Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Getting back to the grass roots: harnessing specialized metabolites for improved crop stress resilience

Yezhang Ding[1], Trent R Northen[1,2], Ahmed Khalil[3], Alisa Huffaker[3] and Eric A Schmelz[3]

Roots remain an understudied site of complex and important biological interactions mediating plant productivity. In grain and bioenergy crops, grass root specialized metabolites (GRSM) are central to key interactions, yet our basic knowledge of the chemical language remains fragmentary. Continued improvements in plant genome assembly and metabolomics are enabling large-scale advances in the discovery of specialized metabolic pathways as a means of regulating root-biotic interactions. Metabolomics, transcript coexpression analyses, forward genetic studies, gene synthesis and heterologous expression assays drive efficient pathway discoveries. Functional genetic variants identified through genome wide analyses, targeted CRISPR/Cas9 approaches, and both native and non-native overexpression studies critically inform novel strategies for bioengineering metabolic pathways to improve plant traits.

Addresses
[1] Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[2] Joint BioEnergy Institute, Emeryville, CA 94608, USA
[3] Section of Cell and Developmental Biology, University of California at San Diego, La Jolla, CA, USA

Corresponding author: Schmelz, Eric A (eschmelz@ucsd.edu)

Current Opinion in Biotechnology 2021, **70**:174–186

This review comes from a themed issue on **Plant biotechnology**

Edited by **Harro Bouwmeester** and **Rob Schuurink**

https://doi.org/10.1016/j.copbio.2021.05.010

Introduction

Soil ecosystems sustain rich and dynamic communities that include bacteria, fungi, viruses, nematodes, arthropods and plants [1,2]. Given the reliance of heterotrophs on autotrophs for carbon fixation, these complex ecosystems are often structured around plant root systems. The roots of grass and grain crops form interaction networks with diverse soil organisms mediated in part by biologically active specialized metabolites. Grass root specialized metabolites (GRSM) present in tissues and/or exudates play multifunctional roles as endogenous immune signals, facilitators of nutrient acquisition, defenses against detrimental organisms and mediators of complex interactions with neighboring plants and other organisms [3,4]. Among plant tissue types, roots are significant sites of microbial diversity and density. Considering plant diversity at the species and genome-wide levels coupled with diversity in associated root macrobiomes and microbiomes, interactions mediated by GRSM are wildly complex. Given the expansive breadth, challenge and opportunity, we focus this review on monocots, specifically grass models which include maize (*Zea mays*), rice (*Oryza* spp.), sorghum (*Sorghum bicolor*), wheat (*Triticum aestivum*), oats (*Avena sativa*), barley (*Hordeum vulgare*), rye (*Secale cereale*) and switchgrass (*Panicum virgatum*). Collectively grasses cover approximately 40% of the natural and agroecosystems spanning our earths' arable land [5].

In the past five decades, major biological advances have come from the mastery of nucleotide and amino acid polymer chemistries enabling sequencing, relative quantification and synthesis of RNA, DNA and polypeptides. A conspicuous remaining challenge is the accurate annotation of and control over diverse specialized metabolomes. Isoprenoid pathways alone produce more than 80 000 different structures [6]. Compared to protein, DNA and RNA, the vast chemical diversity of core precursors and structural modifications in plant specialized metabolic pathways have made them historically recalcitrant to efficient systematic efforts aimed at annotation and manipulation for beneficial applications. Improvements in analytical chemistry, particularly in mass spectrometry, now enable diverse research groups to interrogate the presence of known and unknown plant specialized metabolites as traits underlying biological processes of interest.

Advances in genomics, transcriptomics, proteomics, genetics, synthetic biology and bioinformatics increasingly provide efficient tools for functional discovery and leveraged application of root specialized metabolism [4,6]. An understanding of the biosynthesis and regulation of GRSM now enables targeted manipulation and functional examination of roles mediating belowground interactions [7•,8••,9,10]. Our collective ability to rapidly connect genotypes to biochemical phenotypes that impact biotic interactions informs breeding and engineering approaches to ensure that key GRSM are present to enhance biotic stress resistance [11,12]. We focus this review on approaches to understand and harness GRSM.

Major GRSM families and functions

Roots continuously interact with both biotic and abiotic factors in the rhizosphere, often responding to stimuli by producing diverse specialized metabolites. GRSM can be produced constitutively or display elicited production [3]. Lipophilic GRSM are not commonly associated with glandular trichomes or secretory cavities, as is common in dicots, and thus are likely to be associated with either lipid droplets present in all forms of life or secreted into the soil [13]. Some GRSM are highly tissue-specific [14–16]; however, GRSM can also commonly co-occur in foliar tissues. Predictably, GRSM display pangenome variation in biochemical diversity consistent with complex selection pressures [11,12]. We briefly review examples of GRSM that mediate plant-biotic interactions, including complex shikimate pathway containing molecules, terpenoids, benzoxazinoids and fatty acid derivatives (Figure 1)

Shikimate pathway containing molecules: flavonoids, phenolics and conjugates

Core shikimate pathway precursors are commonly directed to phenylpropanoid biosynthesis. The phenylpropanoid pathway is encoded by multigene families, such as phenylalanine ammonia lyase, cinnamate 4-hydroxylase, and 4-coumarate-CoA ligase, leading to *p*-coumaroyl-CoA precursors for biosynthesis of complex monolignol pathway products and diverse flavonoids [17]. Flavonoids encompass >10 000 compounds as aglycones and complex glycoside conjugates [18]. While many GRSM are species-specific, all plants contain chalcone synthases which are type III polyketide synthases (PKS) that use 4-coumaroyl-CoA and malonyl-CoA as substrates to produce naringenin chalcone precursors underlying complex arrays of subsequently derived flavonoids [17]. Maize roots challenged with the fungal pathogen *Colletotrichum graminicola* accumulate naringenin-chalcone and simple derivatives such as genkwanin, apigenin and eryodictiol [19]. Similarly, in diverse wheat lines, association with a mycorrhizal fungus (*Funneliformis mosseae*) resulted in dramatic increases in scutellarin and luteolin 7-*O*-glucuronide [20]. In sorghum, high root levels of 3-deoxyanthocyanidins, such as luteolinidin, are associated with a greater diversity of root bacterial taxa [21]. Beyond antifungal roles, as sakuranetin in rice, modes of action of flavonoids, general antioxidant activities and roles in microbial attraction are examined (Figure 1) [22].

Simple phenylpropanoids derived from cinnamic acid occur as GRSM exudates [23] and can further serve as precursors to conjugates with amides, sucrose esters and hydroxycitric acids. The phenylamide family is derived from the pairing of cinnamic acid and derivatives such as ferulic, caffeic, *p*-coumaric and benzoic acids that are conjugated to amines such as tyramine, tryptamine, serotonin, agmatine, and putrescine [24] (Figure 1). Further GRSM in oats include anthranilic acid amides, termed avenanthramides [25]. In barley roots, infection with *Fusarium culmorum* results in the production of cinnamic acid amides with modified tryptamine derivatives such as 9-hydroxy-8-oxotryptamine, 8-oxotryptamine, and (1H-indol-3-yl)methylamine, termed triticamides A, B and C [24]. Phenylamides are commonly antibiotic, but in maize can be inducible susceptibility factors driven by bacterial pathogens such as *Pantoea stewartii* [26]. Beyond amide conjugates, antifungal phenolic sucrose esters, such as smiglaside C (3,6-diferuloyl-2′,3′,6′-triacetylsucrose), are produced in maize roots following *Fusarium graminearum* elicitation (Figure 1) [27]. Additionally, coumaryl-hydroxycitric, caffeolyl-hydroxycitric and feruloyl-hydroxycitric acid conjugates are also found in maize [28•].

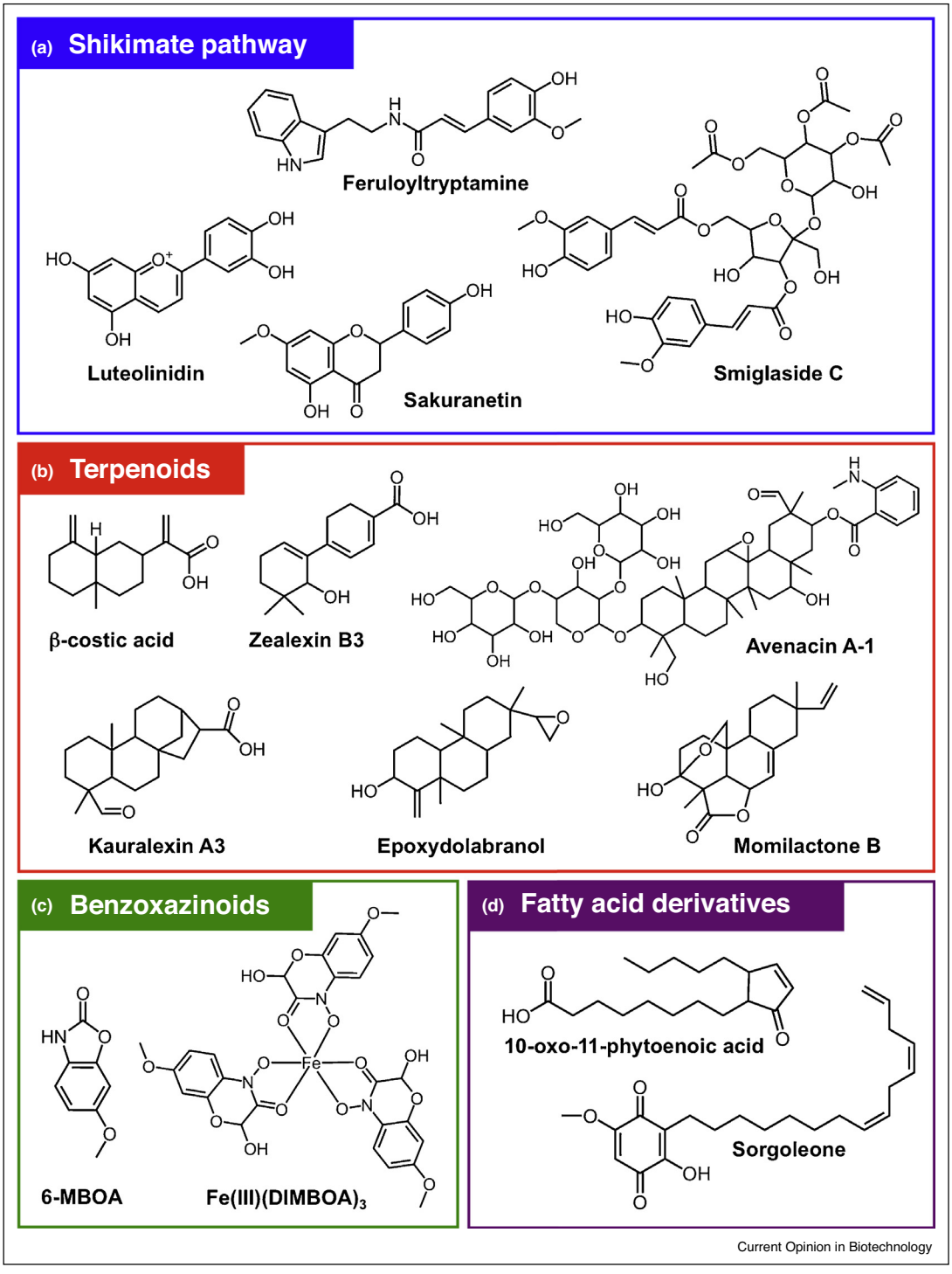
Terpenoids

Similar to phenylpropanoids, terpenoids constitute an expansive family of GRSM [6]. Families of terpene synthases (TPS) are commonly encoded by 40–50 genes and catalyze the dephosphorylation and cyclization of the isoprenoid precursors geranyl(G)-diphosphate (PP) farnesyl-PP (FPP), geranylgeranyl-PP (GGPP), and squalene/2,3-oxidosqualene to generate diverse mono(C10)-terpenoids, sesqui(C15)-terpenoids, di(C20)-terpenoids, and tri(C30)-terpenoids, respectively [29]. Further diversity is generated by cytochrome P450 (CYP) enzymes often in the CYP71, CYP99, CYP701, CYP76 and CYP81 families resulting in diverse oxidized terpenes [8••,29,30••].

Monoterpenes and sesquiterpenes

While ubiquitous in plants, monoterpenes are rarely observed as dominant GRSM, although recently switchgrass PvTPS04 was discovered to account for borneol biosynthesis [31]. Select terpene products can be favored depending on substrate availability. For example, the products of maize ZmTPS1 are substrate dependent and can be either monoterpenes, sesquiterpenes or diterpenes [32]. Sesquiterpenoids are more commonly encountered GRSM. In switchgrass, 10 different *TPS* transcripts accumulated in roots following jasmonate treatment and tracked emission of (E)-β-caryophyllene, cycloisosativene, β-elemene, α-humulene, α-selinene, germacrene D and δ-cadinene [31]. In maize, western corn rootworm (*Diabrotica virgifera*) larvae elicit root emission of (E)-β-caryophyllene, which aids in the attraction of entomophagous nematodes as an indirect defense [12]. *D. virgifera* larvae can similarly use (E)-β-caryophyllene to identify host plants [33]. Maize roots further produce non-volatile sesquiterpenoid antibiotics derived from 5 TPS enzymes functioning as α/β-selinene synthases and β-bisabolene/β-macrocarpene synthases [8••,11]. Resulting microbe-induced production of β-costic acid and β-bisabolene/β-macrocarpene derived acids, termed zealexins (ZX), protect against fungal pathogens and negatively influence the growth of *Diabrotica balteata* larvae (Figure 1) [8••,11]. In maize roots, ZX are produced by enzymes encoded by duplicated families of *TPS*, *CYP71Z* and *CYP81A* genes which drive

Figure 1



Representative GRSM examples.
(a) The shikimate pathway underlies production of complex phenylpropanoids and diverse derivatives that include phenylamides (feruloyltryptamine), flavonoids (luteolinidin, sakuranetin) and phenolic sucrose esters (smiglaside C). **(b)** Terpenoid derived GRSM antibiotics include sesquiterpenoids (β-costic acid and zealexin B3 in maize), diterpenoids (kauralexin A3 and epoxydolabranol in maize; momilactone B in rice) and triterpenoid saponins (avenacin A-1 in oats). **(c)** In maize, wheat and rye, benzoxazinoid glucoside cleavage products such as 2,4-dihydroxy-7-methoxy-2H-1,4-benzoxazin-3(4H)-one (DIMBOA) can function as iron siderophore complexes [Fe(III)$(DIMBOA)_3$] while smaller bioactive degradation products persist in soils (6-methoxy-benzoxazolin-2-one; 6-MBOA). **(d)** Diverse fatty acid derivatives can display cytotoxic activity and include root-specific alkylresorcinols in sorghum (sorgoleone) and 9-lipoxygenase derived death acids in maize (10-oxo-11-phytoenoic acid).

the endogenous production of at least 17 pathway products [8••]. Maize *zx1 zx2 zx3 zx4* quadruple mutants lacking β-macrocarpene synthases are susceptibile to diverse pathogens and display altered root microbiome compositions [8••].

Diterpenes and triterpenes

Rice, maize and switchgrass all produce root diterpenoids. Most commonly type II diterpene synthases (DiTPS) utilize the precursor GGPP to produce *ent*-copalyl diphosphate (CPP; *ent*-CPP), (+)-CPP, *syn*-CPP and other bicyclic labdadienyl intermediates, which are acted on by type I DiTPS for additional ring closures [34]. Rice diterpenoid pathways have been systematically examined and reviewed [35]. Root secreted momilactones absent in *Oscps4* rice mutants are important for non-host resistance to the fungus *Magnaporthe poae* (Figure 1) [36]. The rice momilactone pathway has been completed and optimized by shifting biosynthesis from the chloroplast to the cytosol [37••]. Biosynthesis of the rice diterpenoid 5,10-diketo-casbene was also recently described; however, its root biology remains unclear [38•]. Maize roots deploy two *ent*-CPP synthases (*ent*-CPS), a root associated (+)-CPS synthase (ZmCPS3) and a labda-8,13-dien-15-yl PP synthase (ZmCPS4) [34]. Maize roots deploy an *ent*-CPS, termed anther ear2 (ZmAN2), kaurene synthase like 4 and two CYP71Z enzymes to produce root-specific dolabradiene-derived antibiotics termed dolabralexins [15]. Epoxydolabranol displays potent growth inhibition to multiple *Fusarium* species [15]. Kauralexins constitute a second maize diterpenoid family (Figure 1). Kauralexin biosynthesis relies on an *ent*-isokaurene synthase and a steroid 5α reductase enabling the indirect production of predominant *ent*-kaurane associated defenses that circumvent use of gibberellin precursors [30••]. Gene family discovery, driven by gene synthesis and combinatorial enzyme assays, revealed additional diterpenoid complexity in switchgrass and foxtail millet (*Setaria italica*) consistent with common roles in grasses [39,40]. Beyond antibiotics. the diterpenoid 3-*epi*-brachialactone exists as a potent biological nitrification inhibitor produced by the grass *Brachiaria humidicola* [41]. Two GGPP diterpenoid precursors can be further condensed to phytoene (C40) enabling β-carotene production that then serves as a precursor for strigolactone biosynthesis with multifunctional roles in root architecture, hyphal branching in mycorrhizal fungi and germination cues for parasitic plants from the genus *Striga* [42]. Complex arrays of non-cannonical strigolactones, including zeapyranolactone, continue to be elucidated in GRSM exudates [42]. Triterpenoids are also common GRSM and exemplified by oat glycosylated triterpenoids, termed avenacins, that function as potent antifungal agents with broad impacts on eukaryotic rhizosphere communities (Figure 1) [43]. Recent analyses further demonstrate that steroidal (C27) saponins and terpenoid glycosides derived from diosgenin, oxydiosgenin and anhydrodiosgenin sapogenin cores can predominate in switchgrass roots [44].

Benzoxazinoids (BXs)

BXs also common GRSM exemplified by antifeedant, insecticidal, antimicrobial and allelopathic activities in maize, wheat and rye [45]. Specific breakdown products of BX-glucoside conjugates, such as 2,4-dihydroxy-7-methoxy-2H-1,4-benzoxazin-3(4H)-one (DIMBOA) are biocidal and function as signals triggering callose deposition associated with chitin induced immunity [46]. BXs can occur at high concentrations in roots and *bx* pathway mutants display significant differences in rhizosphere and root communities of pathogenic and beneficial microbes [7•,10]. The degradation product of DIMBOA, 6-methoxy-benzoxazolin-2-one (6-MBOA), can persist in soils driving plant defense [9]. DIMBOA can also form siderophore complexes with iron, improving nutrient acquisition in both plants and root herbivores [47] (Figure 1). GRSM have context-specific multifunctional roles-dependent upon the interactions examined [3].

Fatty acid derivatives

Oxygenated fatty acids, termed oxylipins, occur in all plant tissues. Pathways derived from the action of 9/13-lipoxygenase (LOX), allene oxide synthase and hydroperoxy-lyase are commonly studied for roles in defense regulation mediated by jasmonate biosynthesis [48]. In wounded maize roots, levels of the linolenic acid 13-LOX derived jasmonate precursor 12-oxo-11-phytodienoic acid can be exceeded by 9-LOX derived positional isomers, 10-oxo-11-phytodienoic acid and linoleic acid derived 10-oxo-11-phytoenoic acid, termed Death Acids (Figure 1) [49]. 10-OPEA is broadly toxic to insects, fungi and plants, in part through activation of cysteine proteases. Microbially elicited GRSM are not limited to below ground interactions. In maize roots, *Trichoderma virens* promotes induced systemic resistance against the foliar pathogen *C. graminicola*. Transfusions of xylem sap containing increased levels of the 13-LOX α-ketol product of octadecadienoic acid (9-hydroxy-10-oxo-12(Z),15(Z)-octadecadienoic acid) triggers receiver plant resistance [50•]. Initially derived from palmitoleoyl (C16:1)-CoA, the alkylresorcinol termed sorgoleone is a dominant sorghum root metabolite (Figure 1) [16]. Diverse alkylresorcinols further occur in wheat, rye, barley and rice. Alkylresorcinols are antimicrobial, allelochemicals suppressing neighboring plants, drivers of mycorrhizal colonization and mediators of rhizosphere compositions of nitrifying microorganisms [51,52]. Consistent with phytotoxic action, sorgoleone pathway engineering in tobacco (*Nicotiana benthamiana*) using a fatty acid desaturase, alkylresorcinol synthase family type III PKS, *O*-methyltransferase and SbCYP71AM1 resulted in necrosis and cell death [16].

Roles for GRSM in structuring rhizosphere communities

Plant metabolites significantly influence root associated microbial communities and the topic has been extensively

reviewed [53]. For GRSM, maize has been the predominant model. For example, integrated metabolite profiling and 16s rRNA sequencing of wild type maize lines and three different *bx* pathway mutants have revealed GRSM roles in constraining the composition of soil microbial taxa [7•,10]. This approach was further used to investigate the roles of root diterpenoids and demonstrated that *Zman2* mutants deficient in kauralexins [54] and dolabralexins displayed altered Alphaproteobacteria abundance [55]. Tools to understand GRSM effects include the combined use of pathway mutants, isolated microbial community members [56] and reconstituted synthetic microbe communities (SynCom). For example, Lebeis *et al.* used a 38 member SynCom to identify salicylic acid as required for the assembly of normal root bacterial communities by functioning as a signal or carbon source [57]. Similarly, exometabolite profiling of rhizosphere and soil bacteria samples demonstrated that aromatic organic acids from wild oats (*Avena barbata*) can be preferentially consumed by rhizosphere bacteria providing a mechanism for host selection [23].

Integrative approaches for GRSM pathway discovery

Plant metabolomics

Metabolomics combining gas chromatography (GC) mass spectrometry (MS) (GC/MS), liquid chromatography MS (LC/MS) and nuclear magnetic resonance (NMR) spectroscopy approaches are essential in the identification of GRSM. This topic has been comprehensively reviewed [58]. GC/MS is especially well suited for tracking many small molecules under a 350 daltons and was leveraged in the expanded analyses of maize zealexins (Figure 2a). To facilitate progress, the annotation of thousands of unknown GRSM present in extracts is urgently needed. Recently a computational tool, termed class assignment and ontology prediction using mass spectrometry (CANOPUS), was developed to combine MS fragmentation spectra and deep neural networks to accurately assign annotation for 2000 compound classes [59•].

Tools connecting GRSM to genotypes, reference genomes and pan-genomes are increasingly available for grasses including maize, rice, wheat, sorghum, switchgrass, *Miscanthus*, *Brachypodium distachyon* and others [60,61]. While GRSM biosynthetic enzyme classes and gene families can now be partially predicted, specific enzymes of interest often remain unclear. Recent innovations in genomics, RNA-seq based transcriptome coexpression analyses, proteomics, bioinformatics, synthetic biology, forward and reverse genetics can be cross-leveraged to make GRSM pathway discovery an efficient process. Advances in core biochemical technologies are ongoing [58,62]; thus, we primarily focus on connecting GRSM and biology.

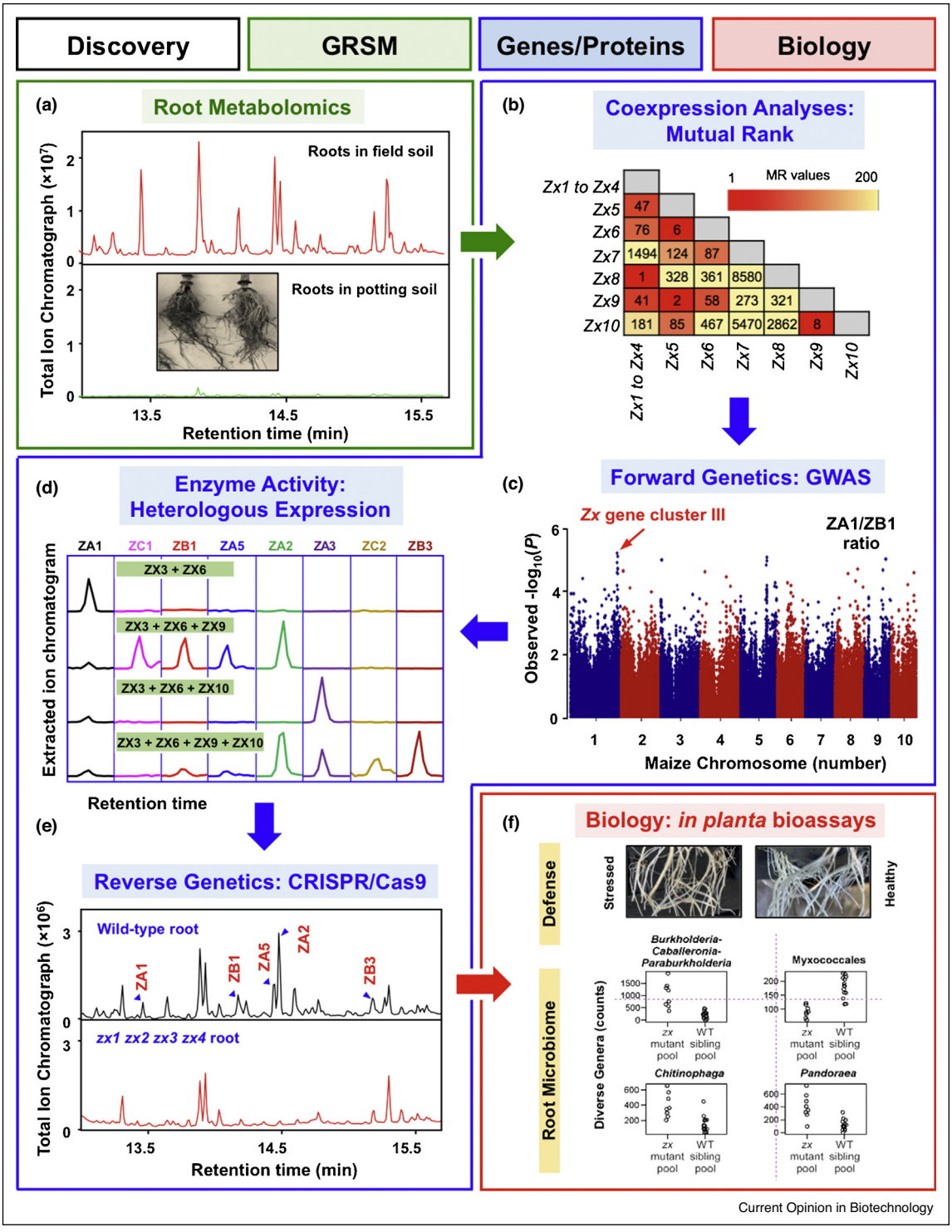
Coexpression, genomic organization and gene families

GRSM biosynthetic genes are commonly co-regulated in a spatiotemporal-dependent and/or environmental-dependent manner [63]. Transcriptional coexpression patterns can be simply interrogated via Mutual Rank analyses [63] for hypothesis testing using public R shiny web-applications to identify candidate pathway genes of interest [64]. Coexpression analyses prioritized momilactone biosynthetic pathway candidates in rice and contributed to demonstrating interconnections between sesquiterpenoid and diterpenoid pathways in maize [8••,30••,37••,64,65•]. Beyond transcript coexpression, correlation analyses can be conducted with dataset combinations [66]. In 1997, discovery of the maize BX biosynthetic pathway revealed the first GRSM gene cluster in plants and initiated interrogation of many biosynthetic gene clusters [67]. Subsequently GRSM gene clusters of different sizes have aided in the pathway discovery of avenacins, momilactones, phytocassanes, zealexins and *S. italica* diterpenoids [30••,39,65•,68]. While gene clusters merit consideration, examples of both functionally irrelevant gene clusters and broad genomic scattering are common [30••,63]. To address this challenge, machine learning is being applied to accurately separate genes predicted in specialized and generalized metabolic pathways [69••]. Prioritization of candidate biosynthetic genes is ideally paired with phylogenetic analyses and consideration of genetic variation including duplications. For example, in maize two CYP71Z subfamily of P450s have identical catalytic activity in the biosynthesis of kauralexins and zealexins, however, transcript co-regulation suggests unequal association with the ZX pathway (Figure 2b) [8••]. Phylogenetic analysis can prevent pathway genes from being overlooked while pan-genome analyses can reveal significant expansions and contractions of GRSM pathway genes [8••].

Forward genetics

Transcript coexpression linked to forward genetic approaches bridge the gap from simple pathway candidates to a high confidence targets. Genome-wide association studies (GWAS) and quantitative trait locus (QTL) analyses powerfully leverage single nucleotide polymorphisms (SNPs) to statistically link genetically variable traits to candidate genes. GWAS and metabolite GWAS (mGWAS) are mature research tools with over 1000 crop plant studies and approximately 50 annual studies conducted on rice and maize alone [70,71]. A starting point for mGWAS is the use of Trait Analysis by Association, Evolution and Linkage (TASSEL) [72]. The current version (TASSEL 5; www.maizegenetics.net/tassel) is continuously updated and is simple enough to be utilized is undergraduate laboratory classes and directly interfaced through R for larger applications [73]. Currently over 30 monocots with reference genomes and summaries of SNP resources exist (http://www.gramene.org). Pangenomes contain thousands of genes absent from single reference genomes [74] and can complicate interpretation of GWAS results. Recent progress now comes closer to

Figure 2



Integrated approaches for the discovery of GRSM pathways and functions: example, the maize zealexin (ZX) pathway. **(a)** Root metabolome differences revealed by gas chromatography mass spectrometry (GC/MS) analyses of maize plants grown in field soil versus potting soil. **(b)** Transcript coexpression analyses can prioritize candidate genes in GRSM pathways. Mutual Rank (MR) coexpression heatmap displaying correlations between the sum of four β-macrocarpene synthase genes (*Zx1* to *Zx4*) with cytochrome P450 (*CYP*) genes encoding enzymes in the CYP71Z (*Zx5*, *Zx6*, *Zx7*) and CYP81A (*Zx8*, *Zx9*, *Zx10*) families. **(c)** Forward genetic approaches using metabolite-based Genome Wide Association Studies (mGWAS) in elicited tissues. Using the ratio of ZA1 to ZB1 as a trait, a Manhattan plot displays statistically significant single nucleotide polymorphisms (SNPs) associated with a ZX pathway gene cluster. **(d)** For the validation of candidate genes via efficient enzyme function studies, *Agrobacterium* mediated transient heterologous expression of ZX pathway enzymes were conducted in

reference-free association mapping with the development of Practical Haplotype Graphs that capture genotype variation with modest file sizes [75]. The use of short DNA sequences, termed *k*-mers, is likewise a reference genome-independent approach to powerfully link traits to genomic regions [76••]. Gene coexpression analyses and mGWAS are complementary discovery approaches and have been combined in an open-source and validated framework termed Coanalysis of molecular components (Camoco) [77]. Non-targeted mGWAS commonly yields hundreds of significant associations between metabolites and biosynthetic or pathway regulatory genes [28•,70]. In maize the combined integrative transcript coexpression and mGWAS and was recently used to endogenously support and identify ZX biosynthetic genes (Figure 2b and c) and two other GRSM antibiotic pathways [8••,11,30••].

Proteomics

With sequenced genomes, proteomics can facilitate prioritization of candidate biosynthetic enzymes for specialized metabolic pathways [8••]. However, due to lower sensitivity and greater costs, proteomics are used less frequently than transcriptomics for gene discovery. Collective analyses of transcriptome, proteome and metabolite production from target tissues can verify metabolic pathway inter-conversions and drive gene discoveries in difficult non-model species [78]. Proteomics approaches are likely to be expanded as mass spectrometry-independent sequencing technologies are optimized [62]

Validation of enzyme function

Following candidate gene identification, verification of enzyme function is desirable before pursuing mutants in most grass models. DNA synthesis is becoming a cost-effective approach for the rapid assembly of gene candidates into expression vectors for functional analysis [79]. The U.S. Department of Energy Joint Genome Institute (JGI) supports large-scale gene synthesis proposals for the discovery of GRSM [79] enabling the systematic interrogation of gene families and functions [31,39,40]. Biochemical approaches are commonly used for functional analyses following heterologous protein expression, purification, and *in vitro* enzymatic assays with chemical substrates when available. Challenges include insufficient protein expression, low enzymatic activity and lack of specialized substrates as pathway intermediates [61]. Improvements in heterologous bacteria and yeast expression platforms are ongoing; however, the predominant tool for GRSM pathway discovery involves use of *N. benthamiana* [80]. *Agrobacterium*-infiltration of small binary vectors designed for transient heterologous protein expression, termed pEAQ, in *N. benthamiana* are commonly employed [81] with well-established protocols [82]. Advantages include speed, existence of biosynthetic pathway precursors and the ability to interrogate enzyme activity without purification. Many GRSM are grass specific, thus *N. benthamiana* affords a clean background void of existing target metabolites. Recent GRSM advances using this approach include demonstration of the 10-gene maize ZX pathway (Figure 2b and d), the large-scale production of rice momilactones and other valuable plant natural products [8••,37••,83].

Reverse genetics to establish *in planta* GRSM mutants for bioassays

GRSM pathways can be proven by obtaining genetic mutants through mining for genome-wide variation, classical ethyl methanesulfonate-induced mutations, T-DNA insertion lines or expanding transposon-insertion mutant collections [84]. However, predominant tools for precisely targeted mutations are the use of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 9 (Cas9) genome editing approaches (CRISPR/Cas9) [85] and RNA-guided gene silencing techniques, now commonly used to create stable and/or transient modifications for functional studies *in planta* [86]. Creation of a CRISPR/Cas9 derived maize *zx1 zx2 zx3 zx4* quadruple mutant demonstrated a lack ZX metabolites (Figure 2e) and alterations in root microbiome communities (Figure 2f) [8••]. More broadly, CRISPR/Cas9 mutagenesis efforts in 10 diverse monocot families and methods have been recently summarized [87]. Of similar importance are tools for transient protein overexpression in grasses. Towards this goal, a modified sugarcane mosaic virus (SCMV) vector was developed in maize for the transient overexpression of defense associated proteins and functional assessment *in planta* [88]. Potential applications include the expression of non-native genes to explore the production of novel defense chemistries.

Harnessing GRSM for useful agricultural traits

Conventional and molecular breeding

As genetically variable phenotypic traits that contribute to biotic-stress resilience, GRSM are exciting targets to consider for modification. Across pangenomes, thousands of unique genes can exist, and can be deployed to complement germplasm lacking these genes [74]. Likewise deleterious mutations exist in GRSM pathways and

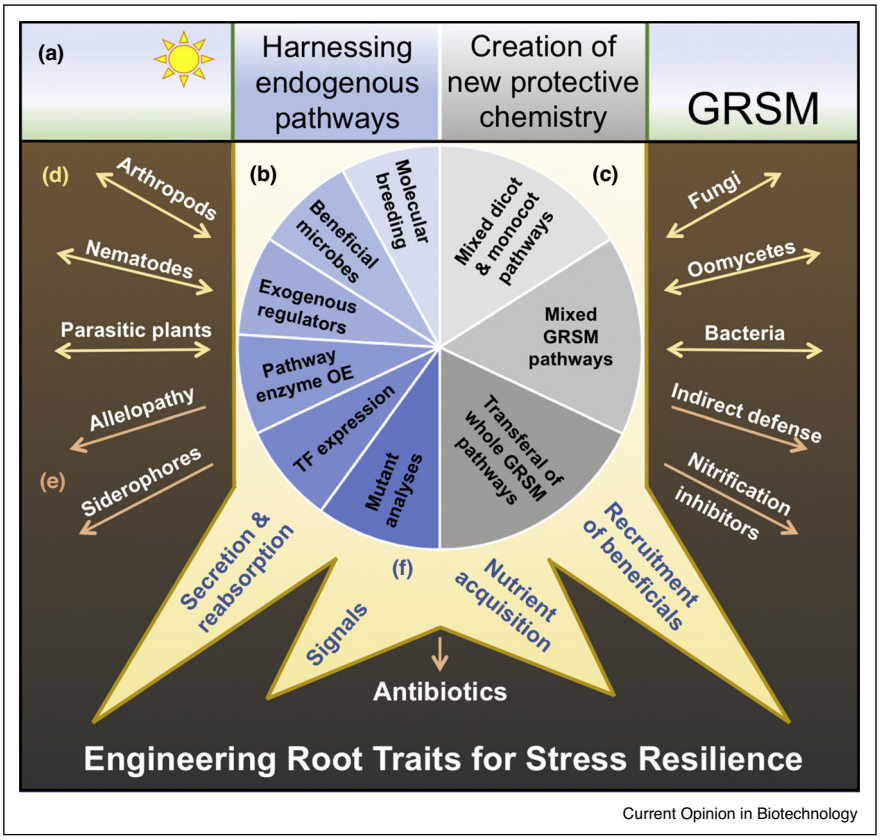
tobacco (*N. benthamiana*). Extracted ion chromatograms show early ZX pathway products (ZA1) and further downstream products (ZC1, ZB1, ZA5, ZA2, ZA3, ZC2) following combinatorial ZX pathway protein expression (ZX3, Zx6, ZX9, ZX10). **(e)** Reverse genetic approaches for *in planta* pathway mutant analyses. The GC/MS total ion chromatograms of wild type maize plants and CRISPR/Cas9 derived *zx1 zx2 zx3 zx4* quadruple mutants lacking root ZX production. **(f)** Functional roles for GRSM in plant biotic interactions. Root microbiome analyses of maize *zx1 zx2 zx3 zx4* quadruple mutants reveal altered abundance of microbial taxa with plants displaying diminished defense resulting in increased susceptibility to fungal and bacterial pathogens.

can be ultimately corrected using marker-assisted breeding [11,89]. For use on a global agricultural scale, breeding approaches lacking transgenes remain favored and are likely to be greatly advanced by precision gene editing that is free of transgenes [90].

Prospects for targeted microbial-mediated control of GRSM

Fungal pathogens in the genus *Fusarium* are potent regulators of plant antibiotic pathways. In asymptomatic non-pathogenic interactions, *Fusarium virguliforme* colonization of maize seeding roots significantly upregulates expression of all characterized maize terpenoid antibiotic pathway genes [8••,91]. This suggests that non-pathogenic microbes can be leveraged to drive GRSM production. As endogenous plant signals, the Plant Elicitor Peptides (Pep) and their cognate receptors (PepR) control complex immune responses partly mediated by jasmonate and ethylene signaling pathways which result in the activation of inducible specialized metabolism [92]. For potato (*Solanum tuberosum*) root delivery, *Bacillus subtilis* were recently engineered to secrete StPep1 which significantly reduced galling by the plant pathogenic nematode, *Meloidogyne chitwoodi* [93•]. Similarly, transient expression of ZmPep1 and ZmPep3 via a sugarcane mosaic virus vector activated maize pathways for

Figure 3



Simplified diagram of approaches to harness GRSM as mediators of complex plant biotic interactions for improved crop stress resilience. **(a)** A mixture of conventional and transgenic approaches exist to control the production of biologically active GRSM. **(b)** Current approaches to harness and understand GRSM include molecular breeding for GRSM pathways genes present in select germplasm. Non-pathogenic microbes can be used to modulate the expression of GRSM genes. Engineered bacteria or virus can be used to control the exogenous delivery of regulators/signals and drive GRSM production. Overexpression (OE) and/or replacement of missing GRSM pathway enzymes can be accomplished with transgenic approaches. OE of key GRSM transcription factors (TFs) and creation of defined GRSM pathway mutants are key tools to empirically test functions. **(c)** Using transgenic approaches, GRSM can be expanded by mixing dicot and monocot pathways, mixing monocot pathways, or conceptually transferring whole GRSM biosynthetic pathways. **(d)** GRSM production will be influenced by diverse stresses imposed by arthropod herbivory, plant-parasitic nematodes, parasitic plants, pathogenic fungi, oomycetes, and bacteria. In many cases diverse beneficial nematodes (entomophagous), fungi and bacteria can be recruited and/or promoted by GRSM production, suppress biotic threats and influence GRSM production. **(e)** Secreted GRSM can act as allelopathic agents by suppressing neighboring plant growth, indirect defenses to promote the attraction of entomophagous nematodes, siderophores for improved solubilization of iron (Fe), bacterial nitrification inhibitors that delay soil conversion of ammonium to nitrite, and broadly act as antibiotics. **(f)** GRSM can be secreted and reabsorbed, aid in nutrient acquisition, function as external and internal signals, and either recruit or promote the growth of diverse beneficial organisms. Whether examined or ignored, complex plant-biotic interactions mediated by GRSM will underlay root trait optimization in field settings.

specialized metabolism and strongly induced resistance against diverse insects [88].

Overexpression

To engineer maize GRSM, a ubiquitin promoter was used to drive expression of an oregano (E)-β-caryophyllene synthase to replace the native unexpressed gene (*ZmTPS23*) in B73 creating constitutive root emission of (E)-β-caryophyllene [12]. Resulting plants exhibited significantly less root damage from *D. virgifera* larvae than the non-(E)-β-caryophyllene-emitting plants via attraction of entomophagous nematodes. Additional studies revealed costs of constitutive-(E)-β-caryophyllene in field settings highlighting both the importance of ecological studies and multifunctional roles of GRSM [94]. In rice, overexpression of the (*S*)-limonene synthase (OsTPS19) resulted in increased metabolite production and enhanced resistance against the fungal pathogen *Magnaporthe oryzae* [95]. Regulatory transcription factors (TFs) offer additional means of manipulating GRSM. In maize, 6000 public RNA-seq samples were recently re-analyzed to link TFs to target genes in predicted Gene Regulatory Networks (GRN) thereby narrowing candidate TFs involved in diverse GRSM pathways including BX, flavonoids, and terpenoids [96••]. As an application, overexpression of a rice bZIP TF (OsTGAP1) led to enhanced root momilactone production and increased allelopathic action on barnyard grass [97]. Given biological complexities, use of inducible promoters enabling targeted activation may be essential to avoid unintended costs and loss of signal information in a field context. With gene discovery projects increasingly completed, the transfer of entire complex pathways between crops is now envisioned. One strategy is to deploy the oat avenacin pathway by engineering biosynthesis in wheat to shield against the devastating 'take-all' disease caused by *Gaeumannomyces graminis* var. *tritici* [6].

Heterologous expression of modular pathway enzymes for new GRSM

Biosynthetic enzymes often lack perfect substrate and product specificity. For example, depending on expression levels, GRSM pathways utilizing modular combinations of type II DiTPS, type I DiTPS and CYPs result in different product profiles with different core structures and sites of oxygenation [8••,15,30••]. P450 substrate promiscuity was recently leveraged in rice to create new antibiotics. Specifically, the expression of maize ZmCYP71Z18 in rice resulted in the modification of endogenous defenses, including the novel production of 15,16-epoxy-*syn*-pimaradien-19-ol, and improved rice disease resistance to *M. oryzae* [98•]. Recent engineering approaches in *N. benthaminana* have demonstrated that the indole-sulfur phytoalexin pathway in crucifers leading to brassinin can be modified further using monocot enzymes to generate novel antifungal agents. Specifically sorghum CYP79A1 and *S. italica* CYP79A2 were used to create novel brassinin-like defenses, termed crucifalexins, with brassinin indole R-groups functionally substituted for novel 4-hydroxybenzyl and benzyl R-groups [99]. The engineered creation of novel biochemicals has great potential to temporarily overcome existing detoxification systems evolved in pests and pathogens and afford new layers of increased protection.

Conclusion and outlook

The discovery of GRSM metabolites, pathway genes and respective biological functions is an increasingly efficient process. While 80–90% of metabolomic features are commonly unknown, expanding literature, MS databases and MSn-based predictions create yearly improvements in class and family level annotations [59•]. While GRSM are complex, commonly encountered chemical convergence will fortify metabolite identifications across diverse plant systems. Sustained efforts in the biochemical (Figure 1) and genetic annotation (Figure 2) of GRSM pathways increasingly provide complete molecular dictionaries that can be tailored, optimized and deployed for the critical analyses of complex yet beneficial biotic and abiotic interactions that govern field traits (Figure 3). With discoveries ongoing, plant scientists currently in training will increasingly contribute to discovering mechanisms of pathway regulation. Specific improvements are still required in defining control points that include signal transduction cascades governing GRSM production, storage and secretion. Highly multifunctional GRSM have been tailored over millions of years of evolution by complex biotic and abiotic selection pressures. Given this reality, targeted optimization of one trait can create deficits in others. With advances in genomics and gene editing using CRISPR/Cas9, a growing limitation for the discovery of how GRSM interact with beneficial and antagonistic bacteria, fungi, nematodes, arthropods and neighboring plants (Figure 3) is in performing detailed field relevant multi-organismal ecological studies. At a more reductionist level, the development of controlled fabricated ecosystems enabling interrogation of metabolites, phenotypes, and microbial interactions in wild type, mutant and engineered plants has great potential for defining GRSM functions. We are now able to delete, add and create novel layers of GRSM. Novel chemistries created by mixing biosynthetic pathways hold great promise and have been demonstrated to afford significant protection against previously adapted pathogens (Figure 3) [98•,99]. While the range of organismal interactions are complex, the discovery, control and improved deployment of GRSM will be fundamental to optimizing biotic and abiotic stress resilience traits.

Conflict of interest statement

Nothing declared.

Acknowledgements

This work was supported by grants from the U.S. Department of Agriculture, National Institute of Food and Agriculture (grant no. 2018-67013-28125, A.H. and E.S.) and the National Science Foundation (NSF), Plant-Biotic Interactions Program (grant no. 1758976, E.S. and Y.D.), NSF Faculty Early Career Development Program (grant no. 1943591, A.H.), the U.S. Department of Energy (DOE) Joint Genome Institute (JGI) Community Science Program (CSP) (grant nos. CSP 2568, E.S. and A.H.) and CSP 503420 (A.H. and E.S.). TN is supported by U.S. DOE, Office of Science, Office of Biological & Environmental Research Microbial Community Analysis and Functional Evaluation in Soils Program and the DOE JGI at Lawrence Berkeley National Laboratory DE-AC02-05CH11231.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Bonkowski M, Villenave C, Griffiths B: **Rhizosphere fauna: the functional and structural diversity of intimate interactions of soil fauna with plant roots**. *Plant Soil* 2009, **321**:213-233.

2. Pascale A, Proietti S, Pantelides IS, Stringlis IA: **Modulation of the root microbiome by plant molecules: the basis for targeted disease suppression and plant growth promotion**. *Front Plant Sci* 2019, **10**:1741.

3. Erb M, Kliebenstein DJ: **Plant secondary metabolites as defenses, regulators, and primary metabolites: the blurred functional trichotomy**. *Plant Physiol* 2020, **184**:39-52.

4. Massalha H, Korenblum E, Tholl D, Aharoni A: **Small molecules below-ground: the role of specialized metabolites in the rhizosphere**. *Plant J* 2017, **90**:788-807.

5. Suttie JM, Reynolds SG, Batello C: *Grasslands of the World*. Food & Agriculture Org.; 2005.

6. Huang AC, Osbourn A: **Plant terpenes that mediate below-ground interactions: prospects for bioengineering terpenoids for plant protection**. *Pest Manag Sci* 2019, **75**:2368-2377.

7. • Cotton TEA, Pétriacq P, Cameron DD, Meselmani MA, Schwarzenbacher R, Rolfe SA, Ton J: **Metabolic regulation of the maize rhizobiome by benzoxazinoids**. *ISME J* 2019, **13**:1647-1658

The authors use maize bx pathway mutants, untargeted mass spectrometry and microbial amplicon sequence analyses to demonstrate BX metabolites significantly influence root-microbe interactions and more globally root metabolism including flavonoid production.

8. •• Ding Y, Weckwerth PR, Poretsky E, Murphy KM, Sims J, Saldivar E, Christensen SA, Char SN, Yang B, Tong A *et al.*: **Genetic elucidation of interconnected antibiotic pathways mediating maize innate immunity**. *Nat Plants* 2020, **6**:1375-1388

The authors combine analyses of transcript coexpression, association mapping, pangenomes, proteomics, enzyme structure–function studies, NMR and CRISPR/Cas9 mediated targeted mutagenesis to identify 10 genes in three zealexin gene clusters responsible in part for the production of diverse maize antibiotics originating from four separate pathways. The study demonstrated that plants lacking zealexins display increased disease susceptibility and alterations in the abundance in root microbiome taxa.

9. Hu L, Robert CAM, Cadot S, Zhang X, Ye M, Li B, Manzo D, Chervet N, Steinger T, van der Heijden MGA *et al.*: **Root exudate metabolites drive plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota**. *Nat Commun* 2018, **9**:2738.

10. Kudjordjie EN, Sapkota R, Steffensen SK, Fomsgaard IS, Nicolaisen M: **Maize synthesized benzoxazinoids affect the host associated microbiome**. *Microbiome* 2019, **7**:59.

11. Ding YZ, Huffaker A, Kollner TG, Weckwerth P, Robert CAM, Spencer JL, Lipka AE, Schmelz EA: **Selinene volatiles are essential precursors for maize defense promoting fungal pathogen resistance**. *Plant Physiol* 2017, **175**:1455-1468.

12. Degenhardt J, Hiltpold I, Köllner TG, Frey M, Gierl A, Gershenzon J, Hibbard BE, Ellersieck MR, Turlings TCJ: **Restoring a maize root signal that attracts insect-killing nematodes to control a major pest**. *Proc Natl Acad Sci U S A* 2009, **106**:13213-13218.

13. Lundquist PK, Shivaiah KK, Espinoza-Corral R: **Lipid droplets throughout the evolutionary tree**. *Prog Lipid Res* 2020, **78**:101029.

14. Louveau T, Orme A, Pfalzgraf H, Stephenson MJ, Melton R, Saalbach G, Hemmings AM, Leveau A, Rejzek M, Vickerstaff RJ *et al.*: **Analysis of two new arabinosyltransferases belonging to the carbohydrate-active enzyme (CAZY) Glycosyl Transferase Family1 provides insights into disease resistance and sugar donor specificity**. *Plant Cell* 2018, **30**:3038-3057.

15. Mafu S, Ding YZ, Murphy KM, Yaacoobi O, Addison JB, Wang Q, Shen ZX, Briggs SP, Bohlmann J, Castro-Falcon G *et al.*: **Discovery, biosynthesis and stress-related accumulation of dolabradiene-derived defenses in maize**. *Plant Physiol* 2018, **176**:2677-2690.

16. Pan Z, Bajsa-Hirschel J, Vaughn JN, Rimando AM, Baerson SR, Duke SO: ***In vivo* assembly of the sorgoleone biosynthetic pathway and its impact on agroinfiltrated leaves of *Nicotiana benthamiana***. *New Phytol* 2021, **230**:683-697.

17. Gomez-Cano L, Gomez-Cano F, Dillon FM, Alers-Velazquez R, Doseff AI, Grotewold E, Gray J: **Discovery of modules involved in the biosynthesis and regulation of maize phenolic compounds**. *Plant Sci* 2020, **291**:110364.

18. Weston LA, Mathesius U: **Flavonoids: their structure, biosynthesis and role in the rhizosphere, including allelopathy**. *J Chem Ecol* 2013, **39**:283-297.

19. Balmer D, de Papajewski DV, Planchamp C, Glauser G, Mauch-Mani B: **Induced resistance in maize is based on organ-specific defence responses**. *Plant J* 2013, **74**:213-225.

20. Bernardo L, Carletti P, Badeck FW, Rizza F, Morcia C, Ghizzoni R, Rouphael Y, Colla G, Terzi V, Lucini L: **Metabolomic responses triggered by arbuscular mycorrhiza enhance tolerance to water stress in wheat cultivars**. *Plant Physiol Biochem* 2019, **137**:203-212.

21. Cloutier M, Chatterjee D, Elango D, Cui J, Bruns MA, Chopra S: **Sorghum root flavonoid chemistry, cultivar, and frost stress effects on rhizosphere bacteria and fungi**. *Phytobiomes J* 2021, **5**:39-50.

22. Górniak I, Bartoszewski R, Króliczewski J: **Comprehensive review of antimicrobial activities of plant flavonoids**. *Phytochem Rev* 2019, **18**:241-272.

23. Zhalnina K, Louie KB, Hao Z, Mansoori N, da Rocha UN, Shi S, Cho H, Karaoz U, Loqué D, Bowen BP *et al.*: **Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly**. *Nat Microbiol* 2018, **3**:470-480.

24. Ube N, Yabuta Y, Tohnooka T, Ueno K, Taketa S, Ishihara A: **Biosynthesis of phenylamide phytoalexins in pathogen-infected barley**. *Int J Mol Sci* 2019, **20**.

25. Li Z, Chen Y, Meesapyodsuk D, Qiu X: **The biosynthetic pathway of major avenanthramides in oat**. *Metabolites* 2019, **9**.

26. Asselin JAE, Lin J, Perez-Quintero AL, Gentzel I, Majerczak D, Opiyo SO, Zhao W, Paek S-M, Kim MG, Coplin DL *et al.*: **Perturbation of Maize phenylpropanoid metabolism by an AvrE family type III effector from *Pantoea stewartii***. *Plant Physiol* 2015, **167**:1117-1135.

27. Zhou S, Zhang YK, Kremling KA, Ding Y, Bennett JS, Bae JS, Kim DK, Ackerman HH, Kolomiets MV, Schmelz EA: **Ethylene signaling regulates natural variation in the abundance of antifungal acetylated diferuloylsucroses and *Fusarium graminearum* resistance in maize seedling roots**. *New Phytol* 2019, **221**:2096-2111.

28. • Zhou S, Kremling KA, Bandillo N, Richter A, Zhang YK, Ahern KR, Artyukhin AB, Hui JX, Younkin GC, Schroeder FC *et al.*: **Metabolome-scale genome-wide association studies reveal chemical diversity and genetic control of maize specialized metabolites**. *Plant Cell* 2019, **31**:937-955

The authors analyzed maize chemical diversity, namely 3991 mass features in 282 inbreds, and established a public resource of GWAS results linking metabolites with underlying maize biosynthetic and/or regulatory genes.

29. Karunanithi PS, Zerbe P: **Terpene synthases as metabolic gatekeepers in the evolution of plant terpenoid chemical diversity**. *Front Plant Sci* 2019, **10**:1166.

30. •• Ding Y, Murphy KM, Poretsky E, Mafu S, Yang B, Char SN, Christensen SA, Saldivar E, Wu M, Wang Q *et al.*: **Multiple genes recruited from hormone pathways partition maize diterpenoid defences**. *Nat Plants* 2019, **5**:1043-1056
The authors used metabolomic profiling, mGWAS, heterologous *Agrobacterium*-mediated enzyme coexpression in *N. benthamiana* and the analyses of diverse mutants to identify all maize kauralexin biosynthetic enzymes. The study revealed a circuitous biosynthetic pathway for *ent*-kaurane-related antibiotics that avoids the production of direct growth hormone precursors.

31. Muchlinski A, Chen X, Lovell JT, Köllner TG, Pelot KA, Zerbe P, Ruggiero M, Callaway L 3rd, 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* 2019, **10**:1144.

32. Fu J, Ren F, Lu X, Mao H, Xu M, Degenhardt J, Peters RJ, Wang Q: **A tandem array of *ent*-kaurene synthases in maize with roles in gibberellin and more specialized metabolism**. *Plant Physiol* 2016, **170**:742-751.

33. Robert CAM, Erb M, Duployer M, Zwahlen C, Doyen GR, Turlings TCJ: **Herbivore-induced plant volatiles mediate host selection by a root herbivore**. *New Phytol* 2012, **194**:1061-1069.

34. Murphy KM, Ma L-T, Ding Y, Schmelz EA, Zerbe P: **Functional characterization of two class II diterpene synthases indicates additional specialized diterpenoid pathways in maize (*Zea mays*)**. *Front Plant Sci* 2018, **9**:1542.

35. Schmelz EA, Huffaker A, Sims JW, Christensen SA, Lu X, Okada K, Peters RJ: **Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins**. *Plant J* 2014, **79**:659-678.

36. Lu X, Zhang J, Brown B, Li R, Rodríguez-Romero J, Berasategui A, Liu B, Xu M, Luo D, Pan Z *et al.*: **Inferring roles in defense from metabolic allocation of rice diterpenoids**. *Plant Cell* 2018, **30**:1119-1131.

37. •• De La Peña R, Sattely ES: **Rerouting plant terpene biosynthesis enables momilactone pathway elucidation**. *Nat Chem Biol* 2021, **17**:205-212
The authors finalized discovery of the rice momilactone biosynthetic pathway and engineered a strategy for dramatically improved heterologous diterpene production in *N. benthamiana* by rerouting biosynthesis from the chloroplast to the cytosol.

38. • Zhan C, Lei L, Liu Z, Zhou S, Yang C, Zhu X, Guo H, Zhang F, Peng M, Zhang M *et al.*: **Selection of a subspecies-specific diterpene gene cluster implicated in rice disease resistance**. *Nat Plants* 2020, **6**:1447-1454
The authors employed mGWAS to identify a rice diterpenoid gene cluster responsible for the biosynthesis of a jasmnate-regulated monocyclic diterpenoid, 5,10-diketo-casbene.

39. Karunanithi PS, Berrios DI, Wang SD, Davis J, Shen T, Fiehn O, Maloof JN, Zerbe P: **The foxtail millet (*Setaria italica*) terpene synthase gene family**. *Plant J* 2020, **103**:781-800.

40. Pelot KA, Chen R, Hagelthorn DM, Young CA, Addison JB, Muchlinski A, Tholl D, Zerbe P: **Functional diversity of diterpene synthases in the biofuel crop switchgrass**. *Plant Physiol* 2018, **178**:54-71.

41. Egenolf K, Conrad J, Schöne J, Braunberger C, Beifuß U, Walker F, Nuñez J, Arango J, Karwat H, Cadisch G *et al.*: **Brachialactone isomers and derivatives of *Brachiaria humidicola* reveal contrasting nitrification inhibiting activity**. *Plant Physiol Biochem* 2020, **154**:491-497.

42. Charnikhova TV, Gaus K, Lumbroso A, Sanders M, Vincken J-P, De Mesmaeker A, Ruyter-Spira CP, Screpanti C, Bouwmeester HJ: **Zeapyranolactone – a novel strigolactone from maize**. *Phytochem Lett* 2018, **24**:172-178.

43. Turner TR, Ramakrishnan K, Walshaw J, Heavens D, Alston M, Swarbreck D, Osbourn A, Grant A, Poole PS: **Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants**. *ISME J* 2013, **7**:2248-2258.

44. Li X, Jones AD, Last RL: **Switchgrass metabolomics reveals striking genotypic and developmental differences in saponins**. *bioRxiv* 2020 http://dx.doi.org/10.1101/2020.06.01.127720. 2020.06.01.127720.

45. Niemeyer HM: **Hydroxamic acids derived from 2-hydroxy-2H-1,4-benzoxazin-3(4H)-one: key defense chemicals of cereals**. *J Agric Food Chem* 2009, **57**:1677-1696.

46. Ahmad S, Veyrat N, Gordon-Weeks R, Zhang YH, Martin J, Smart L, Glauser G, Erb M, Flors V, Frey M *et al.*: **Benzoxazinoid metabolites regulate innate immunity against aphids and fungi in maize**. *Plant Physiol* 2011, **157**:317-327.

47. Hu L, Mateo P, Ye M, Zhang X, Berset JD, Handrick V, Radisch D, Grabe V, Kollner TG, Gershenzon J *et al.*: **Plant iron acquisition strategy exploited by an insect herbivore**. *Science* 2018, **361**:694-697.

48. Wasternack C, Feussner I: **The oxylipin pathways: biochemistry and function**. *Annu Rev Plant Biol* 2018, **69**:363-386.

49. Christensen SA, Huffaker A, Kaplan F, Sims J, Ziemann S, Doehlemann G, Ji LX, Schmitz RJ, Kolomiets MV, Alborn HT *et al.*: **Maize death acids, 9-lipoxygenase-derived cyclopente(a)nones, display activity as cytotoxic phytoalexins and transcriptional mediators**. *Proc Natl Acad Sci U S A* 2015, **112**:11407-11412.

50. • Wang K-D, Borrego EJ, Kenerley CM, Kolomiets MV: **Oxylipins other than jasmonic acid are xylem-resident signals regulating systemic resistance induced by *Trichoderma virens* in maize**. *Plant Cell* 2020, **32**:166-185
The authors demonstrate that root *Trichoderma virens* interactions elicit production and increased xylem levels of 12-oxo-phytodienoic acid and the α-ketol product of octadecadienoic acid that function as symbiont-induced systemic resistance (ISR) signals in maize.

51. de Oliveira IF, Simeone MLF, de Guimarães CC, Garcia NS, Schaffert RE, de Sousa SM: **Sorgoleone concentration influences mycorrhizal colonization in sorghum**. *Mycorrhiza* 2021, **31**:259-264.

52. Sarr PS, Ando Y, Nakamura S, Deshpande S, Subbarao GV: **Sorgoleone release from sorghum roots shapes the composition of nitrifying populations, total bacteria, and archaea and determines the level of nitrification**. *Biol Fertil Soils* 2020, **56**:145-166.

53. Jacoby RP, Koprivova A, Kopriva S: **Pinpointing secondary metabolites that shape the composition and function of the plant microbiome**. *J Exp Bot* 2021, **72**:57-69.

54. Christensen SA, Sims J, Vaughan MM, Hunter C, Block A, Willett D, Alborn HT, Huffaker A, Schmelz EA: **Commercial hybrids and mutant genotypes reveal complex protective roles for inducible terpenoid defenses in maize**. *J Exp Bot* 2018, **69**:1693-1705.

55. Murphy KM, Edwards J, Louie KB, Bowen BP, Sundaresan V, Northen TR, Zerbe P: **Bioactive diterpenoids impact the composition of the root-associated microbiome in maize (*Zea mays*)**. *Sci Rep* 2021, **11**:333.

56. Huang AC, Jiang T, Liu Y-X, Bai Y-C, Reed J, Qu B, Goossens A, Nützmann H-W, Bai Y, Osbourn A: **A specialized metabolic network selectively modulates Arabidopsis root microbiota**. *Science* 2019, **364**:eaau6389.

57. Lebeis SL, Paredes SH, Lundberg DS, Breakfield N, Gehring J, McDonald M, Malfatti S, Glavina del Rio T, Jones CD, Tringe SG *et al.*: **PLANT MICROBIOME. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa**. *Science* 2015, **349**:860-864.

58. Salem MA, Perez de Souza L, Serag A, Fernie AR, Farag MA, Ezzat SM, Alseekh S: **Metabolomics in the context of plant natural products research: from sample preparation to metabolite analysis**. *Metabolites* 2020, **10**:37.

59. • Dührkop K, Nothias LF, Fleischauer M, Reher R, Ludwig M, Hoffmann MA, Petras D, Gerwick WH, Rousu J, Dorrestein PC *et al.*: **Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra**. *Nat Biotechnol* 2021, **39**:462-471
The authors created a computational tool, termed CANOPUS, for systematic compound classification from MS fragmentation spectra. With reference data, CANOPUS outperforms numerous existing methods suggesting broad utility in accurate compound class annotation from MS–MS spectra.

60. Bayer PE, Golicz AA, Scheben A, Batley J, Edwards D: **Plant pan-genomes are the new reference**. *Nat Plants* 2020, **6**:914-920.

61. Tiedge K, Muchlinski A, Zerbe P: **Genomics-enabled analysis of specialized metabolism in bioenergy crops: current progress and challenges**. *Synth Biol (Oxf)* 2020, **5**:ysaa005.

62. Timp W, Timp G: **Beyond mass spectrometry, the next step in proteomics**. *Sci Adv* 2020, **6**:eaax8978.

63. Wisecaver JH, Borowsky AT, Tzin V, Jander G, Kliebenstein DJ, Rokas A: **A global coexpression network approach for connecting genes to specialized metabolic pathways in plants**. *Plant Cell* 2017, **29**:944-959.

64. Poretsky E, Huffaker A: **MutRank: an R shiny web-application for exploratory targeted mutual rank-based coexpression analyses integrated with user-provided supporting information**. *PeerJ* 2020, **8**.

65. • Kitaoka N, Zhang J, Oyagbenro RK, Brown B, Wu Y, Yang B, Li Z, Peters RJ: **Interdependent evolution of biosynthetic gene clusters for momilactone production in rice**. *Plant Cell* 2021, **33**:290-305
The authors use gene-coexpression analysis, gene cluster analyses and genetics to demonstrate that a rice cytochrome P450, namely CYP76M8, is important for momilactone biosynthesis.

66. Desmet S, Brouckaert M, Boerjan W, Morreel K: **Seeing the forest for the trees: retrieving plant secondary biochemical pathways from metabolome networks**. *Comput Struct Biotechnol J* 2021, **19**:72-85.

67. Frey M, Chomet P, Glawischnig E, Stettner C, Grun S, Winklmair A, Eisenreich W, Bacher A, Meeley RB, Briggs SP *et al.*: **Analysis of a chemical plant defense mechanism in grasses**. *Science* 1997, **277**:696-699.

68. Field B, Osbourn AE: **Metabolic diversification-independent assembly of operon-like gene clusters in different plants**. *Science* 2008, **320**:543-547.

69. •• Moore BM, Wang P, Fan P, Leong B, Schenck CA, Lloyd JP, Lehti-Shiu MD, Last RL, Pichersky E, Shiu SH: **Robust predictions of specialized metabolism genes through machine learning**. *Proc Natl Acad Sci U S A* 2019, **116**:2344-2353
The authors developed a computational model using machine learning methods which were applied to diverse datasets and metrics to predict Arabidopsis genes involved in specialized metabolism. Over 1000 genes with previously unknown functions where further predicted to contribute to specialized metabolism.

70. Fang C, Luo J: **Metabolic GWAS-based dissection of genetic bases underlying the diversity of plant metabolism**. *Plant J* 2019, **97**:91-100.

71. Liu HJ, Yan J: **Crop genome-wide association study: a harvest of biological relevance**. *Plant J* 2019, **97**:8-18.

72. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES: **TASSEL: software for association mapping of complex traits in diverse samples**. *Bioinformatics* 2007, **23**:2633-2635.

73. Monier B, Casstevens TM, Buckler ES: **rTASSEL: an R interface to TASSEL for association mapping of complex traits**. *bioRxiv* 2020 http://dx.doi.org/10.1101/2020.07.21.209114. 2020.07.21.209114.

74. Hirsch CN, Foerster JM, Johnson JM, Sekhon RS, Muttoni G, Vaillancourt B, Penagaricano F, Lindquist E, Pedraza MA, Barry K *et al.*: **Insights into the maize pan-genome and pan-transcriptome**. *Plant Cell* 2014, **26**:121-135.

75. Franco JAV, Gage JL, Bradbury PJ, Johnson LC, Miller ZR, Buckler ES, Romay MC: **A maize practical haplotype graph leverages diverse NAM assemblies**. *bioRxiv* 2020. 2020.2008.2031.268425.

76. •• Voichek Y, Weigel D: **Identifying genetic variants underlying phenotypic variation in plants without complete genomes**. *Nat Genet* 2020, **52**:534-540
The authors use short DNA sequences, termed k-mers, in maize GWAS studies to powerfully link a wider range of genetic variations to phenotypic variations. Importantly progress using this approach is possible in plants without completed genomes.

77. Schaefer RJ, Michno JM, Jeffers J, Hoekenga O, Dilkes B, Baxter I, Myers CL: **Integrating coexpression networks with GWAS to prioritize causal genes in maize**. *Plant Cell* 2018, **30**:2922-2942.

78. Irmisch S, Jo S, Roach CR, Jancsik S, Man Saint Yuen M, Madilao LL, O'Neil-Johnson M, Williams R, Withers SG, Bohlmann J: **Discovery of UDP-glycosyltransferases and BAHD-acyltransferases involved in the biosynthesis of the antidiabetic plant metabolite montbretin A**. *Plant Cell* 2018, **30**:1864-1886.

79. Blaby IK, Cheng J-F: **Building a custom high-throughput platform at the Joint Genome Institute for DNA construct design and assembly—present and future challenges**. *Synth Biol* 2020, **5**.

80. Bally J, Jung H, Mortimer C, Naim F, Philips JG, Hellens R, Bombarely A, Goodin MM, Waterhouse PM: **The rise and rise of *Nicotiana benthamiana*: a plant for all reasons**. *Annu Rev Phytopathol* 2018, **56**:405-426.

81. Sainsbury F, Thuenemann EC, Lomonossoff GP: **pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants**. *Plant Biotechnol J* 2009, **7**:682-693.

82. Bach SS, Bassard J, Andersen-Ranberg J, Møldrup ME, Simonsen HT, Hamberger B: **High-throughput testing of terpenoid biosynthesis candidate genes using transient expression in *Nicotiana benthamiana***. *Methods Mol Biol* 2014, **1153**:245-255.

83. Reed J, Stephenson MJ, Miettinen K, Brouwer B, Leveau A, Brett P, Goss RJM, Goossens A, O'Connell MA, Osbourn A: **A translational synthetic biology platform for rapid access to gram-scale quantities of novel drug-like molecules**. *Metab Eng* 2017, **42**:185-193.

84. Marcon C, Altrogge L, Win YN, Stöcker T, Gardiner JM, Portwood JL 2nd, Opitz N, Kortz A, Baldauf JA, Hunter CT *et al.*: **BonnMu: a sequence-indexed resource of transposon-induced maize mutations for functional genomics studies**. *Plant Physiol* 2020, **184**:620-631.

85. Zhu H, Li C, Gao C: **Applications of CRISPR-Cas in agriculture and plant biotechnology**. *Nat Rev Mol Cell Biol* 2020, **21**:661-677.

86. Mei Y, Whitham SA: **Virus-induced gene silencing in maize with a foxtail mosaic virus vector**. *Methods Mol Biol* 2018, **1676**:129-139.

87. Shan S, Soltis PS, Soltis DE, Yang B: **Considerations in adapting CRISPR/Cas9 in nongenetic model plant systems**. *Appl Plant Sci* 2020, **8**:e11314.

88. Chung SH, Bigham M, Lappe RR, Chan B, Nagalakshmi U, Whitham SA, Dinesh-Kumar SP, Jander G: **A sugarcane mosaic virus vector for rapid in planta screening of proteins that inhibit the growth of insect herbivores**. *Plant Biotechnol J* 2021 http://dx.doi.org/10.1111/pbi.13585. (Online ahead of print) PMID: 33763921.

89. Pott DM, Durán-Soria S, Osorio S, Vallarino JG: **Combining metabolomic and transcriptomic approaches to assess and improve crop quality traits**. *CABI Agric Biosci* 2021, **2**:1.

90. Wada N, Ueta R, Osakabe Y, Osakabe K: **Precision genome editing in plants: state-of-the-art in CRISPR/Cas9-based genome engineering**. *BMC Plant Biol* 2020, **20**:234.

91. Baetsen-Young A, Chen H, Shiu S-H, Day B: **Contrasting transcriptional responses to *Fusarium virguliforme***

colonization in symptomatic and asymptomatic hosts. *Plant Cell* 2021, **33**:224-247.

92. Poretsky E, Dressano K, Weckwerth P, Ruiz M, Char SN, Shi D, Abagyan R, Yang B, Huffaker A: **Differential activities of maize plant elicitor peptides as mediators of immune signaling and herbivore resistance**. *Plant J* 2020, **104**:1582-1602.

93. • Zhang L, Gleason C: **Enhancing potato resistance against root-knot nematodes using a plant-defence elicitor delivered by bacteria**. *Nat Plants* 2020, **6**:625-629
The authors demonstrated that engineered *Bacillus subtilis* bacteria can deliver secreted Plant elicitor peptide (Pep) signals, specifically StPep1, to potato roots to elicit inducible immune responses for improved root-knot nematodes resistance.

94. Robert CA, Erb M, Hiltpold I, Hibbard BE, Gaillard MD, Bilat J, Degenhardt J, Cambet-Petit-Jean X, Turlings TC, Zwahlen C: **Genetically engineered maize plants reveal distinct costs and benefits of constitutive volatile emissions in the field**. *Plant Biotechnol J* 2013, **11**:628-639.

95. Chen X, Chen H, Yuan JS, Köllner TG, Chen Y, Guo Y, Zhuang X, Chen X, Zhang YJ, Fu J *et al.*: **The rice terpene synthase gene OsTPS19 functions as an (S)-limonene synthase in planta, and its overexpression leads to enhanced resistance to the blast fungus *Magnaporthe oryzae***. *Plant Biotechnol J* 2018, **16**:1778-1787.

96. •• Zhou P, Li Z, Magnusson E, Gomez Cano F, Crisp PA, Noshay JM, Grotewold E, Hirsch CN, Briggs SP, Springer NM: **Meta gene regulatory networks in maize highlight functionally relevant regulatory interactions**. *Plant Cell* 2020, **32**:1377-1396
The authors reanalyzed over 6000 publically available maize RNA-seq samples to create gene regulatory networks (GRNs) linking transcription factors with target genes. This study shows the potential of GRN analyses in identifying candidate regulators of specialized metabolic pathways.

97. Yoshida Y, Miyamoto K, Yamane H, Nishizawa Y, Minami E, Nojiri H, Okada K: **OsTGAP1 is responsible for JA-inducible diterpenoid phytoalexin biosynthesis in rice roots with biological impacts on allelopathic interaction**. *Physiol Plant* 2017, **161**:532-544.

98. • Shen QQ, Pu QY, Liang J, Mao HJ, Liu J, Wang Q: **CYP71Z overexpression confers elevated blast resistance in transgenic rice**. *Plant Mol Biol* 2019, **100**:579-589
The authors leveraged a highly substrate promiscuous maize terpenoid antibiotic pathway enzyme, namely ZmCYP71Z18, to create stable transgenic rice heterologous overexpression lines resulting in novel diterpenoid antibiotics for improved disease resistance. This study highlights the utility of mixing monocot defense pathways to engineer new biochemicals for enhanced stress resistance.

99. Calgaro-Kozina A, Vuu KM, Keasling JD, Loqué D, Sattely ES, Shih PM: **Engineering plant synthetic pathways for the biosynthesis of novel antifungals**. *ACS Cent Sci* 2020, **6**:1394-1400.