 245–255.
 114. Irmisch,S., Ruebsam,H., Jancsik,S., Man Saint Yuen,M., Madilao,L.L. and Bohlmann,J. (2019) Flavonol biosynthesis genes and their use in engineering the plant antidiabetic metabolite montbretin A. *Plant Physiol.*, 180, 1277–1290.
 115. Reski,R., Bae,H. and Simonsen,H.T. (2018) *Physcomitrella patens*, a versatile synthetic biology chassis. *Plant Cell Rep.*, 37, 1409–1417.
 116. Goren,M.A., Nozawa,A., Makino,S.-i., Wrobel,R.L. and Fox,B.G. (2009) Cell-free translation of integral membrane proteins into unilamellar liposomes. *Methods Enzymol.*, 463, 647–673.
 117. Moore,S.J., Lai,H.-E., Needham,H., Polizzi,K.M. and Freemont,P.S. (2017) *Streptomyces venezuelae* TX-TL—a next generation cell-free synthetic biology tool. *Biotechnol. J.*, 12, 1600678.
 118. Zhang,N., Han,Z., Sun,G., Hoffman,A., Wilson,I.W., Yang,Y., Gao,Q., Wu,J., Xie,D., Dai,J. et al. (2014) Molecular cloning and characterization of a cytochrome P450 taxoid 9a-hydroxylase in *Ginkgo biloba* cells. *Biochem. Biophys. Res. Commun.*, 443, 938–943.
 119. Andersen-Ranberg,J., Kongstad,K.T., Nielsen,M.T., Jensen,N.B., Pateraki,I., Bach,S.S., Hamberger,B., Zerbe,P., Staerk,D., Bohlmann,J. et al. (2016) Expanding the landscape

- of diterpene structural diversity through stereochemically controlled combinatorial biosynthesis. *Angew. Chem. Int. Ed. Engl.*, 55, 2142–2146.
120. Jia, M., Potter, K.C. and Peters, R.J. (2016) Extreme promiscuity of a bacterial and a plant diterpene synthase enables combinatorial biosynthesis. *Metab. Eng.*, 37, 24–34.
121. Morrone, D., Hillwig, M.L., Mead, M.E., Lowry, L., Fulton, D.B. and Peters, R.J. (2011) Evident and latent plasticity across the rice diterpene synthase family with potential implications for the evolution of diterpenoid metabolism in the cereals. *Biochem. J.*, 435, 589–595.
122. Afendi, F.M., Okada, T., Yamazaki, M., Hirai-Morita, A., Nakamura, Y., Nakamura, K., Ikeda, S., Takahashi, H., Altaf-Ul-Amin, M., Darusman, L.K. et al. (2012) KNAPsAcK family databases: integrated metabolite-plant species databases for multifaceted plant research. *Plant Cell Physiol.*, 53, e1.
123. Andolfo, G., Iovieno, P., Frusciante, L. and Ercolano, M.R. (2016) Genome-editing technologies for enhancing plant disease resistance. *Front. Plant Sci.*, 7, 1813.
124. Aksenov, A.A., da Silva, R., Knight, R., Lopes, N.P. and Dorrestein, P.C. (2017) Global chemical analysis of biology by mass spectrometry. *Nat. Rev. Chem.*, 1, 0054.
125. Silva, N.V. and Patron, N.J. (2017) CRISPR-based tools for plant genome engineering. *Emerg. Top Life Sci.*, 1, 135–149.
126. Oliver, N.J., Rabinovitch-Deere, C.A., Carroll, A.L., Nozzi, N.E., Case, A.E. and Atsumi, S. (2016) Cyanobacterial metabolic engineering for biofuel and chemical production. *Curr. Opin. Chem. Biol.*, 35, 43–50.
127. Peralta-Yahya, P.P., Ouellet, M., Chan, R., Mukhopadhyay, A., Keasling, J.D. and Lee, T.S. (2011) Identification and microbial production of a terpene-based advanced biofuel. *Nat. Commun.*, 2, 483.