nature plants

ARTICLES

https://doi.org/10.1038/s41477-020-00787-9

Genetic elucidation of interconnected antibiotic pathways mediating maize innate immunity

Yezhang Ding[1,15], Philipp R. Weckwerth[1,15], Elly Poretsky[1], Katherine M. Murphy[2], James Sims[3], Evan Saldivar[1], Shawn A. Christensen[4], Si Nian Char[5], Bing Yang[5,6], Anh-dao Tong[1], Zhouxin Shen[1], Karl A. Kremling[7], Edward S. Buckler[7,8], Tom Kono[9], David R. Nelson[10], Jörg Bohlmann[11], Matthew G. Bakker[12,13], Martha M. Vaughan[12], Ahmed S. Khalil[1], Mariam Betsiashvili[1], Keini Dressano[1], Tobias G. Köllner[14], Steven P. Briggs[1], Philipp Zerbe[2], Eric A. Schmelz[1] and Alisa Huffaker[1]✉

Specialized metabolites constitute key layers of immunity that underlie disease resistance in crops; however, challenges in resolving pathways limit our understanding of the functions and applications of these metabolites. In maize (*Zea mays*), the inducible accumulation of acidic terpenoids is increasingly considered to be a defence mechanism that contributes to disease resistance. Here, to understand maize antibiotic biosynthesis, we integrated association mapping, pan-genome multi-omic correlations, enzyme structure–function studies and targeted mutagenesis. We define ten genes in three zealexin (*Zx*) gene clusters that encode four sesquiterpene synthases and six cytochrome P450 proteins that collectively drive the production of diverse antibiotic cocktails. Quadruple mutants in which the ability to produce zealexins (ZXs) is blocked exhibit a broad-spectrum loss of disease resistance. Genetic redundancies ensuring pathway resiliency to single null mutations are combined with enzyme substrate promiscuity, creating a biosynthetic hourglass pathway that uses diverse substrates and in vivo combinatorial chemistry to yield complex antibiotic blends. The elucidated genetic basis of biochemical phenotypes that underlie disease resistance demonstrates a predominant maize defence pathway and informs innovative strategies for transferring chemical immunity between crops.

At present, 50% of global arable land is allocated to agriculture. In the absence of yearly improvements in cereal germplasm and productivity, comparable land consumption today would be greater than 60% (refs. [1,2]). Given human reliance on maize (*Zea mays*) and a few related grasses, genetically encoded mechanisms that provide stress protection have been sought[3–5]. In particular, fungal diseases—such as those caused by *Fusarium* species including *Fusarium graminearum*—are widely devastating to poaceous crops, and further result in grain contamination with harmful mycotoxins[6,7]. Understanding the innate immune responses, genetic variation and endogenous pathway interactions that underlie broad-spectrum disease resistance[8] represents foundational knowledge that is necessary for sustained crop improvement and trait optimization.

Plants are protected from pest and pathogen attack by interconnected layers of physical barriers, pattern-recognition receptors, defence proteins and bioactive specialized metabolites[9–11]. Specialized metabolic pathways are often unique to individual species, display specificity in regulated production and mediate cryptic, albeit impactful, phenotypes[9,12]. Benzoxazinoids (BXs) are the most broadly shared and widely studied poaceous chemical defences. Constitutively produced in seedlings, BXs contribute to resistance against insects and fungi, such as northern corn leaf blight (*Setosphaeria turcica*)[5,13–15]. In contrast to BXs and other largely constitutive defences that are present before attack, many specialized metabolites are produced exclusively on demand, are highly localized to the site of challenge and often evade analytical detection[16]. Maize relies on a combination of dynamically regulated BXs, phenylpropanoids and terpenoids for biotic and abiotic stress protection[5,17–19]. Although the biosynthesis and roles of BXs and terpene volatiles in anti-herbivore defences are increasingly understood[13,20], the genetic and biochemical complexities that underlie maize protection against fungal pathogens have remained a challenge to resolve[18,21–23].

Terpenoids are the most structurally diverse class of plant specialized metabolites, and are typically produced from the combined activities of terpene synthase (TPS) and cytochrome P450 monooxygenase (P450) enzymes[24,25]. Many catalytic activities and biological roles have been assigned to the 43 TPSs encoded in the maize B73 genome[26,27]. Known maize terpenoid antibiotics include

[1]Section of Cell and Developmental Biology, University of California at San Diego, La Jolla, CA, USA. [2]Department of Plant Biology, University of California Davis, Davis, CA, USA. [3]ETH Zurich, Institute of Agricultural Sciences, Zurich, Switzerland. [4]Chemistry Research Unit, Center for Medical, Agricultural and Veterinary Entomology, Department of Agriculture, Agricultural Research Service, Gainesville, FL, USA. [5]Division of Plant Sciences, Bond Life Sciences Center, University of Missouri, Columbia, MO, USA. [6]Donald Danforth Plant Science Center, St Louis, MO, USA. [7]Department of Plant Breeding and Genetics, Cornell University, Ithaca, NY, USA. [8]Robert W. Holley Center for Agriculture and Health, Ithaca, US Department of Agriculture, Agricultural Research Service, New York, NY, USA. [9]Minnesota Supercomputing Institute, University of Minnesota, Minneapolis, MN, USA. [10]University of Tennessee Health Science Center, Memphis, TN, USA. [11]Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, Canada. [12]National Center for Agricultural Utilization Research, US Department of Agriculture, Agricultural Research Service, Peoria, IL, USA. [13]Department of Microbiology, University of Manitoba, Winnipeg, Manitoba, Canada. [14]Max Planck Institute for Chemical Ecology, Jena, Germany. [15]These authors contributed equally: Yezhang Ding, Philipp R. Weckwerth. ✉e-mail: ahuffaker@ucsd.edu

sesquiterpenoids, which are represented by zealexins (ZXs) and α/β-costic acids[28,29], and diterpenoids, such as kauralexins and dolabralexins[18,30]. Although enzymes that are active in the oxidation of β-selinene to β-costic acid remain unknown, separate studies have demonstrated that both *Zm*CYP71Z16 and *Zm*CYP71Z18 have the ability to oxidize diverse hydrocarbon olefins to yield ZXs, kauralexins and dolabralexins[18,30,31]. Despite advances in elucidating the diterpenoid biosynthetic pathways[18], acidic sesquiterpenoid derivatives of β-macrocarpene, termed ZXs, represent the single-largest class of defensive terpenoids that are known in the genus *Zea*[29], and yet remain the least understood. Highly localized to the site of elicitation, ZX accumulation correlates with the expression of the β-macrocarpene synthase-encoding transcripts Zm*TPS6* and Zm*TPS11*, which display substantial increases after challenge with diverse fungal pathogens[29,32–34]. Consistent with their roles in crop protection, virus-induced gene silencing of Zm*TPS6*/*11* revealed that they are required for restricting smut fungus (*Ustilago maydis*) infection and tumour formation[35]. Although correlations between Zm*TPS6*/*11* transcripts, ZX production and fungal resistance exist[29,35], no single biosynthetic pathway node has been proven in planta, and all known maize lines produce ZXs. Moreover, the structural diversity of ZXs, range of ZX pathway genes and endogenous protective functions remain unresolved[30,31]. Recent advances in -omic tools, coregulation analyses, genetic resources, in vivo protein biochemistry and gene-editing approaches now enable the critical examination and engineering the complex protective pathways that underlie crop resistance.

To define the genetic basis and pan-genome complexity that underlie maize biochemical immunity, here we identified 17 metabolites that are products of the core *Zx* gene network (*Zx1* to *Zx10*). This network consists of three functionally distinct gene clusters that encode TPSs responsible for hydrocarbon olefin production and P450s in the CYP71Z and CYP81A families that facilitate oxygenation and desaturation. ZXs occur within a larger interconnected metabolic network that also produces kauralexins, dolabralexins and α/β-costic acids. Diverse precursors for each family converge on a single node of catalytically promiscuous CYP71Z enzymes that act on multiple endogenous substrates with partially overlapping functions to generate diverse products that can be subsequently decorated further by pathway-specific enzymes. The result is a broad cocktail of metabolites that are generated by a modest number of enzymes through an hourglass-shaped pathway. Metabolic complexity is only one layer of induced responses—after fungal challenge, ZX production co-occurs with changes in half of the measurable proteome. Despite this complexity, analyses of *zx1 zx2 zx3 zx4* quadruple mutants, which are abolished in ZX biosynthesis, demonstrate substantial protective roles against multiple pathogens. Given that all plants produce terpenoid precursors, the promiscuous enzyme activities described here are amenable to genetic transfer. A foundational understanding of genetic and biochemical mechanisms in maize lays the groundwork for diversifying the chemical defences that underlie disease-resistance phenotypes in phylogenetically distant grain crops[36].

Results

Maize harbours a functionally variable gene cluster of four β-macrocarpene synthases. Two pathogen-regulated β-macrocarpene synthases, termed TPS 6 and TPS 11, were previously assigned as the B73 (RefGen_V4) genes *Zm00001d024207* and *Zm00001d024210* (www.maizegdb.org), respectively[32,33]. Current indirect evidence supports that *Zm*TPS6/11 has a role in the production of diverse antibiotics, termed ZXs[29,34]. As a visual aid, we summarized the biochemicals, genes and proteins with examined relevance to the ZX pathway that are present in the genus *Zea* (Supplementary Figs. 1 and 2 and Supplementary Tables 1–4). Analyses of the B73 genome for all TPSs reveal that Zm*TPS6*/*11* are components of a four-gene cluster[37] on chromosome 10 (Fig. 1a), sharing >84% protein identity with each other (Supplementary Figs. 3 and 4). Following the BX pathway nomenclature[5], we adopted unified B73 ZX pathway abbreviations, starting with *Zx1*/*TPS6* (*Zm00001d024207*), *Zx2*/*TPS12* (*Zm00001d024208*), *Zx3*/*TPS11* (*Zm00001d024210*) and *Zx4*/*TPS13* (*Zm00001d024211*) on the basis of sequential chromosome order (Fig. 1a,b). Unless otherwise noted, gene and protein abbreviations refer to B73 (RefGen_V4) reference sequences. RNA sequencing (RNA-seq) analyses of stems from B73, Mo17 and W22 inbred lines, damaged either alone or additionally treated with heat-killed *Fusarium*, demonstrate that fungal-elicited transcript accumulation occurs in an inbred-specific manner (Fig. 1c–e and Supplementary Table 2). To understand the contribution of *Zx* gene cluster I to the production of β-bisabolene and β-macrocarpene, individual genes *Zx1* to *Zx4* from both B73 and W22 lines (Supplementary Figs. 4 and 5 and Supplementary Table 4) were functionally analysed using transient, *Agrobacterium*-mediated expression in *Nicotiana benthamiana*. Expression of B73 (ZX1, ZX3, ZX4) and W22 (ZX2, ZX3, ZX4) revealed similar, albeit different, combinations of functional β-macrocarpene synthases, yielding low levels of β-bisabolene (Fig. 1f,g). In both cases, the lack of β-bisabolene and β-macrocarpene production by W22 ZX1 and B73 ZX2 occurred during otherwise sufficient transcript expression (Supplementary Fig. 5).

To understand common mutations that cause a loss of function in ZX1 to ZX4, we examined amino acid sequence variations in W22 ZX1 and we observed a W274R substitution that is predicted to negatively impact the catalytic site[38] (Supplementary Fig. 4). Reversion of inactive W22 ZX1 back to R274W or mutation of active W22 ZX4 (R274) to W274R reactivated and inactivated, respectively, the enzymes in *N. benthamiana* transient expression assays (Supplementary Fig. 6). W22-like *Zx1* non-synonymous single-nucleotide polymorphisms (SNPs) at chromosome 10 position 56448050 (A to G) that underlie the ZX1 W274R null mutation are common in maize germplasm and are present in >10% of examined inbreds[39,40] (Supplementary Table 5). Analysis of *Zx* gene cluster I in 28 maize genomes demonstrates that predicted β-macrocarpene synthase genes typically exist as four copies but can vary from three to six copies in select inbred lines (Supplementary Figs. 7 and 8, and Supplementary Tables 6 and 7). Beyond functional variation in *Zx1*, we observed that B73 and eight additional inbred lines (32%) contain *Zx2* mutations that are predicted to result in non-functional enzymes (Supplementary Tables 6 and 7). In our analyses here, the lack of significant Mo17 *Zx2* transcript after elicitation (Fig. 1e) corresponds to the absence of detectable Mo17 *Zx2* nucleotides on the basis of gene cluster analyses (Supplementary Fig. 7).

To demonstrate endogenous relationships, we used mutual-rank-(MR)-based global gene coexpression analyses to associate transcriptional patterns of maize sesquiterpene synthases in a large RNA-seq dataset[40]. Analyses revealed that the highest degree of coregulation between *Zx1* and *Zx3*, and between *Zx2* and *Zx4* together with partial Zm*TPS21* coregulation is responsible for β-selinene derived antibiotics[28] (Fig. 1h). Genome-wide analyses of *Zx1* to *Zx4* expression levels in diverse inbred lines are consistent with the complex patterns (Fig. 1i) witnessed in B73, Mo17 and W22 (Fig. 1c–e). Analyses of *Zx* gene cluster I demonstrate that inbred-line-specific combinations of functional ZX1 to ZX4 proteins are common (Supplementary Fig. 7) and contribute to β-bisabolene and β-macrocarpene production (Fig. 1a–i).

A second *Zx* gene cluster contains and encodes three promiscuous CYP71Z-family cytochrome P450s. Increases in *Zx1* to *Zx4* accumulation are among the largest fold changes in transcription after pathogen challenge[29,32]. In an early analysis of stems, we observed that a member of the cytochrome P450 CYP71 family, Zm*CYP71Z18* (NM_001147894), is among the most *F. graminearum*

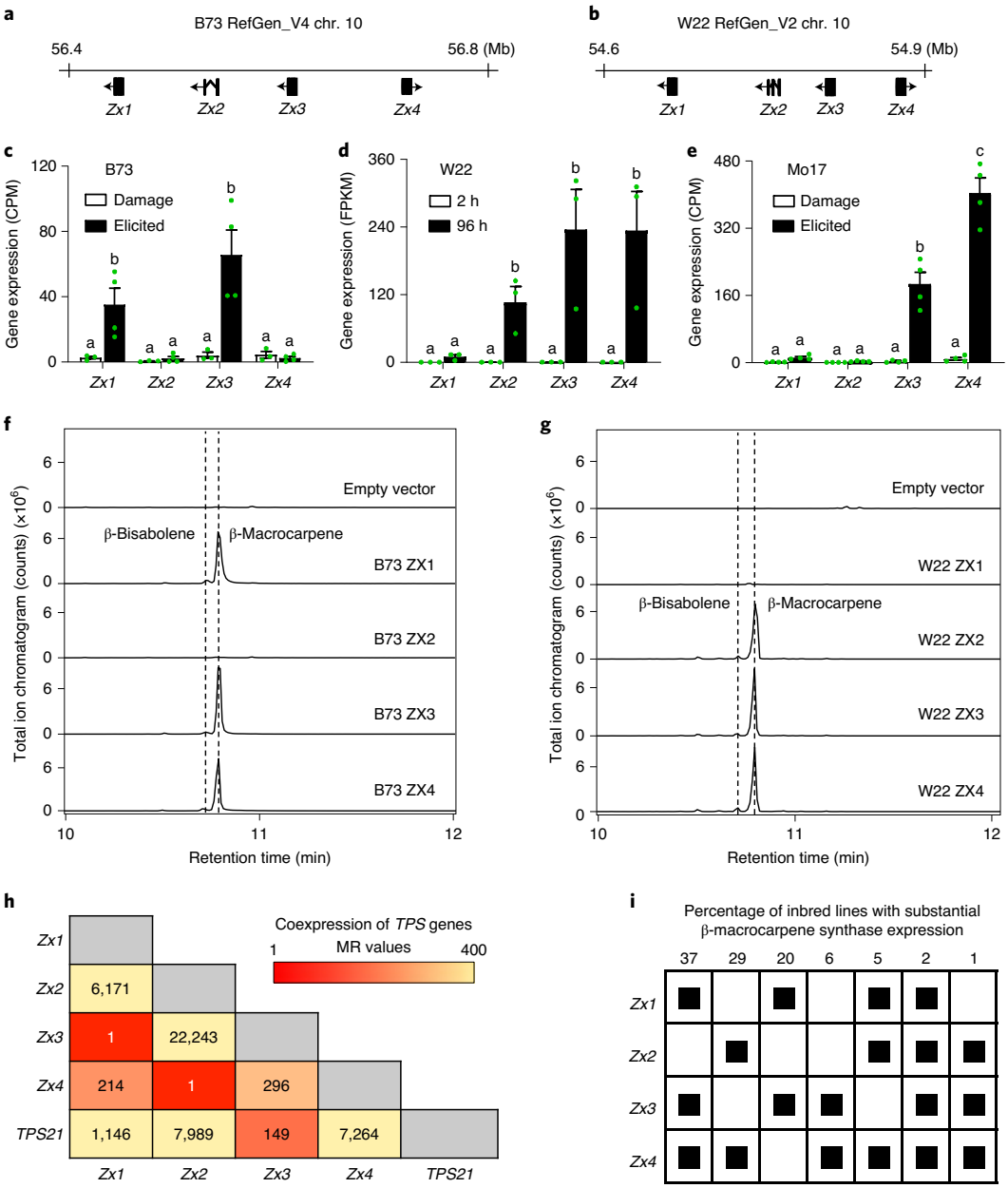


Fig. 1 | A genetically variable cluster of four maize TPSs ensures the production of ZX precursors. a–e, Tandem array of four β-macrocarpene synthase genes, termed *Zx1* to *Zx4*, on chromosome (chr.) 10 of the B73 genome (RefGen_V4) (**a**) and the W22 genome (RefGen_V2) (**b**). *Zx1* to *Zx4* transcript abundance derived from RNA-seq analyses of five-week-old B73 (**c**), W22 (**d**) and Mo17 (**e**) stems damaged (damage) only or additionally treated with heat-killed *F. venenatum* hyphae (elicited). For Mo17 and B73, harvests were performed at 36 h and 3′ RNA-seq gene expression is given as counts per million mapped reads (CPM). For W22 RNA-seq results, the average of three early (0 h, 2 h and 4 h) and late (72 h, 96 h and 120 h) elicitation time points with gene expression is given as fragments per kilobase of transcript per million mapped reads (FPKM). For **c**–**e**, data are mean ± s.e.m. B73 damaged, $n = 3$ biologically independent samples; B73 elicited, $n = 4$ biologically independent samples; Mo17 damaged and elicited, $n = 4$ biologically independent samples. Within plots, different letters (a–c) represent significant differences; statistical analysis was performed using one-way analysis of variance (ANOVA), $P < 0.05$; Tukey tests were used to correct for multiple comparisons, $P < 0.05$. **f**,**g**, GC–MS total ion chromatograms are shown for leaf volatiles emitted after *Agrobacterium*-mediated transient *N. benthamiana* expression assays of B73 (**f**) and W22 (**g**) encoded β-macrocarpene synthases ZX1, ZX2, ZX3 and ZX4. An empty vector was used for the *Agrobacterium*-infiltrated control. Four biological repeats were performed and showed similar results. **h**, Heat map of the coexpression of *Zx1, Zx2, Zx3, Zx4* and an β-selinene synthase (Zm*TPS21*) present in a dataset of 1,960 RNA-seq samples. The low numbers in the squares (<250) indicate supportive MR scores. **i**, Expression matrix of *Zx1* to *Zx4* in 100 inbred lines from a dataset of 1,960 RNA-seq samples. Individual *Zx1, Zx2, Zx3* and *Zx4* transcripts with expression levels of ≥5% of the total sum expression of *Zx1* to *Zx4* were counted as expressed and are represented by filled black squares. The empty white squares indicate low levels of gene expression (<5% of the sum of *Zx1* to *Zx4* expression).

upregulated genes that co-occurred with *Zx1* to *Zx4* members[29]. Subsequently, both *Zm*CYP71Z18 and the adjoining *Zm* CYP71Z16 were shown to be catalytically active in the oxidation of β-macrocarpene to ZA1 (refs. [30,31]). To consider roles for additional P450s, we performed a global gene coexpression analysis of the summed expression of *Zx1* to *Zx4* with all of the predicted maize

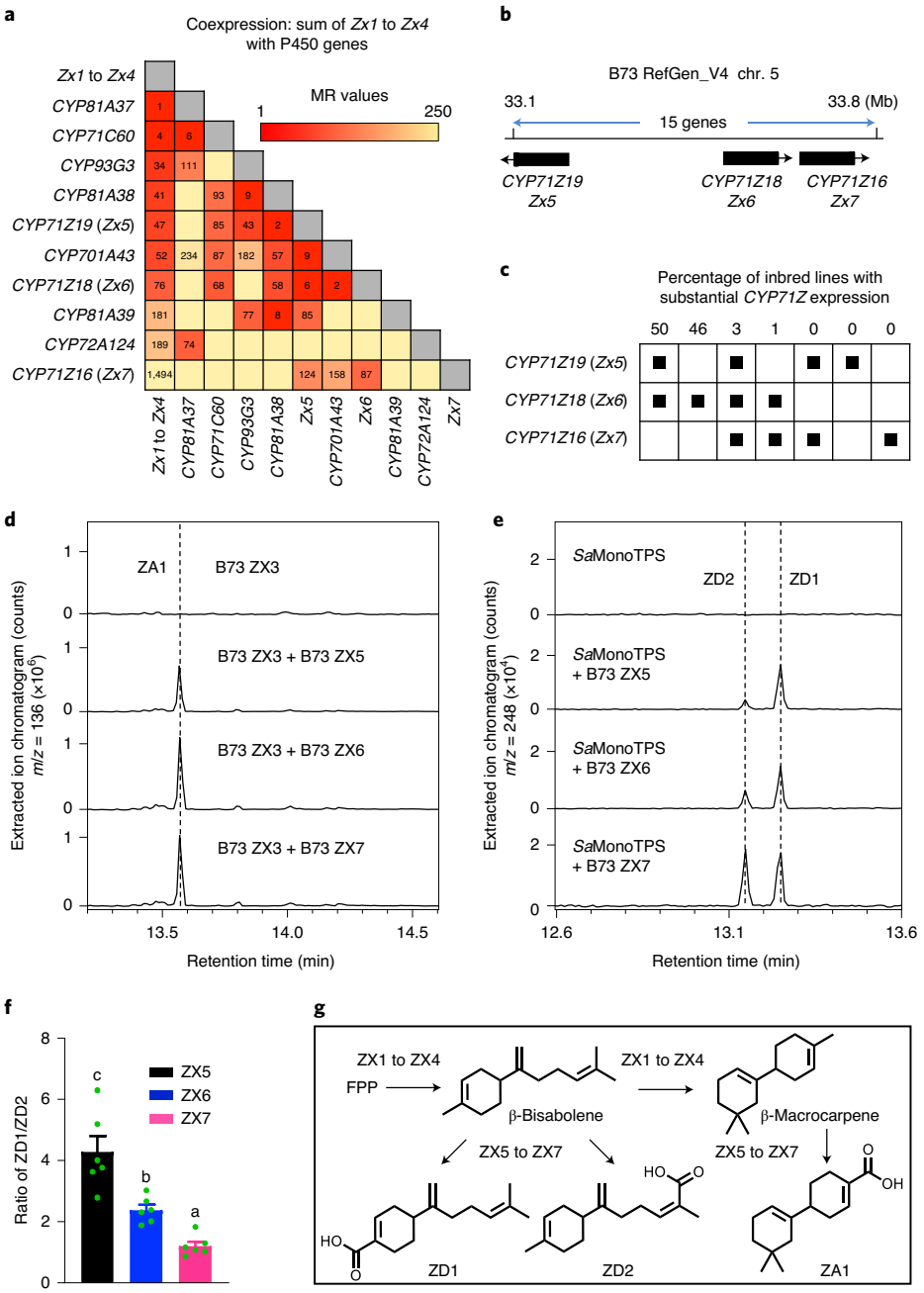


Fig. 2 | *Zx* gene cluster II contains three 71Z-family cytochrome (CYP) P450s that catalyse the production of A- and D-series ZXs. a, Heat map of the summed expression of *Zx1* to *Zx4* and coexpression with all maize P450s in a dataset of 1,960 RNA-seq samples. The low numbers in the squares (<250) indicate supportive MR scores. Weak MR correlations of >250 were omitted. **b**, The physical position of the chromosome 5 gene cluster containing Zm*CYP71Z19* (*Zx5*), Zm*CYP71Z18* (*Zx6*) and Zm*CYP71Z16* (*Zx7*) referenced to the B73 genome (RefGen_V4). **c**, Expression matrix of *Zx5*, *Zx6* and *Zx7* in 100 inbred lines present in a dataset of 1,960 RNA-seq samples. Genes with expression levels of ≥5% of the sum *Zx5* to *Zx7* expression were counted as being expressed (filled black squares). **d**, GC–MS-extracted ion chromatograms of hexane extracts derived from *Agrobacterium*-mediated transient *N. benthamiana* coexpression assays of ZX3 individually paired with ZX5, ZX6 or ZX7 all resulted in the production of ZA1. **e**, Parallel coexpression assays and extracted ion chromatograms of the β-bisabolene synthase from *Santalum album* (*Sa*MonoTPS) with ZX5, ZX6 or ZX7 all resulted in the production of ZD1 and ZD2 in variable proportions. Four biological repeats were performed and showed similar results. **f**, Ratios of ZD1 to ZD2 from *Agrobacterium*-mediated transient *N. benthamiana* coexpression assays of *Sa*MonoTPS with ZX5, ZX6 or ZX7. Data are mean ± s.e.m. $n = 6$ biologically independent replicates. Different letters (a–c) represent significant differences; statistical analysis was performed using one-way ANOVA, $P < 0.05$; Tukey tests were used to correct for multiple comparisons, $P < 0.05$. **g**, Schematic of farnesyl pyrophosphate (FPP) cyclization reactions catalysed by ZX3 (a β-bisabolene-dependent β-macrocarpene synthase) and ZX5, ZX6 and ZX7 each yield ZD1, ZD2 and ZA1.

P450 transcripts and found nine candidates with low MR scores (<250), including Zm*CYP71Z18* and Zm*CYP71Z19* (Fig. 2a), that were further supported by replicated RNA-seq data (Supplementary Table 2). Zm*CYP71Z19* is phylogenetically most closely related to Zm*CYP71Z16/18* (ref. [18]) and is located within the same 15-gene interval on chromosome 5 (Fig. 2b and Supplementary Fig. 9).

Similar to *Zx1* to *Zx4*, Zm*CYP71Z16*/*18*/*19* each display variable relative expression between inbred lines (Fig. 2c). We name the B73 P450 genes Zm*CYP71Z19* (*Zm00001d014121*) *Zx5*, Zm*CYP71Z18* (*Zm00001d014134*) *Zx6* and Zm*CYP71Z16* (*Zm00001d014136*) *Zx7* on the basis of chromosome order, with each sharing >71% protein sequence identity (Supplementary Fig. 10). Gene-cluster-II analyses of 28 maize genomes identified only modest variation but, interestingly, a single inbred line (CML228) contains duplications of *Zx5*, *Zx6* and *Zx7* to yield a total of six *CYP71Z*-family genes (Supplementary Figs. 8 and 11, and Supplementary Tables 6 and 7). Consistent with a shared role in ZX biosynthesis, *Agrobacterium*-mediated enzyme coexpression assays with B73 ZX3 in *N. benthamiana* demonstrate that ZX5, ZX6 and ZX7 each independently catalyse the oxidization of β-macrocarpene to ZA1 (Fig. 2d,g), supporting pathway redundancy.

As a less abundant product of ZX1 to ZX4 (Fig. 1f,g and Supplementary Fig. 5), β-bisabolene predictably contributes to the array of 13 established candidate ZXs[29,33]. Importantly, levels of β-bisabolene production by β-macrocarpene synthases vary, and alterations in pH and Mn^{2+} availability can result in the coequal production β-bisabolene and β-macrocarpene by ZX1 (*Zm*TPS6) when expressed in *E. coli*[33]. Diseased maize sheath tissue was used to purify, isolate and NMR-identify two acidic β-bisabolene derivatives, namely zealexin D1 (ZD1; 4-(6-methylhepta-1,5-dien-2-yl) cyclohex-1-ene-1-carboxylic acid) and ZD2 (2-methyl-6-(4-methylcyclohex-3-en-1-yl)hepta-2,6-dienoic acid), that produce diagnostic gas chromatography coupled mass spectrometry (GC–MS) electron ionization (EI) spectra as methyl ester derivatives (Supplementary Table 8 and Supplementary Fig. 12). To specifically examine the catalytic oxidation of β-bisabolene, we performed *N. benthamiana* coexpression assays using *Santalum album* mono TPS (*Sa*MonoTPS, EU798692), which utilizes the precursor *E*/*E*-farnesyl diphosphate (FDP) to produce β-bisabolene[41]. Similar to ZA1 biosynthesis, ZX5, ZX6 and ZX7 each catalysed the complete oxidation of β-bisabolene at the C1 and C15 positions, yet resulted in significant differences in the final ratios of ZD1/ZD2 produced (Fig. 2d–f and Supplementary Fig. 13). Our collective findings demonstrate that ZX6 and ZX7 support promiscuous catalytic activity on established ZX1 to ZX4 products (Fig. 1f,g) and also diterpenoid defence precursors *ent*-isokaurene (Supplementary Fig. 14) and dolabradiene[18,30]. ZX5 enzyme coexpression analyses demonstrate activity on sesquiterpene olefins, (Fig. 2d,e) further including conversion of the substrate β-selinene to β-costic acid (Supplementary Fig. 14); however, no appreciable activity in kauralexin biosynthesis was found[18,28] (Supplementary Fig. 14). As previously observed maize sesquiterpene acids, termed analytes 2 and 4 (ref. [29]), our results clarify that low levels of β-bisabolene-derived ZD2 and ZD1 co-occur with A, B and C-series ZXs.

To consider the evolutionary origin and catalytic potential of *Zx* gene cluster II, we examined the single-copy *Sorghum bicolor* gene (*Sobic.001G235500*) Sb*CYP71Z19*, which exists as a ZX5 syntenic orthologue sharing 89% amino acid identity (Supplementary Fig. 9) despite at least 12 million years[42] of phylogenetic divergence between the two genera. Both maize and *Sorghum* gene evolution estimates (Supplementary Fig. 15) and enzyme coexpression studies—which demonstrate that *Sb*CYP71Z19 can oxidize both sesquiterpene and diterpene precursors (Supplementary Fig. 14)—are consistent with the existence of a *CYP71Z* progenitor gene that possessed sufficient promiscuity to produce diverse terpenoid defences before gene duplication and divergence in maize. Similarly, highland teosinte (*Z. mays* spp. *mexicana*) contains only two *CYP71Z* genes near cluster II, and they are most similar to *Zx5* and *Zx7* (Supplementary Fig. 11).

To consider whether gene cluster II provides endogenous ZX pathway redundancy, we examined the W22 *zx5 Ds* insertion mutant (*dsgR102G05*) for fungal-elicited ZXs and found no measurable deficits (Supplementary Fig. 16). While suggestive, the absence of an altered ZX chemotype requires a detailed characterization of W22 *zx5* (*dsgR102G05*) at the transcriptional and enzymatic level before any potential impact of the transposon insertion can be stated with certainty. Of *Zx5* to *Zx7*, *Zx5* displays the highest degree of genome-wide coregulation with *Zx1* to *Zx4* (Fig. 2a) as well as the greatest degree of catalytic specificity towards sesquiterpene substrates (Supplementary Fig. 14). Our results support that enzyme promiscuity and gene-cluster-II redundancy enable partially interchangeable enzymes to be shared by at least four different maize defence pathways[18,30].

Forward genetics reveals a third *Zx* biosynthetic gene cluster. Beyond carboxylic acid derivatives, ZXs contain additional oxidations, desaturations and aromatized variants[29]. To identify the enzyme(s) that are responsible for these modifications, we screened germplasm for differences in fungal-elicited ratios of ZB1 to ZA1 and focused on inbred lines that have previously been used for biparental crosses in the establishment of mapping populations[43,44]. Compared with other examined inbred lines, Mo17 uniquely displays low ZB1/ZA1 ratios (Fig. 3a). Using the intermated B73×Mo17 (IBM) recombinant inbred lines (RILs)[43], we used the ratio of ZB1/ZA1 as a mapping trait in mature field roots and identified highly significant SNPs on chromosome 1 (Fig. 3b and Supplementary Table 9). Similarly, a genome-wide association study using the Goodman association panel[39] further supported colocalized SNPs (Supplementary Fig. 17) spanning the same interval.

To systematically narrow candidates, near-isogenic lines (NILs) derived from B73 and Mo17 (ref. [45]) were used for fine-mapping, and resulted in a narrow 100 kb region containing three B73 *CYP81A* genes (Fig. 3c–e) named Zm*CYP81A37* (*Zm00001d034095*) *Zx8*, Zm*CYP81A38* (*Zm00001d034096*) *Zx9* and Zm*CYP81A39* (*Zm00001d034097*) *Zx10*. MR analyses of the combined expression of *Zx1* to *Zx4* and *Zx5* to *Zx7* in relation to the mapping interval confirmed strong *Zx*-pathway coexpression (Fig. 3d). In contrast to B73, the Mo17 genome uniquely contains an 8 kb insertion in *Zx8* (Fig. 3f) and the Mo17 *Zx8* transcript displays no fungal-elicited accumulation (Fig. 3g). In the B73 genome, ZX8 shares 99% and 72% amino acid identity with ZX9 and ZX10, respectively (Supplementary Fig. 18).

To examine gene functions, combinations of the representative B73 pathway genes *Zx3* and *Zx6* were coexpressed in *N. benthamiana* with combinations of *Zx8*, *Zx9* and *Zx10*. Both ZX8/9 pairings resulted in the hydroxylation of ZA1 at carbon C1 to produce ZA2, hydroxylation at C6 to produce ZA5 (2-hydroxy-5′,5′-dimethyl-[1,1′-bi(cyclohexane)]-1′,3-diene-4-carboxylic acid), desaturation of the C1–C6 bond resulting in ZB1 and the predicted autoaromatization product ZC1 (Fig. 4a,b, Supplementary Figs. 12 and 13, and Supplementary Table 8). ZX10 resulted in the hydroxylation of ZA1 at the C8 position to yield ZA3. Parallel microbial coexpression of B73 enzymes ZX3 and ZX7 with ZX8 and ZX9 in *E. coli* also demonstrated the conversion of ZA1 to ZB1, while ZX10 similarly produced ZA3 (Supplementary Fig. 19). In *N. benthamiana*, the combined activities of B73 ZX3, ZX6, ZX8/9 and ZX10 yielded the additional additive product termed ZB3 (6′-hydroxy-5′,5′-dimethyl-[1,1′-bi(cyclohexane)]-1,1′,3-triene-4-carboxylic acid; Fig. 4a,b, Supplementary Figs. 12 and 13 and Supplementary Table 8) and low levels of the aromatic variant ZC2 (refs. [29,46]). Novel and established ZXs produced characteristic retention times and EI spectra when produced in maize and *N. benthamiana* tissues (Supplementary Figs. 12 and 13). Identifications in maize followed from purification and NMR elucidation (Supplementary Table 8). ZX biosynthetic deficiencies in Mo17 are partially explained by a loss of transcript accumulation of *Zx8* but not *Zx9* (Fig. 3g). *N. benthamiana* coexpression of B73 *Zx3* and *Zx6* with Mo17 *Zx9* resulted in equal transcript levels on the basis of quantitative PCR with

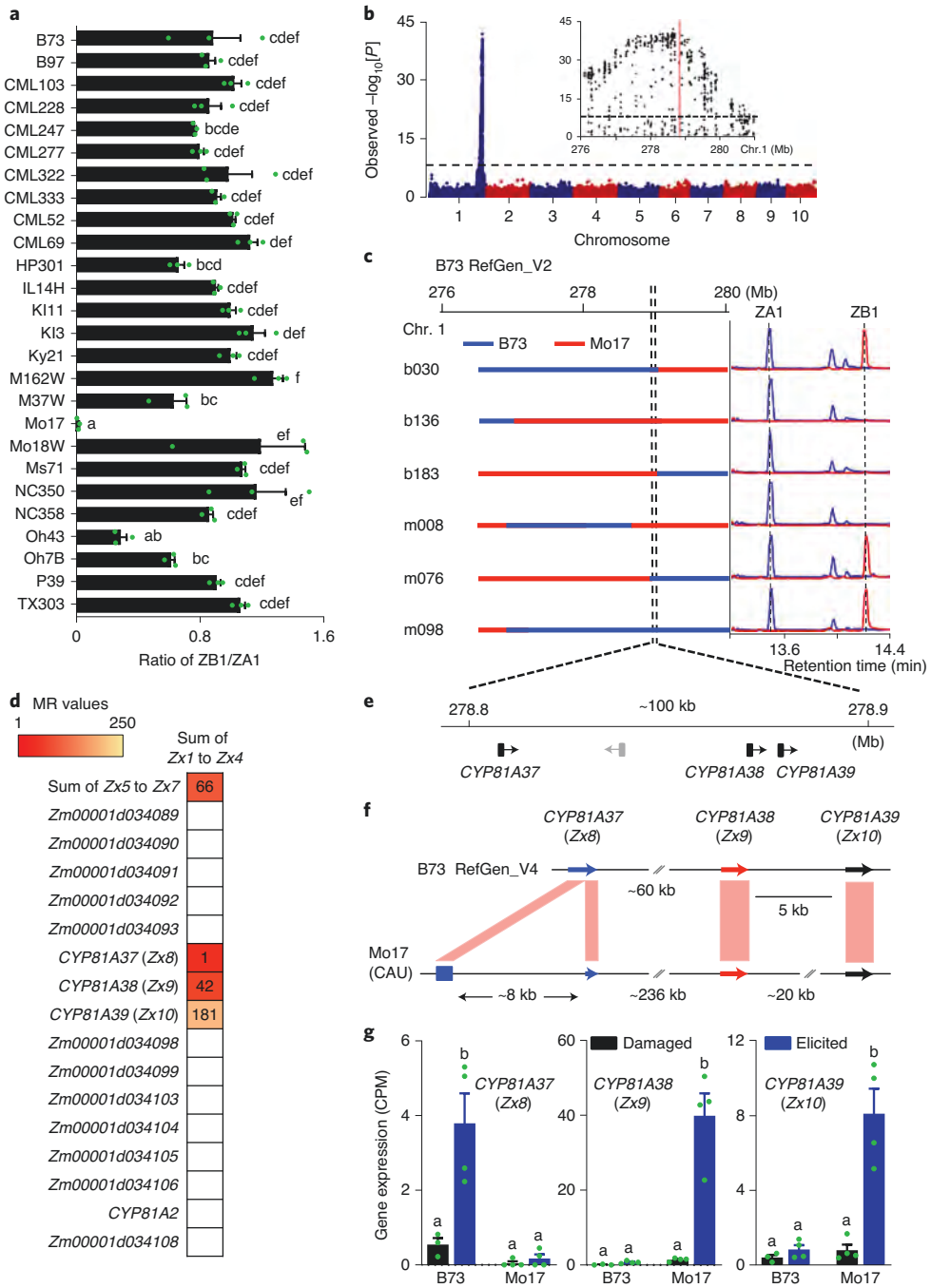


Fig. 3 | Association mapping reveals that *Zx* gene cluster III contains three CYP81A-family P450s. a, The ratio of ZB1 to ZA1 in stems of 25 different four-week-old inbred lines treated with heat-killed *F. venenatum* hyphae for 3 d identified that Mo17 is a unique parent. Data are mean ± s.e.m. $n = 3$ biologically independent samples. **b**, Association analysis of the ratio of ZB1 to ZA1 using the IBM RILs with the general linear model and 173,984 SNPs. The most statistically significant SNPs are located on chromosome 1 (B73 RefGen_V2). The black dashed line denotes the false discovery rate (< 0.05 at $-\log_{10}[P]$) using a Bonferroni correction. Inset: local Manhattan plot surrounding the peak on chromosome 1. The red line denotes the initial estimated position of the candidate gene(s). **c**, B73 and Mo17 chromosomal segments in B73 × Mo17-derived NILs represented by blue and red, respectively, paired with chemotypes indicated as GC–MS-extracted ion chromatograms (ZA1, $m/z = 136$, blue; ZB1, $m/z = 246$, red). **d**, Heat map of the coexpression of genes in the mapping region with the summed expression of *Zx1* to *Zx4* and *Zx5* to *Zx7* present in a dataset of 1,960 RNA-seq samples. The low numbers in the squares (<250) indicate supportive MR scores. Weak MR correlations of >250 were omitted. **e**, The locus was fine-mapped to a 100-kb region on B73 bacterial artificial chromosomes AC202436 and AC196018 on chromosome 1 containing four genes (B73 RefGen_v2). **f**, Tandem array of *CYP81A37* (*Zx8*), *CYP81A38* (*Zx9*), *CYP81A39* (*Zx10*) on chromosome 1 of the B73 genome (RefGen_V4) and the Mo17 genome (China Agricultural University (CAU)). Compared with B73 *Zx8*, Mo17 *Zx8* contains an 8-kb insertion. **g**, 3′ RNA-seq results derived from five-week-old B73 and Mo17 stems that were either damaged or additionally treated with heat-killed *F. venenatum* hyphae (elicited) and harvested 36 h later. Gene expression is given as counts per million mapped reads. Data are mean ± s.e.m. B73 damaged, $n = 3$ biologically independent samples; B73 elicited, $n = 4$ biologically independent samples; Mo17 damaged and elicited, $n = 4$ biologically independent samples. Within the plots, different letters (a–f) represent significant differences; statistical analysis was performed using one-way ANOVA; $P < 0.05$; Tukey tests were used to correct for multiple comparisons; $P < 0.05$.

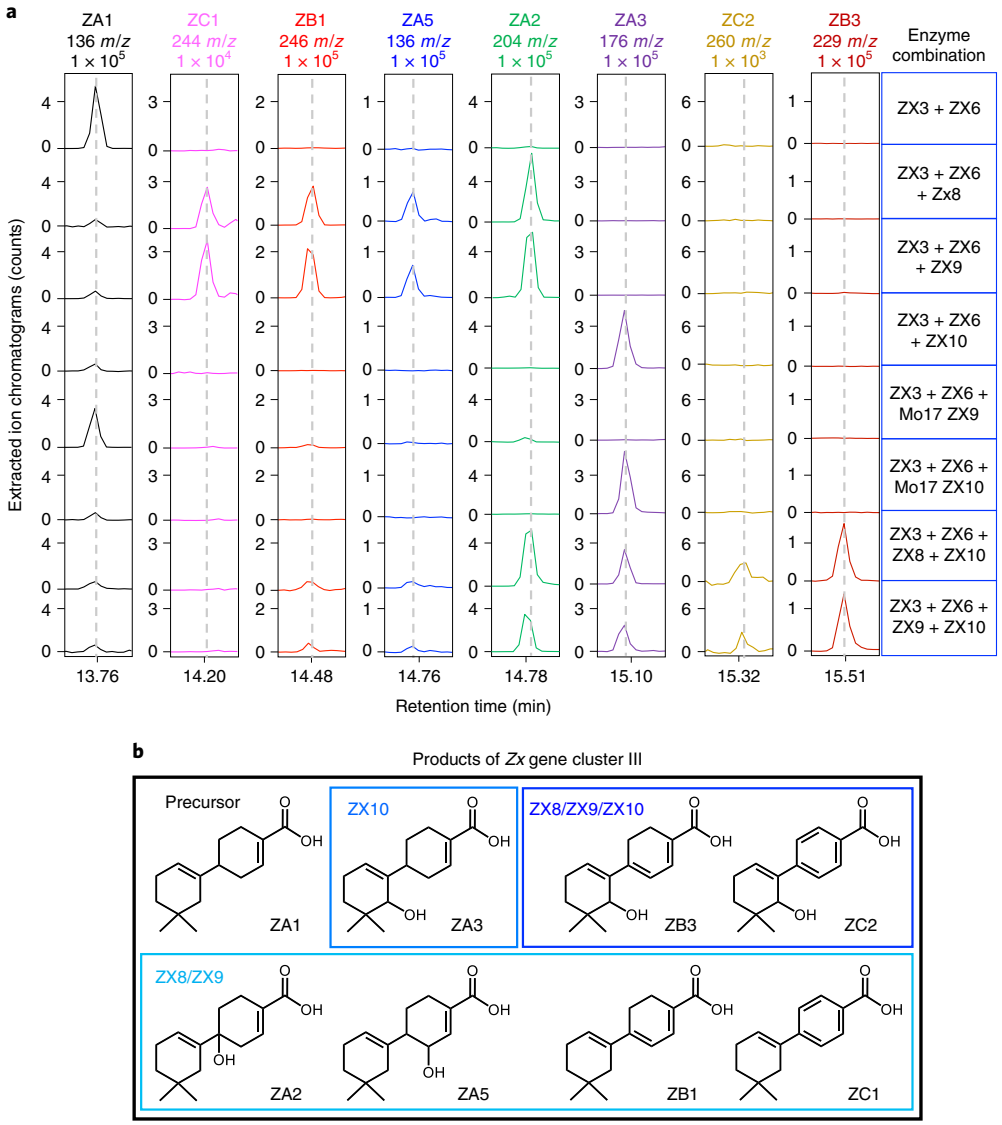


Fig. 4 | Enzyme coexpression defines the role of *Zx* gene cluster III in antibiotic biosynthesis. a, GC–MS-extracted ion chromatograms of extracts derived from *Agrobacterium*-mediated transient *N. benthamiana* enzyme coexpression assays using representative *Zx* pathway genes from gene cluster I (*Zx3*), gene cluster II (*Zx6*) and combinations from gene cluster III, including *Zx8*, *Zx9* and *Zx10*. Combinatorial in vivo enzyme assays in the presence of ZA1 ($m/z=136$, product of ZX3 and ZX6) yielded seven additional ZXs with diagnostic EI m/z ions as follows: ZC1, $m/z=244$; ZB1, $m/z=246$; ZA5, $m/z=136$; ZA2, $m/z=204$; ZA3, $m/z=176$; ZC2, $m/z=260$; and ZB3, $m/z=229$. ZA5 and ZB3 represent new compounds. To address the association mapping results, functionality of Mo17 ZX9 was included and supported impaired activity in ZB1 synthesis. In contrast to Mo17 ZX9, Mo17 ZX10 remains fully functional. Four independent experiments were performed and showed similar results. **b**, Structures of ZXs derived from the activity of ZX8, ZX9 and ZX10 on ZA1 as a substrate.

reverse transcription (RT–qPCR) analysis, yet a significant, more than tenfold decrease in ZA2, ZA5 and ZB1 production (Fig. 4a and Supplementary Fig. 20), consistent with separate deleterious mutations in both Mo17 *Zx8* and Mo17 *Zx9*.

To consider roles for ZX8 and ZX9, we purified ZB1 and observed significant antifungal activity against three important maize pathogens, namely *Fusarium verticillioides*, *F. graminearum* and *Aspergillus flavus*, similar to ZA1 (Supplementary Fig. 21). Gene-cluster-III analyses of 28 maize genomes identified select inbred lines that displayed either a contraction or expansion of *Zx8* gene copies yielded a range of 2–5 predicted *Zx* cluster-III genes (Supplementary Figs. 8 and 22, and Supplementary Tables 6 and 7). *Zx* gene cluster III expands the established roles of CYP81 enzymes beyond glucosinolate, isoflavonoid, lignin and xanthone biosynthesis to include sesquiterpenoid antibiotics (Supplementary Fig. 23).

Elicited antibiotic production occurs during large-scale transcriptomic, proteomic and metabolomic reprogramming. For more than 60 years, BXs have been extensively examined as the predominant chemical defences that protect maize seedlings from herbivores and pathogens[5,15,47,48]. To understand the context in which acidic terpenoids predominate, we applied heat-killed *Fusarium venenatum* hyphae to wounded maize stem tissues at 11 d, 25 d and 35 d after planting. After 3 d of elicitation, seedlings (aged 14 d) maintained predominantly BX metabolites, whereas plants aged 38 d displayed predominantly complex mixtures of acidic terpenoids and flavonoids (Fig. 5a). Experimental variables, such as the plant age and stress conditions considered, probably contributed to a historical focus on BX metabolites. To better understand defence activation in mature plants, we conducted a time-course experiment over a period of 120 h after applying heat-killed *Fusarium* hyphae to stem tissues of the

W22 inbred line and measured changes in transcripts, proteins and defence metabolites (Supplementary Tables 1, 2, 10 and 11). Early (0–4 h) versus late (72–120 h) fold changes in protein levels were averaged to provide statistical patterns. A combination of wounding and fungal elicitation resulted in 52% of proteins (5,501 out of 10,508) that displayed either significantly positive (2,694) or negative (2,807) changes in abundance after treatment (Supplementary Table 11). Protein–transcript pairs (10,508) were analysed using complete-linkage hierarchical clustering (Fig. 5b) and assigned to 11 modules (0–10) using weighted gene coexpression network analysis (WGCNA) of protein and RNA fold changes after rank-order normalization (Fig. 5c, Supplementary Fig. 24 and Supplementary Table 11). Many acidic terpenoid and flavonoid biosynthetic pathway genes grouped in module 3 (1,534) containing an enrichment in Gene Ontology (GO) terms relating to response to stimulus and secondary and phenylpropanoid metabolic processes (Fig. 5b,d and Supplementary Tables 11 and 12). Flavonoids are predominantly protective biochemicals in nearly all plants, are fungal-regulated in maize and commonly cooccur with terpenoids[49–51]. The upregulated production of simple flavonoids, such as naringenin and apigenin, is associated with increased protein levels of phenylalanine ammonia lyase, cinnamate 4-hydroxylase, 4-coumarate CoA ligase, chalcone synthase, chalcone isomerase and flavone synthase family members[52], many of which parallel ZX pathway activation (Fig. 5d). In contrast to terpenoid and flavonoid pathways, early BX pathway transcripts and enzymes were assigned to module 1 (4,787) and display rapid cosuppression (Fig. 5c,d and Supplementary Tables 2 and 11). By contrast, the accumulation of more terminal BX metabolites and biosynthetic proteins is pathogen-elicited (Fig. 5d), exists in modules with parallel RNA/protein increases (Supplementary Table 11 and Supplementary Fig. 24) and highlights a shift from general defences to those with increased reactivity[14].

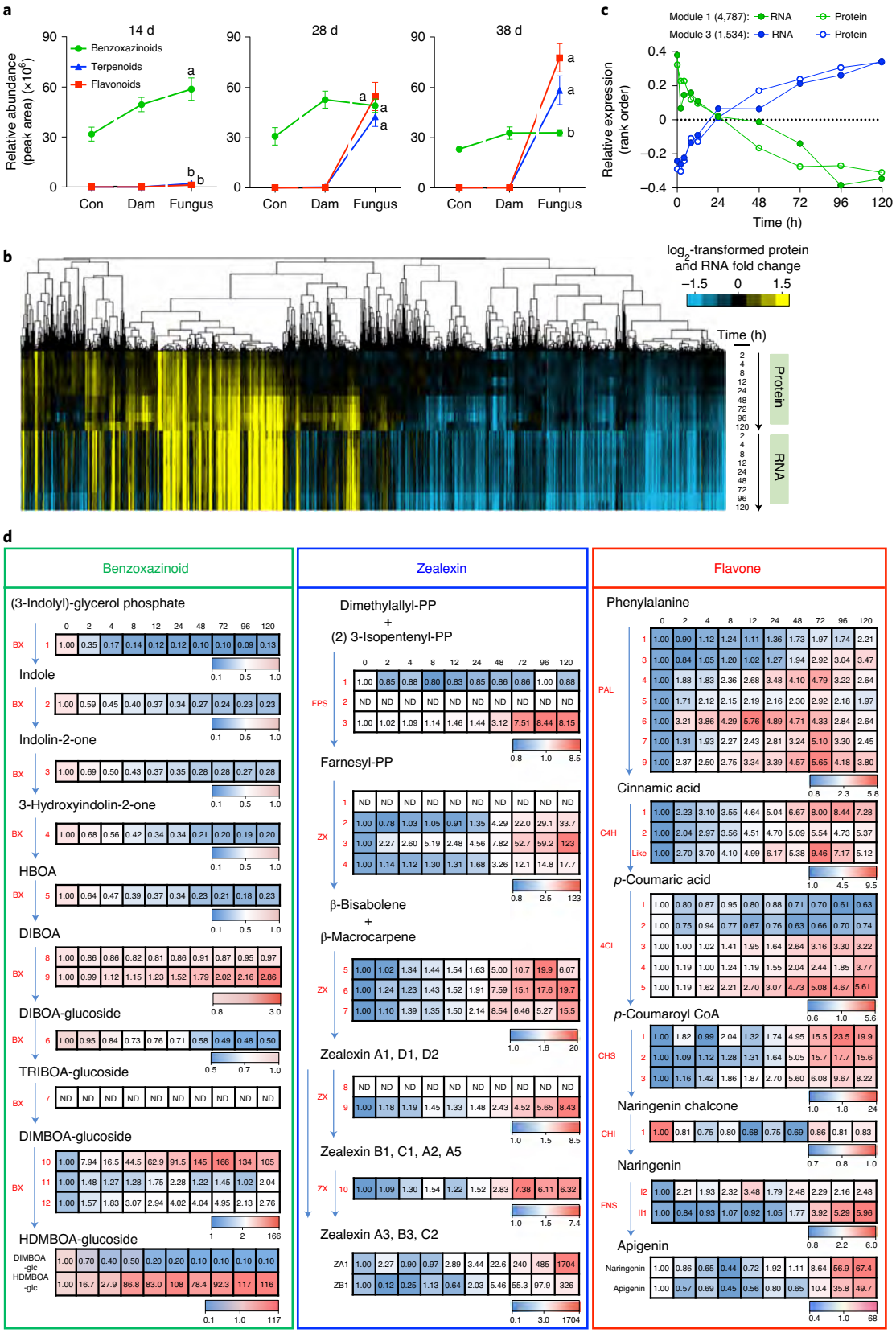
Diverse maize antibiotics produced through an hourglass-shaped biosynthetic pathway drive pathogen resistance. To better understand ZX diversity, we conducted large-scale stem inoculations with a necrotrophic fungal pathogen, namely southern leaf blight (SLB, *Cochliobolus heterostrophus*), and isolated additional related metabolites. Beyond the previously known ZXs[29,34], ZD1, ZD2, ZA5 and ZB3 (Figs. 2 and 4 and Supplementary Table 8), we describe four additional structures, namely ZA6 (1,4′-dihydroxy-5′,5′-dimethyl-[1,1′-bi(cyclohexane)]-1′,3-diene-4-carboxylic acid), ZA7 (4′,6′-dihydroxy-5′,5′-dimethyl-[1,1′-bi(cyclohexane)]-1′,3-diene-4-carboxylic acid), ZA8 (1-hydroxy-5′,5′-dimethyl-4′-oxo-[1,1′-bi(cyclohexane)]-1′,3-diene-4-carboxylic acid) and ZA9 (6′-hydroxy-5′,5′-dimethyl-4′-oxo-[1,1′-bi(cyclohexane)]-1′,3-diene-4-carboxylic acid) (Fig. 6a, Supplementary Table 8, and Supplementary Figs. 12 and 25) derived from β-macrocarpene. In SLB elicitation experiments of maize stem tissues, at least 15 ZX pathway products were detectable, produced diagnostic EI spectra and significantly accumulated over time (Supplementary Figs. 12, 25 and 26).

Our current and collective research[18,28–30,34] enabled the construction of the maize ZX pathway and shared functions (Fig. 6a). Gene duplications resulting in gene clusters I, II and III combine with enzyme promiscuity (Figs. 1, 2 and 4, and Supplementary Fig. 14) to create a complex hourglass pathway in which diverse endogenous substrates of independent origins share enzymes at an intermediate biosynthetic step and subsequently reconnect with pathway-specific enzymes. To examine interactive roles of the *Zx* gene cluster II, we performed a global gene coexpression analysis of the summed expression of *Zx5* to *Zx7* with all predicted maize P450 and TPS transcripts (Supplementary Fig. 27). The results support the strong coexpression of *Zx5* to *Zx7* with ZX (*Zx1*, *Zx2*, *Zx3*, *Zx4* and *Zx9*) and diterpenoid defence pathways (Zm*An2*, Zm*KSL2*, Zm*KO2* and Zm*KSL4*)[18,30] with a limited number of additional TPS and P450s transcripts (Supplementary Fig. 27). Heterologous expression studies in *N. benthamiana* support specificity in late pathway steps. For example, expression of the kauralexin enzymes, kaurene oxidase 2 (*Zm*KO2) and kauralexin reductase 2 (*Zm*KR2)[18] fail to significantly reduce the accumulation of ZA1 resulting from the expression of ZX3 and ZX6 (Supplementary Fig. 27). Similarly, expression of ZX9 has no impact on the accumulation of KB1 produced by the combination of *Zm*An2, *Zm*KSL2 and *Zm*CYP71Z18 (ZX6). As positive controls, in both cases, more terminal pathway-specific enzymes significantly reduced the accumulation of ZX and kauralexin precursors and resulted in the accumulation of late pathway products (Supplementary Fig. 27).

With an emphasis on ZXs and kauralexins, in which late pathway enzymes are now known[18], we present an integrated pathway model comprising three layers of interacting components: (1) a collection of sesquiterpene and diterpene synthases generating structurally diverse hydrocarbon olefin precursors that together (2) function as substrates for a family of catalytically promiscuous CYP71Z P450 enzymes, enabling partially overlapping pathway functions and diverse products that (3) are subsequently further decorated and diversified after reconnection with a final layer of pathway-specific enzymes (Supplementary Fig. 27). Collectively, our data support the hypotheses that a breadth of TPS-derived metabolites converges on a single CYP71Z node before again being expanded through pathway-specific downstream enzyme activities. Thus, maize defensive terpenoid metabolism seems to occur through an hourglass-shaped biosynthetic network rather than distinct individual pathways to produce a cocktail of oxidized antibiotics (Fig. 6a).

Fig. 5 | ZX-pathway activation occurs during the large-scale reprograming of fungal-induced defences. a, Biochemical defence activation on the basis of the relative amounts of major BXs, acidic terpenoids and flavonoids present—estimated from liquid chromatography coupled with MS (LC–MS) peak areas—in intact maize (var. Golden Queen) stem tissues (control (Con)) or those slit and treated with either H_2O (damaged (Dam)) or heat-killed *F. venenatum* hyphae (Fungus) aged 11 d, 25 d and 35 d and harvested 3 d later. Data are mean ± s.e.m. $n = 4$ biologically independent replicates. Within the plots, different letters (a and b) represent significant differences for the *F. venenatum* treatment; statistical analysis was performed using one-way ANOVA, $P < 0.05$; Tukey tests were used to correct for multiple comparisons, $P < 0.05$. **b**, Complete-linkage hierarchical clustering of 10,508 unique protein fold changes paired with mapping of $\log_2$-transformed RNA fold changes during a *F. venenatum* elicitation time course of 120 h in W22 stems using plants aged 38 d. The vertical lines correspond to individual gene IDs. Rows are organized by time point, data type and colours (blue, underexpressed; yellow, overexpressed). **c**, WGCNA of the proteomic and transcriptomic data from the W22 *F. venenatum* elicitation time course identified modules with distinct regulation patterns. Module 1 (4,787 gene IDs) contains highly cosuppressed gene–protein pairs, including early steps in BX biosynthesis, whereas module 3 (1,534 gene IDs) contains highly elicited coaccumulating gene–protein pairs, including the ZX biosynthetic pathway. The vertical axes indicate expression values relative to the mean expression across all of the time points. **d**, Heat maps of normalized protein fold changes in the W22 stem *F. venenatum* elicitation time course for BX, ZX and flavone pathways. The corresponding metabolite fold changes for representative BXs (2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one glucoside (DIMBOA-glc); 2-hydroxy-4,7-dimethoxy-1,4-benzoxazin-3-one-Glc (HDMBOA-glc)) and flavone pathway metabolites (naringenin, apigenin) were analysed using LC–MS, whereas ZXs (ZA1 and ZB1) were analysed using GC–MS. PP, pyrophosphate. Coloured scale bars in **d** indicate the fold change of normalized protein. The definitions of B73 RefGen_V4 gene IDs and all abbreviations are provided in Supplementary Table 2.

Previous studies have indirectly linked β-macrocarpene synthases, ZX production and disease resistance[29,32–35]. To examine endogenous relationships, we generated *zx1 zx2 zx3* triple and *zx1 zx2 zx3 zx4* quadruple insertion- and deletion-based mutants using CRISPR–Cas9 gene editing (Supplementary Fig. 28). Ten days after stalk inoculation with *F. graminearum*, *zx1 zx2 zx3 zx4*



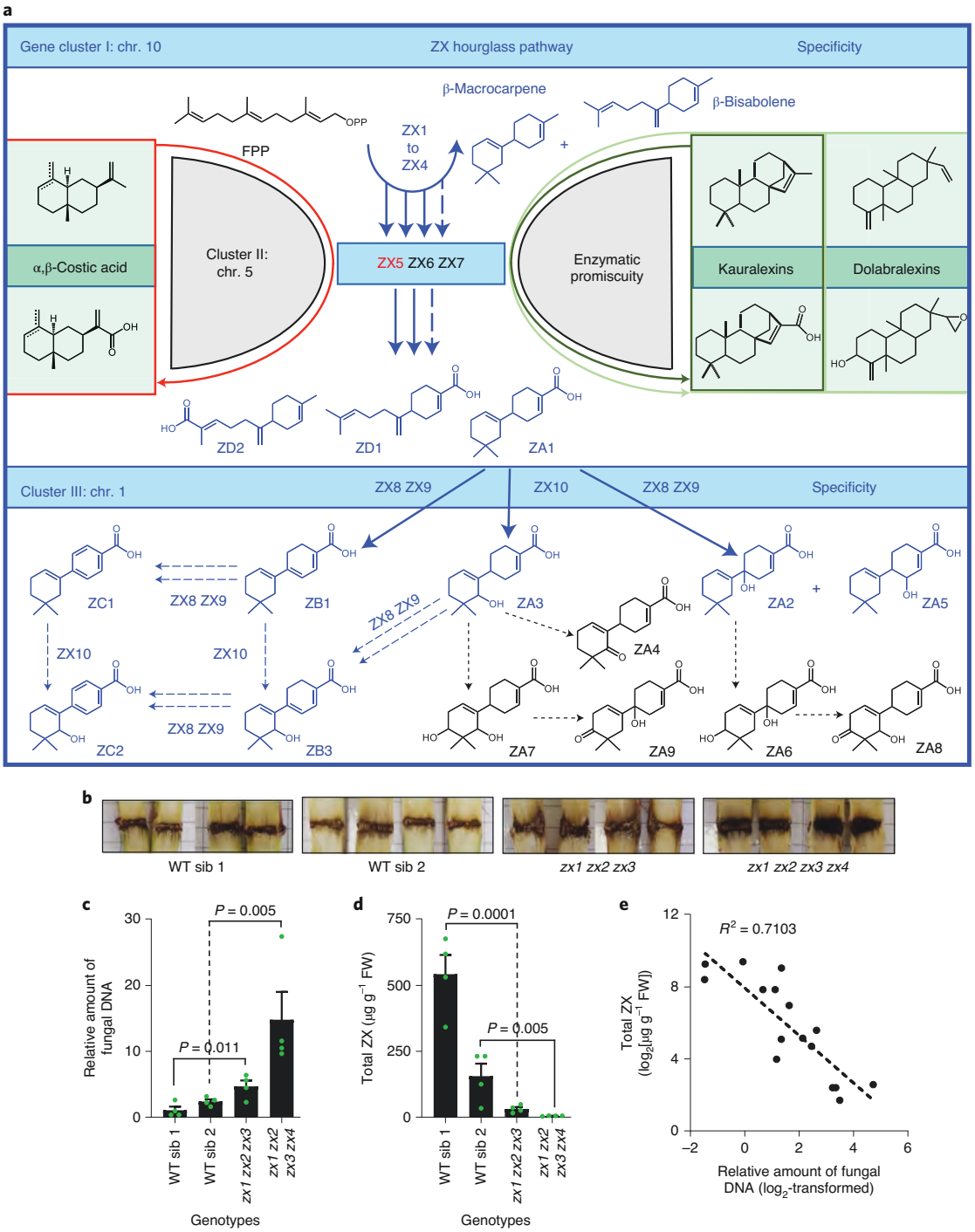


Fig. 6 | The maize ZX pathway is a biosynthetic hourglass with genetic redundancy and enzyme promiscuity that ensures the production of protective antibiotic cocktails. a, Schematic of enzyme activities encoded by *Zx* gene clusters I (*Zx1* to *Zx4*), II (*Zx5* to *Zx7*) and III (*Zx8* to *Zx10*), including interactions with additional pathways, such as kauralexins (dark green), dolabralexins (light green) and costic acids (red). Combined genomic, transcriptomic, proteomic, association mapping, enzyme coexpression studies and NMR-based structure elucidation support the genome-wide role of ZX1 to ZX10 in the production of 17 pathway products. The solid blue line arrows represent enzyme catalysis supported by genetics and in vivo expression studies. The dashed blue line arrows indicate demonstrated activities of gene cluster III that require enzyme feeding studies to clarify the predominant order of catalysis. The dashed black line arrows represent undefined enzyme activities. **b**, Representative disease levels in stems of CRISPR-Cas9-derived triple (*zx1zx2zx3*) and quadruple (*zx1zx2zx3zx4*) mutants and the respective wild-type siblings (WT sib), which were grown for 25 d and were stem-inoculated with *F. graminearum* (10 μl of 1.5×10^5 conidia per ml) for 10 d. Eight biological replicates were performed and showed similar results. **c**, The relative amount of fungal DNA in triple (*zx1zx2zx3*) and quadruple (*zx1zx2zx3zx4*) mutants and the respective wild-type siblings. RT-qPCR was used to determine the ratio of fungal DNA (Fg*Tri6*) to maize DNA (Zm*RPL17*) in stems 10 d after inoculation. **d**, Total ZX (μg g^{-1} fresh weight (FW)) present in triple (*zx1zx2zx3*) and quadruple (*zx1zx2zx3zx4*) mutants and the respective wild-type siblings as measured using GC–MS. **e**, $\log_2$-transformed correlation of endogenous ZX with the relative amount of fungal DNA in triple (*zx1zx2zx3*) and quadruple (*zx1zx2zx3zx4*) mutants and the respective wild-type siblings. For **c** and **d**, data are mean ± s.e.m. $n = 4$ biologically independent replicates. *P* values were determined using Student's *t*-tests, two-tailed distribution, equal variance.

mutant plants displayed visible (Fig. 6b) and quantitative increases in disease susceptibility, as estimated by the relative amount of fungal DNA (Fig. 6c). In contrast to wild-type plants and *zx1 zx2 zx3* plants containing a single β-macrocarpene synthase, *zx1 zx2 zx3 zx4* mutants consistently displayed a lack of detectable ZXs after inoculation with *F. graminearum* (Fig. 6d), yet were not impaired in kauralexin production (Supplementary Fig. 29). Across all samples, ZX production was negatively correlated ($R^2 = 0.71$) with *F. graminearum* DNA levels (Fig. 6e). Initially derived from HiII—a complex genetic background containing B73 and A188—wild-type siblings for zx1 zx2 zx3 and *zx1 zx2 zx3 zx4* mutants displayed differential biosynthetic abilities for ZX accumulation, emphasizing the importance of multiple controls (Fig. 6c,d). Enhanced disease susceptibility to Stewart's wilt, which is caused by the xylem-dwelling bacteria *Pantoea stewartii*, was similarly observed in *zx1 zx2 zx3 zx4* mutants (Supplementary Fig. 30), demonstrating protective roles against diverse pathogens. Suppression of root ZX production in *zx1 zx2 zx3* and *zx1 zx2 zx3 zx4* mutants (Supplementary Fig. 31) also significantly altered the root bacterial microbiome associated with plants grown in field soil, lowering evenness and altering the abundances of particular taxa (Supplementary Fig. 32 and Supplementary Table 13). Collectively, the ZX pathway and complex array of resulting antibiotics has a significant role in maize interactions with microorganisms.

Discussion

An understanding of the genetic, biosynthetic and regulatory machinery that controls innate immunity is essential for optimizing biochemical defences and crop resistance traits. For insights into maize disease resistance, we leveraged multi-omic approaches to elucidate hidden biochemical layers of immunity. Here we characterized ten genes that are present in three distinct gene clusters that ensure the production of 17 ZX pathway metabolites with collective antibiotic action (Fig. 6a, Supplementary Table 2, and Supplementary Figs. 1, 2 and 21). Our data collectively support the existence of a biosynthetic hourglass pathway (Fig. 6a) whereby maize antibiotics production relies on enzymes encoded by three Zm*CYP71Z*-family genes (*Zx5* to *Zx7*) on chromosome 5 that contribute to multiple distinct families of sesquiterpenoids and diterpenoids[18,30]. Maize antibiotic biosynthesis seems to be highly interconnected, and provides a model for how complex combinatorial blends are biosynthesized, contributing to immunity against diverse microorganisms.

Association mapping efforts to uncover genetic loci responsible for quantitative resistance to diverse maize pathogens commonly result in the discovery of multiple loci with comparatively small effects that explain 1–3% of trait variation[21,53,54]. ZX product complexity, pathway redundancy and overall resiliency to mutations are consistent with multiple disease-resistance quantitative trait loci that are commonly too small to be detected individually[22]. In contrast to qualitative resistance genes—such as wall-associated kinase (Zm*WAK*), which protects against head smut (*Sporisorium reilianum*)[55]—ZX biosynthesis as a trait is not controlled by a single Mendelian locus. At each of the three *Zx* gene clusters, pan-genome expression, sequence-level or gene-copy-level variation exists, which can impact individual pathway enzymes; however, gene-cluster redundancies ensure ZX biosynthesis.

Zx gene cluster I, which is encoded on chromosome 10, most commonly contains four tandem duplicate β-bisabolene/β-macrocarpene synthase genes, termed *Zx1* to *Zx4*. Comparative coexpression in *N. benthamiana* was used to prove pathway products (Fig. 1) and to understand the catalytic differences in ZX1 controlled by a single SNP common among inbred lines (Supplementary Fig. 6 and Supplementary Table 5). With genetic variation driving exonic changes and observed differences in relative expression, varying functional copies of *Zx1* to *Zx4* are commonly maintained in the maize pan-genome with inbred-specific patterns (Fig. 1a–i, Supplementary Fig. 7 and Supplementary Table 7). ZX pathway redundancy contrasts with both the BX and kauralexin pathways, for which single gene mutations in indole-3-glycerol phosphate lyase (*benzoxazinless1*, *bx1*) and kaurene synthase-like 2 (Zm*ksl2*) can reduce pathway metabolites to 1% of the wild-type levels and impair biotic stress resistance[5,18,56]. ZX-pathway resiliency to single-gene mutations, coupled with reduced pathogen resistance in *zx1 zx2 zx3 zx4* quadruple mutants, supports the hypothesis that maize relies on ZXs as key biochemical defences.

To generate acidic non-volatile antibiotics, *Zx* gene cluster II on chromosome 5 contains three neighbouring duplicated *CYP71Z* genes (*Zx5* to *Zx7*) that each display variable coexpression with *Zx1* to *Zx4* (Fig. 2a and Supplementary Fig. 27) and drive the production of ZA1, ZD1 and ZD2 (Fig. 2d,e). Previously associated with the synthesis of kauralexins, dolabralexins and ZA1, ZX6 and ZX7 contribute to a powerful in vivo system for combinatorial chemistry[18,30]. We now demonstrate that ZX5 also acts on broader sesquiterpene olefins, including the Zm*TPS21* product β-selinene to produce β-costic acid. However, ZX5 lacks appreciable kauralexin biosynthetic activity (Supplementary Fig. 14), highlighting specific differences in the product-diversifying roles of enzymes contained in *Zx* gene cluster II. Syntenic to *Zx5*, the closely related *Sorghum* gene Sb*CYP71Z19* encodes an enzyme that similarly generates acids from β-selinene, β-bisabolene and β-macrocarpene, indicating that the biosynthetic ability of related genes to oxidize diverse terpenoid precursors existed in a common ancestor of maize and *Sorghum* (Supplementary Figs. 14 and 15). Our present systematic analyses of gene cluster II also uncovered the presence of only two related Zm*CYP71Z* genes in teosinte—which, together with synteny and activity in Sb*CYP71Z19*—provides new insights into the origin of ZXs (Supplementary Figs. 9, 11, 14 and 15). The high degree of coregulation between the sum expression of *Zx1* to *Zx4* with *Zx5*, and comparative sesquiterpene substrate specificity are consistent with an important role for ZX5 in ZX biosynthesis while maintaining the flexible resiliency of gene cluster II to deleterious mutation (Fig. 2a,b and Supplementary Figs. 11 and 14).

Genetic fine-mapping on chromosome 1 identified *Zx* gene cluster III, which unexpectedly revealed three related CYP81A-family P450s. Although the CYP81 subfamily have established roles in the specialized metabolism surrounding glucosinolate, isoflavonoid, lignan and xanthone biosynthesis, none have previously been demonstrated to utilize terpenoid substrates (Supplementary Fig. 23 and Supplementary Table 3). ZX8 and ZX9 are functionally redundant, acting on ZA1 to produce four oxidized products, namely ZA2, ZA5, ZB1 and ZC1 (Fig. 4a,b). Successive rounds of oxidation at the C6 position probably yield the observed C1–C6 desaturation in ZB1 (ref. [57]). *Zx10* represents the sole non-redundant pathway gene, encoding *Zm*CYP81A39, which is responsible for ZX C8 oxidation to an alcohol and the combined variants ZA3, ZB3 and ZC2 (Fig. 4a,b). Additional ZXs, namely ZA6 to ZA9, displaying C10 oxidations to alcohols and ketones, were also identified in maize tissues; however, the final enzymes that are responsible remain unknown. Collectively ZX1 to ZX10 account for the production of 12 out of the 17 identified ZX pathway precursors and end products (Fig. 6a). As a major product of ZX1 to ZX9 action, ZB1 exhibits significant antifungal activity at 25 μg ml^{-1} against two key *Fusarium* pathogens of maize (Supplementary Fig. 21).

ZX biosynthesis is a highly coregulated pathway, fully contained in WGCNA module 3 that includes 1,534 transcript–protein pairs that are enriched for predominant GO terms 'response to stimulus' and 'secondary metabolic processes' (Fig. 5c and Supplementary Tables 10 and 11). Beyond terpenoids, phenylpropanoid defences, including naringenin, chalcone, apigenin and apigenin 7-*O*-methyl ether, are known to accumulate in maize after anthracnose stalk rot (*Colletotrichum graminicola*) infection and reduce fungal

growth[51]. Our current research demonstrates that fungal-elicited flavonoid-pathway activation in maize is highly coordinated with terpenoid defences (Fig. 5a,d). Multi-omic analyses place ZX biosynthesis in the context of massive reorganization of the transcriptome and proteome, including a fivefold suppression of early BX biosynthetic enzymes (BX1 to BX5; Fig. 5d) and, more generally, an enrichment in GO terms that define processes surrounding 'DNA/RNA metabolism' in module 1 (Fig. 5c and Supplementary Tables 10 and 11). Together, our experiments address a fifteen-year-old hypothesis that sesquiterpenoids mediate maize disease resistance[32], consider ZXs in the context of multiple biochemical defence pathways and place ZXs among predominant antibiotics that contribute to defence (Figs. 5d and 6a–e).

Independent of genetic mechanisms pursued, the leveraged application of durable multiple-disease-resistance traits is a key goal in crop protection[22]. Efforts in *Sorghum* have resulted in the identification of complete defence pathways, defined enzyme organization in biosynthetic metabolons and enabled the relocation of pathways to specific organelles in heterologous plants[58,59]. Capturing the full breadth of resistance traits provided by complex pathways requires an understanding of interconnections and biosynthetic nodes. To demonstrate the comparative importance of individual maize antibiotics pathways and other defence mechanisms, a combinatorial series of pathway knockouts would be needed for *zx1 zx2 zx3 zx4*, Zm*tps21* (ref. [28]), Zm*ksl2* (ref. [18]), Zm*ksl4* (ref. [30]) and established resistance genes of interest[60] in a single inbred line for diverse field trials. In each terpenoid pathway, evidence supports roles in resistance to *Fusarium* spp.[18,28,30]; however, assignments of relative importance between pathways are not presently possible. While speculative, gene duplications throughout the ZX biosynthetic pathway suggest additional selection pressures and protective roles[16]. Uncoupling pathways at different biosynthetic nodes presents a challenge given the evidence for ZX biosynthetic interactions with multiple terpenoid pathways mediated by promiscuous *Zm*CYP71Z enzymes. Although the full transfer of maize terpenoid antibiotic defences to a non-native crop model would require a series of pathway genes, leveraging single-gene transfers and knowledge of enzyme promiscuity with existing modular pathways has recently provided enhanced levels of innate immunity in rice against fungal pathogens[36].

Pathogen-elicited terpenoid antibiotics have been studied in crops for more than 40 years and led to the discovery of TPS-mediated plant defences[16,61,62]. Despite a massive growth of comparative -omics, delineating clear connections between genotypes, chemotypes and phenotypes has remained a challenge owing to genetic redundancies and enzyme promiscuity. Our use of association analyses paired with transcriptional coregulation patterns, combinatorial biochemical studies and targeted mutant analyses using CRISPR–Cas9 collectively provides powerful tools to narrow and interrogate the metabolic pathways that control plant innate immunity. Heterologous enzyme coexpression studies efficiently define candidate gene functions, impact of genomic variation, promiscuity, redundancy and endogenous pathway interactions leading to antibiotic complexity. Comprehensive proteomics confirms the existence of endogenous translation products and gives a more complete context to the regulation of multiple pathways that are known to, or suspected to, impact defence phenotypes. Our present elucidation of highly interconnected antibiotic pathways illuminates complex combinatorial strengths in the genus *Zea*, which can now be considered in breeding and additional pathway engineering approaches to effectively enhance disease resistance in crops[36].

Methods

Plant and fungal materials. Maize seeds for the IBM RILs[43] and the Goodman diversity panel[39] were provided by G. Jander (Boyce Thompson Institute) and P. Balint-Kurti (US Department of Agriculture, Agricultural Research Service (USDA-ARS)). Nested association mapping[44] parental line seeds and B73×Mo17 NILs were obtained from the Maize Genetics Cooperation Stock Center. A list of all of the maize lines used for genetic mapping efforts is provided in Supplementary Table 14. *Zea perennis* (Ames 21874), *Zea diploperennis* (PI 462368), *Zea luxurians* (PI 422162), *Zea mays parviglumis* (PI 384069) and *Z. m. mexicana* (Ames 21851) were provided by the USDA-ARS (North Central Regional Plant Introduction Station). Maize inbred lines that were used for replicated elicitation experiments were germinated in Metro-Mix 200 (Sun Gro Horticulture Distribution) supplemented with 14-14-14 Osmocote (Scotts Miracle-Gro) and grown in a greenhouse as previously described[28]. Fungal cultures of *F. graminearum* (NRRL 31084), *F. verticillioides* (NRRL 20956), *A. flavus* (NRRL 3357) and SLB (*C. heterostrophus*) were grown on V8 agar for 12 d before the spores were quantifi d and used[29]. Heat-killed *F. venenatum* (strain PTA-2684) hyphae were commercially obtained (Monde Nissin) and used as a uniform non-infectious elicitor to avoid complexities in diverse germplasm responses, which are likely to occur across a spectrum of susceptible and resistant interactions with live pathogens. In total, maize experiments utilized multiple pathogens, including *Fusarium*, heat-killed *Fusarium*, *Aspergillus, Cochliobolus, Pantoea* and combined microbial complexities associated with fi ld soil to address a range of interactions in which the ZX pathway is present and is likely to be activated.

Maize stem challenge with heat-killed *Fusarium* and live fungi. Using a scalpel, plants (aged 35 d) were slit in the centre, spanning both sides of the stem, to create a 10 cm longitudinal incision. The incision wounded the upper nodes, internodes and the most basal portion of unexpanded leaves. For replicated ($n=3$–4) 36 h experiments using B73 and Mo17, the ten-point ($n=1$) 0–120 h time course with W22 (ref. [18]) and the three-day treatment of the Goodman diversity panel, approximately 500 µl of commercial heat-killed *F. venenatum* hyphae was introduced into each slit stem, and then the site was sealed with clear plastic packing tape to minimize desiccation of the treated tissues. B73 and Mo17 experiments included parallel wound control plants lacking fungal hyphae treatment. For the quantification of ZX diversity after *C. heterostrophus* inoculation, maize (*Z. mays* var. Golden Queen) plants were wounded as described above and treated with either 100 µl of H_2O or an aqueous *C. heterostrophus* spore (1×10^7 ml^{-1}) suspension. Damage controls and *C. heterostrophus*-treated plants were harvested each day for three consecutive days. For the stalk-rot resistance assay, a hole (diameter, 1 mm) was created through the second aboveground node in the stalk of plants (aged 35 d) and inoculated with either 10 µl of H_2O or 10 µl of a *F. graminearum*-spore (1.5×10^5 ml^{-1}) suspension. After 10 d, stems were longitudinally slit with a scalpel, photographed and harvested using pool of two individual plants for each of the four final harvested replicates. Within each experiment, treated maize stem tissues were harvested into liquid N_2 at specific time points as indicated.

***Pantoea stewartii* resistance assay.** *P. stewartii* ssp. *stewartii* strain DC283 harbouring the plasmid pHC60 encoding GFP^{S65T} (DC283-GFP; nalR and tetR) was used as described previously[63]. Nalidixic acid (30 µg ml^{-1}) and tetracycline (20 µg ml^{-1}) were used for selection of DC283-GFP when grown in Luria-Bertani (LB) agar and LB broth at 28 °C. Bacteria were subcultured by diluting 1:10 into 10 ml final volume with antibiotics and grown to an optical density at 600 nm (OD_{600}) of 0.7. Bacteria were harvested by centrifuging at 2,800*g* for 10 min and resuspended in PBS supplemented with 0.01% Tween-20 (PBST buffer) three times. Final bacterial OD was adjusted to an OD_{600} of 0.2 and used for infiltration. Maize seedlings (aged 12 d) were punctured with a needle (diameter, 1 mm) in the internode between aboveground node 1 and node 2 and infiltrated with 10 µl PBS buffer (mock) or 10 µl *P. stewartii*. Plants were evaluated after 5 d for bacterial growth by GFP quantification and 16 d after infection for visual symptoms by counting the number of dying leaves per plant due to bacterial wilting. Progression of *P. stewartii*-GFP bacteria in veins was visualized by illumination with blue light using a Dark Reader Spot Lamp (DRSL; Clare Chemical Research) as previously described[64]. To quantify *P. stewartii*-GFP, total protein from infected leaf tissue was extracted in PBST buffer. GFP fluorescence intensity was measured using a Synergy H1 Multi-Mode microplate reader (BioTek) equipped with a green filter cube (excitation 485/20 nm, emission 528/20 nm) using total-protein extract from mock-inoculated plants as a blank[64]. GFP fluorescence intensity was normalized to the highest fluorescence value and is shown as relative fluorescence units (RFU). For each extraction, two technical replicates were averaged and used to calculate RFU.

RNA-seq analyses of fungal-elicited genes. To examine B73 and Mo17 defence transcript changes after elicitation with heat-killed *Fusarium* hyphae, total RNA was isolated using the NucleoSpin RNA Plant Kit (Takara Bio) according to the manufacturer's protocol. RNA quality was assessed on the basis of RNA integrity number using an Agilent Bioanalyzer. 3′ RNA-seq library construction and sequencing were performed at Cornell University's Genomics Facility, Institute of Biotechnology (Ithaca; http://www.biotech.cornell.edu/brc/genomics-facility/services). Approximately 500 ng of total RNA was used to construct the 3′ RNA-seq libraries using the QuantSeq 3′ mRNA-seq Library Prep Kit FWD (Lexogen) according to the manufacturer's instructions. All of the libraries, each with their own unique adapter sequences, were pooled together and sequenced in one lane

of an Illumina NextSeq 500 to generate 90-bp single-end reads. Trimmomatic (v.0.39) was used to remove Illumina TruSeq adapter sequences and trim the first 12 bp (ref. [65]). Trimmed reads were aligned to the maize B73 V4 reference genome (ensemble 4.44) using Hisat2 (v.2.0.0)[66] and sorted using Sambamba (v.0.6.8)[67]. Raw mapped reads were quantified using featureCounts (v.1.6.4)[68]. Counts per million were processed in R with DESeq2 to generate normalized read counts using the default method and transformation of the normalized counts was performed using the rlog transformation method[69].

For the analyses of W22 tissues, RNA-seq library construction and sequencing were performed by Novogene. The mRNA was first enriched from total RNA using oligo (dT) magnetic beads and then fragmented randomly into short sequences followed by first-strand cDNA synthesis with random hexamer-primed reverse transcription. Second-strand cDNA synthesis was performed by nick-translation using RNase H and DNA polymerase I. After adapter end repair and ligation, cDNA was amplified using PCR and purified to create the final cDNA library. cDNA concentration was quantified using a Qubit (v.2.0) fluorometer (Life Technologies) and then diluted to 1 ng μl^{-1}, before assessing insert size using an Agilent Bioanalyzer 2100. Library preparations were sequenced using an Illumina platform and paired-end reads were obtained. Image analysis and base calling were performed using the standard Illumina pipeline. Raw reads were filtered to remove reads containing adapters or reads of low quality. Qualified reads were then aligned to *Z. mays* AGPv4 reference genome using TopHat (v.2.0.12)[70]. Gene expression values, calculated as fragments per kilobase per million reads, were analysed using HTSeq (v.0.6.1)[71].

Statistical analyses. Statistical analyses were conducted using JMP Pro v.13.0 (SAS Institute) and Prism v.8.0 (GraphPad). One-way ANOVA was performed to evaluate statistical differences. Tukey tests were used to correct for multiple comparisons between control and treatment groups. Student's unpaired two-tailed *t*-tests were conducted for pairwise comparisons. $P < 0.05$ were considered to be statistically significant.

Further information on the research methods is available in the Supplementary Information.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Publicly available datasets used in the study include the National Center for Biotechnology Information (NCBI) Sequence Read Archive project ID SRP115041 and the MaizeGDB BLAST database (https://maizegdb.org/popcorn/main/index.php). Raw read sequences have been deposited at the NCBI Gene Expression Omnibus under accession numbers GSE138961 and GSE138962. Raw sequence data from the root microbiome are available at NCBI BioProject under accession number PRJNA580260. Raw proteomic mass spectra have been deposited at the Mass Spectrometry Interactive Virtual Environment repository (ftp://massive.ucsd.edu/MSV000084285). Maize-related germplasm used in this research have been previously described[27,39,43–45,72,73] and can be obtained from US Department of Agriculture Germplasm Resources Information Network (https://www.ars-grin.gov) and the Maize Genetics Cooperation Stock Center (http://maizecoop.cropsci.uiuc.edu). Where possible, gene identifiers used throughout the manuscript were in reference to B73 RefGen_v4 (https://www.maizegdb.org/genome/assembly/Zm-B73-REFERENCE-GRAMENE-4.0), which was used as a foundation for the study. All data are available from the corresponding author on request.

Code availability

For the gene-duplication date estimations, scripts to perform translations, alignments and backtranslations as well as data files and BEAST control files have been deposited at GitHub (https://github.com/TomJKono/Zealexin_Dating).

Received: 16 March 2020; Accepted: 11 September 2020;
Published online: 26 October 2020

References

1. Ritchie, H. & Roser, M. Land use. *Our World in Data* https://ourworldindata.org/land-use (2020).
2. Evenson, R. E. & Gollin, D. Assessing the impact of the green revolution, 1960 to 2000. *Science* **300**, 758–762 (2003).
3. Cartwright, D., Langcake, P., Pryce, R. J., Leworthy, D. P. & Ride, J. P. Chemical activation of host defence mechanisms as a basis for crop protection. *Nature* **267**, 511–513 (1977).
4. Snyder, B. A. & Nicholson, R. L. Synthesis of phytoalexins in *Sorghum* as a site-specific response to fungal ingress. *Science* **248**, 1637–1639 (1990).
5. Frey, M. et al. Analysis of a chemical plant defense mechanism in grasses. *Science* **277**, 696–699 (1997).
6. Goswami, R. S. & Kistler, H. C. Heading for disaster: *Fusarium graminearum* on cereal crops. *Mol. Plant Pathol.* **5**, 515–525 (2004).
7. Mueller, D. et al. Corn yield loss estimates due to diseases in the United States and Ontario, Canada from 2012 to 2015. *Plant Health Prog.* **17**, 211–222 (2016).
8. Genetics for a warming world. *Nat. Genet.* **51**, 1195–1195 (2019).
9. Dixon, R. A. Natural products and plant disease resistance. *Nature* **411**, 843–847 (2001).
10. Jones, J. D. G. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–329 (2006).
11. van Loon, L. C., Rep, M. & Pieterse, C. M. J. Signifi ance of inducible defense related proteins in infected plants. *Annu. Rev. Phytopathol.* **44**, 135–162 (2006).
12. Moghe, G. D. & Kruse, L. H. The study of plant specialized metabolism: challenges and prospects in the genomics era. *Am. J. Bot.* **105**, 959–962 (2018).
13. Wouters, F. C., Blanchette, B., Gershenzon, J. & Vassao, D. G. Plant defense and herbivore counter-defense: benzoxazinoids and insect herbivores. *Phytochem. Rev.* **15**, 1127–1151 (2016).
14. Meihls, L. N. et al. Natural variation in maize aphid resistance is associated with 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one glucoside methyltransferase activity. *Plant Cell* **25**, 2341–2355 (2013).
15. Fraenkel, G. S. The raison d'etre of secondary plant substances; these odd chemicals arose as a means of protecting plants from insects and now guide insects to food. *Science* **129**, 1466–1470 (1959).
16. Schmelz, E. A. et al. Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins. *Plant J.* **79**, 659–678 (2014).
17. Vaughan, M. M. et al. Accumulation of terpenoid phytoalexins in maize roots is associated with drought tolerance. *Plant Cell Environ.* **38**, 2195–2207 (2015).
18. Ding, Y. et al. Multiple genes recruited from hormone pathways partition maize diterpenoid defences. *Nat. Plants* **10**, 1043–1056 (2019).
19. Casas, M. I. et al. Identifi ation and characterization of maize salmon silks genes involved in insecticidal maysin biosynthesis. *Plant Cell* **28**, 1297–1309 (2016).
20. Degenhardt, J. Indirect defense responses to herbivory in grasses. *Plant Physiol.* **149**, 96–102 (2009).
21. Zila, C. T., Samayoa, L. F., Santiago, R., Butron, A. & Holland, J. B. A genome-wide association study reveals genes associated with *Fusarium* ear rot resistance in a maize core diversity panel. *G3* **3**, 2095–2104 (2013).
22. Wiesner-Hanks, T. & Nelson, R. Multiple disease resistance in plants. *Annu Rev. Phytopathol.* **54**, 229–252 (2016).
23. Stagnati, L. et al. A genome wide association study reveals markers and genes associated with resistance to *Fusarium verticillioides* infection of seedlings in a maize diversity panel. *G3* **9**, 571–579 (2019).
24. Banerjee, A. & Hamberger, B. P450s controlling metabolic bifurcations in plant terpene specialized metabolism. *Phytochem. Rev.* **17**, 81–111 (2018).
25. Karunanithi, P. S. & Zerbe, P. Terpene synthases as metabolic gatekeepers in the evolution of plant terpenoid chemical diversity. *Front. Plant Sci.* **10**, 1166 (2019).
26. Block, A. K., Vaughan, M. M., Schmelz, E. A. & Christensen, S. A. Biosynthesis and function of terpenoid defense compounds in maize (*Zea mays*). *Planta* **249**, 21–30 (2019).
27. Springer, N. M. et al. The maize W22 genome provides a foundation for functional genomics and transposon biology. *Nat. Genet.* **50**, 1282–1288 (2018).
28. Ding, Y. Z. et al. Selinene volatiles are essential precursors for maize defense promoting fungal pathogen resistance. *Plant Physiol.* **175**, 1455–1468 (2017).
29. Huff ker, A. et al. Novel acidic sesquiterpenoids constitute a dominant class of pathogen-induced phytoalexins in maize. *Plant Physiol.* **156**, 2082–2097 (2011).
30. Mafu, S. et al. Discovery, biosynthesis and stress-related accumulation of dolabradiene-derived defenses in maize. *Plant Physiol.* **176**, 2677–2690 (2018).
31. Mao, H., Liu, J., Ren, F., Peters, R. J. & Wang, Q. Characterization of CYP71Z18 indicates a role in maize zealexin biosynthesis. *Phytochemistry* **121**, 4–10 (2016).
32. Basse, C. W. Dissecting defense-related and developmental transcriptional responses of maize during *Ustilago maydis* infection and subsequent tumor formation. *Plant Physiol.* **138**, 1774–1784 (2005).
33. Kollner, T. G. et al. Protonation of a neutral (*S*)-β-bisabolene intermediate is involved in (*S*)-β-macrocarpene formation by the maize sesquiterpene synthases TPS6 and TPS11. *J. Biol. Chem.* **283**, 20779–20788 (2008).
34. Christensen, S. A. et al. Fungal and herbivore elicitation of the novel maize sesquiterpenoid, zealexin A4, is attenuated by elevated CO_2. *Planta* **247**, 863–873 (2018).
35. van der Linde, K., Kastner, C., Kumlehn, J., Kahmann, R. & Doehlemann, G. Systemic virus-induced gene silencing allows functional characterization of maize genes during biotrophic interaction with *Ustilago maydis*. *N. Phytol.* **189**, 471–483 (2011).
36. Shen, Q. et al. CYP71Z18 overexpression confers elevated blast resistance in transgenic rice. *Plant Mol. Biol.* **100**, 579–589 (2019).

37. Medema, M. H. et al. Minimum information about a biosynthetic gene cluster. *Nat. Chem. Biol.* **11**, 625–631 (2015).
38. Kersten, R. D., Diedrich, J. K., Yates, J. R. 3rd & Noel, J. P. Mechanism-based post-translational modifi ation and inactivation in terpene synthases. *ACS Chem. Biol.* **10**, 2501–2511 (2015).
39. Flint-Garcia, S. A. et al. Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant J.* **44**, 1054–1064 (2005).
40. Kremling, K. A. G. et al. Dysregulation of expression correlates with rare-allele burden and fitness loss in maize. *Nature* **555**, 520–523 (2018).
41. Jones, C. G. et al. Sandalwood fragrance biosynthesis involves sesquiterpene synthases of both the terpene synthase (TPS)-a and TPS-b subfamilies, including santalene synthases. *J. Biol. Chem.* **286**, 17445–17454 (2011).
42. Swigonova, Z. et al. Close split of *Sorghum* and maize genome progenitors. *Genome Res.* **14**, 1916–1923 (2004).
43. Lee, M. et al. Expanding the genetic map of maize with the intermated B73 x Mo17 (IBM) population. *Plant Mol. Biol.* **48**, 453–461 (2002).
44. McMullen, M. D. et al. Genetic properties of the maize nested association mapping population. *Science* **325**, 737–740 (2009).
45. Eichten, S. R. et al. B73-Mo17 near-isogenic lines demonstrate dispersed structural variation in maize. *Plant Physiol.* **156**, 1679–1690 (2011).
46. Suzuki, R., Iijima, M., Okada, Y. & Okuyama, T. Chemical constituents of the style of *Zea mays* L. with glycation inhibitory activity. *Chem. Pharm. Bull.* **55**, 153–155 (2007).
47. Ahmad, S. et al. Benzoxazinoid metabolites regulate innate immunity against aphids and fungi in maize. *Plant Physiol.* **157**, 317–327 (2011).
48. Smissman, E. E., Lapidus, J. B. & Beck, S. D. Corn plant resistance factor. *J. Org. Chem.* **22**, 220–220 (1957).
49. Pollastri, S. & Tattini, M. Flavonols: old compounds for old roles. *Ann. Bot.* **108**, 1225–1233 (2011).
50. Lange, B. M. The evolution of plant secretory structures and emergence of terpenoid chemical diversity. *Annu. Rev. Plant Biol.* **66**, 139–159 (2015).
51. Balmer, D., de Papajewski, D. V., Planchamp, C., Glauser, G. & Mauch-Mani, B. Induced resistance in maize is based on organ-specific defence responses. *Plant J.* **74**, 213–225 (2013).
52. Yang, F. et al. A maize gene regulatory network for phenolic metabolism. *Mol. Plant* **10**, 498–515 (2017).
53. Benson, J. M., Poland, J. A., Benson, B. M., Stromberg, E. L. & Nelson, R. J. Resistance to gray leaf spot of maize: genetic architecture and mechanisms elucidated through nested association mapping and near-isogenic line analysis. *PLoS Genet.* **11**, e1005045 (2015).
54. Kump, K. L. et al. Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nat. Genet.* **43**, 163–168 (2011).
55. Zuo, W. et al. A maize wall-associated kinase confers quantitative resistance to head smut. *Nat. Genet.* **47**, 151–157 (2015).
56. Hamilton, R. H. A corn mutant deficie t in 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one with an altered tolerance of atrazine. *Weeds* **12**, 27–30 (1964).
57. Meunier, B., de Visser, S. P. & Shaik, S. Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes. *Chem. Rev.* **104**, 3947–3980 (2004).
58. Henriques de Jesus, M. P. R. et al. Tat proteins as novel thylakoid membrane anchors organize a biosynthetic pathway in chloroplasts and increase product yield 5-fold. *Metab. Eng.* **44**, 108–116 (2017).
59. Laursen, T. et al. Characterization of a dynamic metabolon producing the defense compound dhurrin in *Sorghum*. *Science* **354**, 890–893 (2016).
60. Nelson, R., Wiesner-Hanks, T., Wisser, R. & Balint-Kurti, P. Navigating complexity to breed disease-resistant crops. *Nat. Rev. Genet.* **19**, 21–33 (2018).
61. Chappell, J. & Hahlbrock, K. Transcription of plant defense genes in response to UV-light or fungal elicitor. *Nature* **311**, 76–78 (1984).
62. Facchini, P. J. & Chappell, J. Gene family for an elicitor-induced sesquiterpene cyclase in tobacco. *Proc. Natl Acad. Sci. USA* **89**, 11088–11092 (1992).
63. Koutsoudis, M. D., Tsaltas, D., Minogue, T. D. & von Bodman, S. B. Quorum-sensing regulation governs bacterial adhesion, biofilm development, and host colonization in *Pantoea stewartii* subspecies *stewartii*. *Proc. Natl Acad. Sci. USA* **103**, 5983–5988 (2006).
64. Doblas-Ibanez, P. et al. Dominant, heritable resistance to Stewart's wilt in maize is associated with an enhanced vascular defense response to infection with *Pantoea stewartii*. *Mol. Plant Microbe Interact.* **32**, 1581–1597 (2019).
65. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flex ble trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
66. Kim, D., Landmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
67. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).
68. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an effici t general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
69. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
70. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
71. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
72. Schnable, P. S. et al. The B73 maize genome: complexity, diversity, and dynamics. *Science* **326**, 1112–1115 (2009).
73. Sun, S. et al. Extensive intraspecific gene order and gene structural variations between Mo17 and other maize genomes. *Nat. Genet.* **50**, 1289–1295 (2018).

Acknowledgements

We thank A. Steinbrenner, J. Chan, K. O'Leary, M. Broemmer, H. Riggleman, S. Reyes and S. Delgado for help with planting, treatments and sampling (UCSD); L. Smith (UCSD) for shared UCSD Biology Field Station management; B. Hamberger (Michigan State University) for the *ElHMGR* gene. This work was partially supported by the USDA-ARS National Programs for Food Safety and Plant Genetic Resources, Genomics and Genetic Improvement (to M.M.V., M.G.B. and S.A.C.). Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA. Research was supported by grants from the National Science Foundation, Division of Integrative Organismal Systems (NSF-IOS) (grant no. 1936492 to B.Y. and grant no. 1546899 to S.P.B.), USDA NIFA AFRI (grant no. 2018-67013-28125 to A.H. and E.S.) for sesquiterpenoids, NSF Plant-Biotic Interactions Program (grant no. 1758976 to E.S. and P.Z.) for diterpenoids, NSF Faculty Early Career Development Program (grant no. 1943591 to A.H.), the DOE Joint Genome Institute Community Science Program (JGI-CSP) (grant nos. CSP 2568 (to P.Z., J.B., E.S. and A.H.) and CSP 503420 (to A.H. and E.S.)) and fellowships provided by the NSF Graduate Research Fellowship Program (to K.M.M.), the U.C. Davis Innovation Institute for Food and Health (IIFH) Fellowship Program (to K.M.M. and P.Z.), the USDA NIFA Predoctoral Fellowship Program (award no. 2019-67011-29544, to K.M.M.) and a Fulbright Research Grant (E0581299, to M.B.).

Author contributions

Y.D., P.R.W., E.P., P.Z., J.S., J.B., K.D., E.A.S. and A.H. designed the experiments and analysed the data. Y.D., E.P., S.A.C., T.G.K., P.Z., K.A.K. and E.S.B. designed, performed and analysed the transcriptome data. Y.D., E.S., A.S.K., K.M.M., P.Z., A.H. and E.A.S. performed MS experiments and MS-related metabolite data analysis. Y.D., E.S., K.M.M., P.Z., E.A.S. and A.H. performed and analysed the enzyme coexpression data. Z.S., A.-D.T. and S.P.B. analysed the combined proteome and transcriptome dataset. T.K. calculated estimates of gene evolution dates. D.R.N. assigned subfamily names for P450 proteins. M.M.V. and M.G.B. generated and analysed the root microbiome data. B.Y., S.N.C. and P.R.W. designed gRNA constructs and generated the *zx1 zx2 zx3* and *zx1 zx2 zx3 zx4* maize mutants. J.S. and M.B. performed metabolite purifications and analysed the NMR data. Y.D. and P.R.W. performed the in vitro and in vivo antibiotic resistance assays. Y.D., P.R.W., E.P., P.Z., E.A.S. and A.H. wrote the manuscript with input from all of the authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41477-020-00787-9.

Correspondence and requests for materials should be addressed to A.H.

Peer review information: *Nature Plants* thanks Peter Balint-Kurti, Hugues Renault and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Alisa Huffaker

Last updated by author(s): 8/25/2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	GC-MS and LC-MS data were collected using Agilent MassHunter Workstation Software vB.08.00. Peptide spectra were acquired using Xcalibur 4.0 (Thermo Scientific) and NMR experiments were conducted on a Bruker Avance III console as well as an Agilent 600-MHz 13C direct detect cryoprobe. qRTPCR data were collected using CFX Manager 3.1 (Bio-Rad) on a Bio-Rad CFX96TM Real-Time PCR Detection System.
Data analysis	Data and statistical analyses were conducted using GraphPad Prism 8.0 (GraphPad Software, Inc.) and JMP Pro 13.0 (SAS Institute Inc). BioEdit v7.0.5 and MEGA7 v7.0.26 were used for phylogenetic analysis. GC-MS and LC-MS data were analyzed using Agilent MassHunter Workstation Software vB.08.00. Peptide data was extracted and analyzed using Spectrum Mill vB.06 (Agilent Technologies). Genome wide association studies were performed initially using the Unified Mixed Linear Model (MLM) in TASSEL 5.0 with final analyses conducted with the R package GAPIT 3.0 using the Compressed MLM. Manhattan plots were constructed in the R package qqman (v0.1.4) (http://cran.r project.org/web/packages/qqman).For comparative gene cluster analyses across inbreds, the genomic sequences of each cluster were extracted and compared using NCBI standalone blastn software, specifically BLAST+ v2.10.1 and further used genoPlotR v0.8.9, MAFFT v7.47 and IQ-TREE v1.6.1216 for final outputs. W22 cDNA concentration was quantified using a fluorometer running Qubit v2.0 firmware to assess and normalize of RNA for RNA-seq analyses. R packages Trimmomatic v0.39, Hisat2 v2.0.0, Sambamba v0.6.8, featureCounts v1.6.4, and DESeq2 were used for 3' RNA-seq data analysis, and R packages, TopHat v2.0.12 and HTSeq v0.6.1, were carried out for the whole transcript RNA-seq data analysis. Weighted Correlation Network Analysis (WGCNA) R package (Version 1.69) was used to cluster genes with similarly expressed proteins and similarly expressed RNA into modules following rank ordering for the W22 time course data. R package topGO (version 2.36.0) was used to find functional enrichment in conjunction with the maize-GAMER data set. R package system PipeR (Version 0.6.1.3) was utilized to predict maize upstream open reading frames. Amplicon sequences from root microbiome profiling were processed with the DADA2 pipeline in R v.3.5. The packages ggplot2 (Version 3.3.2) and pheatmap (Version 1.0.12) were used for root microbiome data analysis. Gene duplication date estimation was carried out using BEAST v2.6.1 and visualized with DensiTree v2.2.7. For the gene duplication date estimation, scripts to perform translation, alignment, and backtranslation as well as data files and BEAST control files have been deposited on GitHub

(https://github.com/TomJKono/Zealexin_Dating).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Publicly available datasets used in the study include the National Center for Biotechnology Information (NCBI) Sequence Read Archive project ID SRP115041 (https://www.ncbi.nlm.nih.gov/sra) and the MaizeGDB BLAST (https://maizegdb.org/popcorn/main/index.php) database. Raw read sequences have been deposited in the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under the accession numbers GSE138961 and GSE138962. Raw sequence data from the root microbiome are available at NCBI BioProject (https://www.ncbi.nlm.nih.gov/bioproject/) accession number PRJNA580260. Raw proteomic mass spectra have been deposited at the Mass Spectrometry Interactive Virtual Environment repository (ftp://massive.ucsd.edu/MSV000084285). Maize related germplasm used in this research have been previously described (1-7) and can be obtained from United State Department of Agriculture Germplasm Resources Information Network (https://www.ars-grin.gov) and the Maize Genetics Cooperation Stock Center (http://maizecoop.cropsci.uiuc.edu). Where possible, gene identifiers used throughout the manuscript were in reference to B73 RefGen_v4 (https://www.maizegdb.org/genome/assembly/Zm-B73-REFERENCE-GRAMENE-4.0) which served as a foundation of the study. All figures have associated raw data available upon request.

References to detailed description of germplasm used:

1 Lee, M. et al. Expanding the genetic map of maize with the intermated B73 x Mo17 (IBM) population. Plant Mol. Biol. 48, 453-461 (2002).
2 Flint-Garcia, S. A. et al. Maize association population: a high-resolution platform for quantitative trait locus dissection. Plant J. 44, 1054-1064 (2005).
3 McMullen, M. D. et al. Genetic Properties of the Maize Nested Association Mapping Population. Science 325, 737-740 (2009).
4 Schnable, P. S. et al. The B73 maize genome: complexity, diversity, and dynamics. Science 326, 1112-1115 (2009).
5 Eichten, S. R. et al. B73-Mo17 Near-Isogenic Lines Demonstrate Dispersed Structural Variation in Maize. Plant Physiol. 156, 1679-1690 (2011).
6 Springer, N. M. et al. The maize W22 genome provides a foundation for functional genomics and transposon biology. Nat. Genet. 50, 1282-1288 (2018).
7 Sun, S. et al. Extensive intraspecific gene order and gene structural variations between Mo17 and other maize genomes. Nat. Genet. 50, 1289-1295 (2018).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No specific statistical methods were used to determine sample size for the experiments reported in this manuscript. Instead, sample sizes were chosen based on our previous experience in with maize and tobacco, variability encountered, availability of plants and space considerations. Unstressed maize plants produce exceedingly low levels of zealexins and fungal elicitation elicits highly consistent qualitative and quantitative differences in production. With a broad dynamic range in zealexin accumulation following elicitation, a modest number of biological replicates are sufficient to detect differences. In every case, a sample size of at least 3 independent biological replicates was selected for purposes of statistical comparisons. These sample sizes proved to be sufficient to detect significant differences between treatments / genotypes and to obtain reproducible results in independent runs of the same experiment.
Data exclusions	No data were excluded from analysis in this study.
Replication	When technical problems were eliminated and methods described were followed all experiments could be reliably reproduced. All experiments were conducted at least two times. The independent experimental replications gave similar results.
Randomization	In this study, maize and tobacco plants were used as the dominant model organisms. Plants were grown in the field, greenhouse and growth chambers and randomly used for experiments.
Blinding	Initial RNA-seq data analyses were done blind. Also blinding was performed for the forward genetic studies and all biochemical analyses. In these cases, plants and plant analytical samples were assigned a simple numerical series for each experiment and the identities of the samples/results were not revealed until final analyses of processed data. Regarding Nicotiana benthamiana plants used for the Agrobacterium mediated heterologous expression of enzymes, blinding was not attempted due to the large number of chemicals, solutions, and vectors that had to be prepared and precisely combined for the plant treatments. While blinding was not practical during treatments, subsequent downstream steps were performed blinded with samples reduced to a simple numerical series and processed by different team members.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

nature research

https://doi.org/10.1038/s41477-020-00787-9

Supplementary information

Genetic elucidation of interconnected antibiotic pathways mediating maize innate immunity

In the format provided by the authors and unedited

Supplementary Information

Genetic elucidation of interconnected antibiotic pathways mediating maize innate immunity

Yezhang Ding[1], Philipp R. Weckwerth[1], Elly Poretsky[1], Katherine M. Murphy[2], James Sims[3], Evan Saldivar[1], Shawn A. Christensen[4], Si Nian Char[5], Bing Yang[5,6], Anh-dao Tong[1], Zhouxin Shen[1], Karl A. Kremling[7], Edward S. Buckler[7,8], Tom Kono[9], David R. Nelson[10], Jörg Bohlmann[11], Matthew G. Bakker[12,13], Martha M. Vaughan[12], Ahmed S. Khalil[1], Mariam Betsiashvili[1], Keini Dressano[1], Tobias G. Köllner[14], Steven P. Briggs[1], Philipp Zerbe[2], Eric A. Schmelz[1], and Alisa Huffaker[1*]

[1]Section of Cell and Developmental Biology, University of California at San Diego, La Jolla, CA USA; [2]Department of Plant Biology, University of California Davis, One Shields Avenue, Davis, CA, USA; [3]ETH Zurich, Institute of Agricultural Sciences, Zurich, Switzerland, [4]Chemistry Research Unit, Center for Medical, Agricultural, and Veterinary Entomology, Department of Agriculture–Agricultural Research Service, Gainesville, FL, USA; [5]Division of Plant Sciences, Bond Life Sciences Center, University of Missouri, Columbia, MO, USA; [6]Donald Danforth Plant Science Center, St. Louis, MO, USA; [7]Department of Plant Breeding and Genetics, Cornell University, Ithaca, NY, USA; [8]United States Department of Agriculture-Agricultural Research Service, Robert W. Holley Center for Agriculture and Health, Ithaca, New York, USA; [9]Minnesota Supercomputing Institute, University of Minnesota, Minneapolis, MN USA; [10]University of Tennessee Health Science Center, Memphis, TN, USA; [11]Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, Canada. [12]National Center for Agricultural Utilization Research, United States Department of Agriculture-Agricultural Research Service, Peoria, IL, USA. [13]Department of Microbiology, University of Manitoba, Winnipeg, Manitoba, Canada. [14]Max Planck Institute for Chemical Ecology, Jena, Germany

*Correspondence to: Alisa Huffaker, ahuffaker@ucsd.edu

This PDF file includes:

1. **Supplementary Methods**
2. **Supplementary Figures 1 to 32**
3. **Supplementary Table 8**
4. **Supplementary References 1-58**

1. Supplementary Methods

5' RACE cDNA library construction and cloning of *Zx* cDNAs. Total RNA was isolated from 35-day-old B73, Mo17 and W22 meristem tissues elicited with heat-killed *F. venenatum* hyphae collected at 48 h as described above. Approximate 2 µg total RNA was used for the construction of a 5' rapid amplification of cDNA ends (RACE) cDNA library with the SMARTer RACE 5'/3' Kit (Clontech) in accordance with the manufacturer's protocol. Genes with full-length open reading frames (ORFs) were amplified using gene-specific oligonucleotides (Supplementary Table 15). For *Agrobacterium*-mediated transient expression in *N. benthamiana*, full-length ORFs, including B73 *Zx3, B73 Zx4*, W22 *Zx3,* W22 Zx*4*, B73 *Zx5*, B73 *Zx7*, Mo17 *Zx9, and* Mo17 *Zx10* were amplified from cDNA library and cloned into the expression vector pLIFE33. Genes, including B73 *Zx1 B73 Zx2*, W22 *Zx1,* W22 *Zx2*, B73 *Zx6*, and Sorghum homolog of maize *Zx5* (*Sobic.001G235500*) were synthesized and subcloned into pLIFE33. In addition, *SaMonoTPS* (*Santalum album*, EU798692) on the plasmid pESC Leu2d was subcloned into pLIFE33[20]. All native and synthetic gene sequences used in this study for enzyme characterization are detailed (Supplementary Table 4).

Transient co-expression assays in *N. benthamiana.* For transient expression in *N. benthamiana*, pLIFE33 constructs carrying individual target genes and pEarleyGate100 with *ElHMGR*$^{159–582}$ construct[35] were electroporated into *Agrobacterium tumefaciens* strain GV3101. To ensure detectable production of sesquiterpenoid pathway products, all assays utilized co-expression of the coding sequence for truncated cytosolic *Euphorbia lathyris* 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR; ElHMGR$^{159–582}$, JQ694150.1)[1]. An *A. tumefaciens* strain encoding the P19 protein was also equally added in order to suppress host gene silencing. *Agrobacterium* cultures were separately prepared at OD_{600} of 0.8 in 10 mM MES pH 5.6, 10 mM $MgCl_2$, mixed together in equal proportion, and then infiltrated into the newly fully expanded leaves of six week old *N. benthamiana* plants using a needleless syringe[2]. Three days post

infiltration (dpi), sesquiterpene volatiles from *Agrobacterium*-inoculated tobacco leaves were collected by passing purified air over the samples at 600 ml min^{-1} and trapped on inert filters containing 50 mg of HayeSep Q (80- to 100-μm mesh) polymer adsorbent (Sigma-Aldrich). Individual samples were then eluted with 150 µl of methylene chloride and analyzed by GC/EI-MS. For analyzing non-volatile sesquiterpenoids, *Agrobacterium*-inoculated leaves were harvested at 5 dpi for further metabolite analysis.

Co-expression of TPSs and P450s in *E. coli*. Microbial co-expression of TPS and P450 enzymes was conducted using an established *E. coli* system engineered for enhanced terpenoid production[3,4]. For functional analysis, an N-terminally truncated Zx3 gene (lacking the predicted plastid transit peptide) and a full-length gene of the maize farnesyl diphosphate synthase (*ZmFPS3*, *Zm0001d043727*) were inserted into the pCOLA-Duet1 expression vector (EMD Millipore) to generate the construct. For co-expression of the P450s ZX6, ZX7, ZX8, ZX9 and ZX10, N-terminally modified[5] and codon-optimized genes were synthesized and subcloned into the pET-Duet1 expression vector (EMD Millipore) carrying the maize cytochrome reductase (ZmCPR2, Zm00001d026483), resulting in the constructs pET-Duet1:ZmCPR2/ZX6/ZX10. These individual constructs were then co-expressed with pCOLA-Duet1:ZmFPS3/ZX3 using the expression of pCOLA-Duet1:ZmFPPS/ZX3 only as a control. For further functional analysis of ZX7 with other P450s an additional pACYC-Duet1:ZmFPS3/ZX7 construct was generated and co-expressed with pET-Duet1:ZmCPR2/ZX8, ZX9 or ZX10. The desired construct combinations were co-transformed into *E. coli* strain BL21DE3-C41 cells (Lucigen) together with pCDFDuet:IRS for enhanced precursor formation[3]. Cultures were grown in 50 ml Terrific Broth (TB) medium to an OD_{600} of ~0.6 at 37°C and cooled to 16°C before protein expression was induced by adding 1 mM isopropyl-thio-galactoside (IPTG), followed by incubation for 72 h with supplement of 25 mM sodium pyruvate, 4 mg l^{-1} riboflavin, and 75 mg l^{-1} δ-aminolevulinic acid as previously described[4]. Organic solvent extraction of enzyme products was

performed with 50 ml of 1:1 ethyl acetate:hexane (v/v), followed by sample concentration under N_2 stream. Samples were resuspended in 200 µl methanol, treated with 10 µl of 1M (trimethylsilyl)diazomethane for one hour to methylate the compounds, then concentrated under N_2 stream again. Samples were then re-suspended in 1 ml hexane for mass spectral analysis.

Gas chromatography/mass spectrometry (GC/MS) analyses of metabolites. Maize and *N. benthamiana* tissue samples were frozen in liquid N_2, ground to a powder and stored at -80°C until further analyses. Tissue aliquots were weighed to 50 mg, solvent extracted in a bead homogenizer, derivatized using trimethylsilyldiazomethane, and collected using vapor phase extraction as described previously[6,7]. For metabolite extraction from *N. benthamiana*, tissue aliquots were subjected to β-glucosidase treatment (Sigma-Aldrich, Co, LLC, USA) in 250 µl 0.1 M sodium acetate buffer (pH=5.5) at a concentration of 100 units ml^{-1} at 37°C for 30 minutes before solvent extraction. GC-MS analysis was conducted using an Agilent 6890 series gas chromatograph coupled to an Agilent 5973 mass selective detector (interface temperature, 250°C; mass temperature, 150°C; source temperature, 230°C; electron energy, 70 eV). The gas chromatograph was operated with a DB-35MS column (Agilent; 30 m, 250 µm i.d., 0.25 µm film). The sample was introduced as a splitless injection with an initial oven temperature of 45°C. The temperature was held for 2.25 min, then increased to 300°C with a gradient of 20°C min^{-1}, and held at 300°C for 5 min. Unless otherwise noted, GC/EI-MS quantification of ZXs pathway products was based on the slope of an external standard curve constructed from β-costic acid (Ark Pharm; no. AK168379) spiked into 50-mg aliquots untreated maize stem tissues identically processed using vapor phase extraction[7]. With consideration of relative retention times on a DB35 column, diagnostic EI fragments used in this study are as follows β-bisabolene (*m*/*z* 248 parent ion, *m*/*z* 93 fragment ion), β-macrocarpene (*m*/*z* 204 parent ion, *m*/*z* 136 fragment ion), ZD1 (*m*/*z* 248 parent ion, *m*/*z* 93/69 fragment ions), ZD1 (*m*/*z* 248 parent ion, *m*/*z* 93 fragment ion), ZA1 (*m*/*z* 248 parent ion, *m*/*z* 136 fragment ion), ZB1 (*m*/*z* 246 parent ion),

ZA5 (*m*/*z* 136 fragment ion), ZA2 (*m*/*z* 204 fragment ion), ZA3 (*m*/*z* 176 fragment ion), ZC2 (*m*/*z* 260 parent ion), and ZB3 (*m*/*z* 229 fragment ion). To analyze complex ZX profiles following *C.h.* inoculation, we used an isobutane-chemical ionization-GC/MS method better suited for deconvolution of co-chromatography challenges[6,7]. In this situation analytes of interest produced the following diagnostic ions (*m*/*z*) and retention times (RT) in order; β-bisabolene $[M+H]^+$ *m*/*z* 205, RT 9.81 min; β-macrocarpene $[M+H]^+$ *m*/*z* 205, RT, 9.85 min; α/β-costic acids $[M+H]^+$ *m*/*z* 249, RT, 12.86 min; ZD2 $[M+H]^+$ *m*/*z* 249, RT 13.14 min; ZD1 $[M+H]^+$ *m*/*z* 249, RT 13.27 min; ZA1 $[M+H]^+$ *m*/*z* 249, RT 13.55 min; ZC1 $[M+H]^+$ *m*/*z* 245, RT 14.12, ZB1 $[M+H]^+$ *m*/*z* 247, RT 14.51; ZA5 fragment $[M-H_2O]^+$ *m*/*z* 247 RT 14.88; ZA2 fragment $[M-H_2O]^+$ *m*/*z* 247, RT 14.93; ZA3 fragment $[M-H_2O]^+$ *m*/*z* 247, RT 15.32; ZC2 $[M+H]^+$ *m*/*z* 261, RT 15.65 min; ZB3 $[M+H]^+$ *m*/*z* 263, RT min 15.87; ZA4 $[M+H]^+$ *m*/*z* 263, RT 16.17 min; ZA6 fragment $[M-2H_2O]^+$ *m*/*z* 245, RT 16.75; ZA7 fragment $[M-2H_2O]^+$ *m*/*z* 245 , RT 17.20 min; ZA8 $[M+H]^+$ *m*/*z* 279, RT 17.21 min; ZA9 $[M+H]^+$ *m*/*z* 279, RT 17.48 min. Analytes were quantified using MassHunter Workstation Software (Agilent) based on an U-^{13}C-linolenic acid internal standard as previously described[6].

GC-MS analysis of *E. coli* expressed enzyme products was performed on an Agilent 7890B GC with a 5977 Extractor XL MS Detector at 70 eV and 1.2 ml min^{-1} He flow, using a HP5-MS column (30 m, 250 µm i.d., 0.25 µm film) with a sample volume of 1 µl and the following GC parameters: pulsed splitless injection at 250°C and 50°C oven temperature; hold at 50°C for 3 min, 20°C min^{-1} to 300°C, hold 3 min. MS data from 90 to 600 *m*/*z* ratio were collected after a 9 min solvent delay. Product identification was conducted using authentic standards and by comparison of reference mass spectra with Wiley, National Institute of Standards and Technology and the Adams libraries.

Liquid chromatography/mass spectrometry (LC/MS) analyses of maize benzoxazinoids, flavonoids and acidic terpenoids. LC/MS analyses were used to estimate the relative abundance of defense metabolite classes present following stem elicitation with heat killed *F. venenatum* in different aged plants.

Stem tissues where ground to a fine powder with liquid N_2 and 50 mg samples were sequentially and additively bead homogenized in 1) 100 µl 1-propanol: acetonitrile: formic acid (1:1:0.01), 2) 250 µl acetonitrile: ethyl acetate (1:1), and 3) 100 µl of H_2O. The co-miscible acidified solvent mixture of contained 1-propanol: acetonitrile: ethyl acetate: H_2O (11:39:28:22) which following centrifugation (15,000 rpm, 20 min) 5 µl was used for LC/MS analysis. The LC consisted of an Agilent 1260 Infinitely series HiP Degasser (G4225A), 1260 binary pump (G1312B), and a 1260 autosampler (G1329B). The binary gradient mobile phase consisted of 0.1% formic acid in H_2O (solvent A) and 0.1% formic acid in MeOH (solvent B). Analytical samples were chromatographically separated on a Zorbax Eclipse Plus C18 Rapid Resolution HD column (Agilent: 1.8 µm, 2.1 x 50 mm) using a 0.35 ml min^{-1} flow rate. The mobile phase gradient was: 0–2 min, 5% B constant ratio; 3 min, 24% B; 28 min, 98% B, 35 min, 98% B, and 36 min 5% B for column re-equilibration before the next injection. Eluted analytes underwent electrospray ionization (ESI) via an Agilent Jet Stream Source with thermal gradient focusing using the following parameters: nozzle voltage (500 V), N_2 nebulizing gas (flow 12 l min^{-1}, 55 psi, 225°C) and sheath gas (350°C, 12 l min^{-1}). The transfer inlet capillary was 3500V interfacing with an Agilent 6460 Triple Quad with both MS1 and MS2 heaters at 100°C. Negative ionization mode scans (0.1 amu steps, 2.25 cycles s^{-1}) from *m*/*z* 100 to 1000 were acquired. Using the conditions defined above the following retention times (min) and ions (m/z) were used to estimate relative changes in the abundance of maize defense metabolites. Benzoxazinoids included DIMBOA-Glc (split peak 5.63/6.58 min) $[M-H]^-$ *m*/*z* 372; HDMBOA-Glc (split peak 7.27/7.89 min) $[M+46 (formate)-H]^-$ *m*/*z* 432; and HDM_2BOA-Glc (split peak 6.66/7.65 min) $[M+46 (formate)-H]^-$ *m*/*z* 462. Flavonoids included genkwanin (18.25 min) $[M-H]^-$ *m*/*z* 282; naringenin (12.98 min) $[M-H]^-$ *m*/*z* 271; naringenin chalcone (13.99 min) $[M-H]^-$ *m*/*z* 271; apigenin (14.94 min) $[M-H]^-$ *m*/*z* 269; tetrahydroxyflavanone (9.33 min) $[M-H]^-$ *m*/*z* 287; and a dimethoxytetrahydroxyflavanone candidate (split peak 10.54/13.28 min) $[M-H]^-$ *m*/*z* 315. Estimates of total acidic terpenoids included A- and B-series kauralexin diterpenoids KB1 (26.59 min) $[M-H]^-$ *m*/*z* 301; KA1

(26.95 min) $[M-H]^-$ *m*/*z* 303; KB3 (21.13 min) $[M-H]^-$ *m*/*z* 315; KA3 (22.21 min) $[M-H]^-$ m/z 317, KA2 (21.14 min) $[M-H]^-$ *m*/*z* 333; KA4 (21.14 min) $[M-H]^-$ *m*/*z* 319 and acidic sesquiterpenoids α/β-costic acids (23.13 min) $[M-H]^-$ *m*/*z* 233; ZC1 (22.60 min) $[M-H]^-$ *m*/*z* 229; ZB1 (22.78 min) $[M-H]^-$ *m*/*z* 231; ZD1+ZD2 (combined peak 23.45 min) $[M-H]^-$ *m*/*z* 233; ZA1 (23.71 min) $[M-H]^-$ *m*/*z* 233; ZB3 (17.16 min) $[M-H]^-$ *m*/*z* 247; ZC2 (17.20 min) $[M-H]^-$ *m*/*z* 245; ZA3 (18.00 min) $[M-H]^-$ *m*/*z* 249; and ZA2 (19.59 min) $[M-H]^-$ *m*/*z* 249.

Homology modeling and ZX1 site-directed mutagenesis. A homology model of B73 β-macrocarpene synthase ZX1 was generated by using the SWISS-MODEL server(https://swissmodel.expasy.org/) based on the template for *Nicotiana tabacum* 5-*epi*-aristolochene synthase (5-EAT)[8]. Protein variants were generated by whole-plasmid PCR amplification with site-specific sense and anti-sense oligonucleotides (Supplementary Table 15), followed by Dpn I treatment to remove the parental template. All genes encoding variant proteins were sequence-verified before co-expression in *N. benthamiana*.

Mutual Rank (MR) and analyses of coregulated transcripts. The Goodman diversity panel RNA-seq dataset (B73 RefGen_V4) was composed of 300 inbred lines constituting 1960 developmentally diverse tissues samples previously deposited in the National Center for Biotechnology Information Sequence Read Archive project ID SRP115041 (https://www.ncbi.nlm.nih.gov/sra) [9]. Using 1960 samples, calculations of Mutual Rank (MR) were used as a measure of coexpression by calculating the geometric mean of the product of two-directional ranks derived from Pearson correlation coefficients across gene pairs[10,11]. To estimate predominant *Zx* expression patterns, expression of all *Zx* genes within in each cluster were summed across all tissue types for each inbred. Inbreds were then sorted to select the top 100 lines displaying the highest total gene expression for each cluster. From the group of 100 inbred lines, the percent of inbred lines with significant individual *Zx* gene expression for each cluster was calculated. Individual *Zx* genes contributing 5% or greater to the total cluster

expression were counted as expressed in each inbred. Individual *Zx* genes with expression less than 5% of the total *Zx* gene cluster expression in each inbred were considered weakly expressed and less likely to significantly contribute to the pathway in the given inbred.

Comparative gene cluster analyses. We used a combination of tools to identify tandem duplicate *Zx* genes in the three defined gene clusters using reference genomes of highland teosinte (*Zea mays* spp. *mexicana*; PI566673), B73, W22, Mo17 and all NAM inbred parents. Individual B73 exon sequences of *Zx1*, *Zx5* and *Zx8* were used as a reference to query the MaizeGDB BLAST (https://maizegdb.org/popcorn/main/index.php) database[12] and generate a list of putative exon sequences for gene clusters I, II and III, respectively. Adjacent and consecutive sequences were combined and used to generate putative *Zx* gene models in each cluster (Supplementary Table 6). The genomic sequences of each cluster were extracted and compared using NCBI standalone blastn software, specifically BLAST+ v2.10.1[13], (word_size 150) and the comparison results were visualized as synteny plots using genoPlotR v0.8.9 (Supplementary Fig. 7, 11 and 23)[14]. The nucleotide sequences of all putative *Zx* genes of each cluster (Supplementary Table 6) were aligned with MAFFT v7.47[15] using default parameters and Maximum Likelihood (ML) phylogenetic trees (Supplementary Fig. 8) were constructed using IQ-TREE v1.6.12[16] under default parameters with 1,000 bootstrap replications. The generated synteny plots and phylogenetic trees of each cluster guided the assignment of *Zx* orthology in reference genomes where available (Supplementary Table 7).

Genetic mapping of *Zx* biosynthetic genes. To search for genetic variation, we first screened the NAM parent founders, B73 and Mo17 for differences in zealexin production after 3 d of heat-killed *Fusarium* stem elicitation. Based on a selective deficit of ZB1 in Mo17, a field grown population of 216 IBM RILs[17] was employed using naturally occurring necrotic root tissues collected 30 days after pollination for analysis of the ratio of ZA1 to ZB1 as a mapping trait. The locus

responsible for ZB1 biosynthesis was further fine-mapped using select B73 x Mo17 NILs[18]. To utilize genetic variation in a larger population, the Goodman diversity panel[19] was grown in the greenhouse and stem tissues were harvested 3 d after elicitation with heat-killed *Fusarium*. Association analyses were conducted in TASSEL 5.0[20] using the General Linear Model (GLM) for the IBM RILs and the unified Mixed Linear Model (MLM) to effectively control for false positives arising from the differential population structure and familial relatedness in the Goodman diversity panel[21]. Differential population structure and familial relatedness are generally not significant features in biparental RIL populations; thus, GLM analyses were selected for the IBM RILS. A list of NAM parents, IBM RILs, B73 x Mo17 NILs, and specific diversity panel lines used for mapping in this study are given (Supplementary Table 14). Genotypic data from imputed IBM RIL SNP markers (July 2012 All Zea GBS final build; www.panzea.org) with less than 20% missing genotypes and a >15% minor allele frequency greater were used to generate 173,984 final SNP markers. GWAS analyses utilized the B73 version 2 referenced HapMap consisting of 246,477 SNPs as described[22]. Final GWAS analyses were conducted with the R package GAPIT 3.0 [23,24] and compressed MLM parameters to identify genomic regions putatively associated with the trait. The kinship matrix (K) was derived from the 246,477 SNPs and used jointly with population structure (Q) to improve association analysis[25]. Manhattan plots were constructed in the R package qqman (v0.1.4) (http://cran.r project.org/web/packages/qqman)[26].

Gene duplication date estimation. Coding sequences for *Zea mays* (B73 RefGenv4) were fetched from Ensembl Plants and *Sorghum bicolor* (v.3.1.1) coding sequences were fetched from Phytozome V12. The coding sequences were translated to AAs with the standard translation table and aligned with clustal-omega 1.2.4[27]. Clustal-omega was run with up to 10 refinement iterations (--iterations=10) and using the full distance matrix during iterations (--full-iter). The resulting AA alignments were back-translated to nucleotides using the original coding sequences as guides. The back-translated alignments were used

to estimate gene duplication dates. Date estimation was carried out using BEAST 2.6.1[28] and a general time reversible nucleotide substitution model and a random local clock[29]. We used a calibrated Yule model as the prior for the gene tree. The maize genes were set to form a monophyletic clade in the tree, and the distribution for the common ancestor of the maize genes was set to be normal with a mean of 11.9 million years ago [30] and a standard deviation of 1. The MCMC routine in BEAST was run for 10 million steps, and runtimes were improved by using the beagle phylogenetics library (https://github.com/beagle-dev/beagle-lib). Trees from BEAST were visualized with DensiTree 2.2.7[31].

Nucleic Acid Isolation and qrtPCR. Total RNA was isolated with a NucleoSpin® RNA Plant Kit (Takara Bio USA) from *N. benthamiana* leaves 2 days post infiltration with the *Agrobacterium tumefaciens* strain (GV3101) according to the manufacturer's protocol. First-strand cDNA was synthesized with SuperScript III First-Strand Synthesis SuperMix (Invitrogen, Grand Island, NY, USA). Quantitative real-time PCR (qrtPCR) was performed using Power SYBR Green Master mix (Applied Biosystems, Waltham, MA, USA), and 250 nM primers on a CFX96TM Real-Time PCR Detection System (Bio-Rad) using CFX Manager v3.1 (Bio-Rad) software. Mean cycle threshold values were normalized to the *N. benthamiana EF-1α*[32]. Fold-change calculations were performed using the equation $2^{-\Delta\Delta Ct}$. The sequences of qrtPCR primers used in the study are listed (Supplementary Table 15). For quantification of the fungal biomass, total DNA was extracted from fungal-inoculated maize stem tissues and subjected to qrtPCR using the *F. graminearum*-specific primers for a deoxynivalenol mycotoxin biosynthetic gene (*FgTri6*) (Supplementary Table 15)[33]. Plant DNA quantification was analyzed using specific primers (Supplementary Table 15) for the maize ribosomal protein *L17* gene (*ZmRLP17b*, *Zm00001d049815*). The relative amounts of fungal DNA were calculated by the $2^{-\Delta\Delta Ct}$ method, normalized to *ZmRLP17b* and expressed relative to those in damage-treated maize stems.

***In vitro* bioassays of ZB1 activity as an antifungal agent.** *In vitro* antifungal assays using purified ZA1 and ZB1 were performed using the Clinical and Laboratory Standards Institute M38-A2 guidelines as detailed[34]. In brief, a 96-well microtiter plate-based method using a Synergy4 (BioTech Instruments) reader was used to monitor fungal growth at 30°C in broth medium through periodic measurements of changes in optimal density (OD_{600} nm) for 48 h. Each well contained 200 μl of initial fungal inoculum (2.5×10^4 conidia ml^{-1}) with 1 μl of either pure dimethyl sulfoxide (DMSO) or DMSO containing 5 μg ZA1 or ZB1.

Identification of the *Zmcyp71z19* (*zx5*) mutant. The *Dsg* insertion (dsgR102G05) in W22 *Zx5* (Zm00001d014121, B73 RefGen_V4) was verified by designing PCR primer pairs, with one gene-specific pair (Supplementary Table 15) from W22 *Zx5* and one primer from the *Dsg GFP* insertion (GFP_AC-DS: TTCGCTCATGTGTTGAGCAT)[35].

Proteomic analysis of W22 stem tissues. As part of a previously described effort, W22 maize plants were grown individually in 1G pots for 35 days[10]. All plants were stem elicited with heat-killed with *F. venenatum* hyphae with staged timing to enable 10 time points (0, 2, 4, 8, 12, 24, 48, 72, 96, and 120 h) to be harvested within the same hour and age. Stem tissues from four plants were harvested and pooled to generate a single homogenous sample per time point, ground in liquid N_2 and stored at -80°C. Briefly, extracted proteins were digested with Lys-C (Wako Chemicals, 125-05061) for 15 min and secondarily digested with trypsin (Roche, 03 708 969 001) for 4 h as described[10]. TMT-10 labeling was performed and checked by LC–MS/MS to confirm >99% efficiency. Labeled peptides from each time point sample were pooled together for 2D-nanoLC–MS/MS analysis as described[10]. Spectra were acquired using Xcalibur 4.0 (Thermo Scientific) software on a Q-Exactive-HF mass spectrometer (ThermoFisher Scientific) and raw data was extracted and searched using Spectrum Mill vB.06 (Agilent Technologies)[10]. MS/MS spectra were searched against maize B73 V4 genome (Ensembl v36) with a concatenated 1:1 decoy

database of 263,022 total protein sequences. Peptides shared among different protein groups were removed before quantitation. False discovery rates (FDR) were set to 0.1% at the peptide level and 1% at the protein level, respectively. In a large-scale expansion an earlier effort, we now combine analyses of 2 separate technical LC/MS replicates, assign peptide sequences to the B73 genome, include 4 new intermediate time points (8, 12, 24 and 48 h) and report 10,749 unique protein groups beyond the original 13 previously reported[10].

Analyses of paired transcriptome and proteome changes following *Fusarium* elicitation. Analyses of proteome changes utilized two technical replicates with intensity values summed between runs. In cases where the fold change at a time point for one technical replicate had $\geqslant$ 5-fold difference from the same time point for other replicate, values from the run with the higher total intensities were used. To generate a co-expression heat map of 10,508 proteins and transcripts, we performed complete linkage hierarchical clustering of the W22 fold-change protein data using Cluster 3.0 [36] and combined the corresponding fold-change RNA-seq values to the cluster table. Uncentered Pearson correlations were used as the similarity metric and results were visualized in Java Treeview[37]. The W22 time course data was analyzed using the Weighted Correlation Network Analysis (WGCNA) R package[38] to cluster genes with similarly expressed proteins and similarly expressed RNA into modules following rank ordering. One-step network construction was performed using the blockwise modules function with a high sensitivity (deep split 4)[39]. Networks and topological overlap matrices were assigned with the minimum module size set to 30 and soft thresholding power of 5. The tree cutting algorithm was adaptive-height tree cut (Dynamic Tree Cut) and average linkage hierarchical clustering was used. To find functional enrichment, we used the R package topGO[40] (R package version 2.36.0) in conjunction with the maize-GAMER data set[41]. The R package system PipeR was used to predict maize upstream open reading frames[42] and topGO[97] (R package version 2.36.0) was used in conjunction with

the maize-GAMER data set[98]. The R package system PipeR was used to predict maize upstream open reading frames[99].

Isolation and NMR identification of ZXs. Field grown maize (*Z. mays* var. Golden Queen) stems (6 kg) 20 days post pollination were harvested, slit in half lengthwise with a scalpel, inoculated with *C. heterostrophus* (SLB) hyphae and allowed to incubate for 5 d at room temperature in the dark at 100% humidity. Husks from the same plants were coated with a thin slurry of heat-killed *F. venenatum* for 5 days. Following incubation, tissues were frozen in liquid N_2 and crushed to a coarse powder with dry ice in a hammer mill and stored at -20°C prior to extraction. Equal portions of the stem and husk tissues were combined (1 kg) and ground to a fine powder in liquid N_2. The powder was then allowed to thaw for 2 min and further ground in 2 l of ethyl acetate. The suspension was filtered through a Buchner funnel with Whatman#1 filter paper and resulting solvent concentrated *en vacuo* on a Buchi rotoevaporator until 20 ml remained. The remaining solution was directly absorbed onto 20 g C18 resin (Discovery®, DSC-18; Sigma Aldrich) by the evaporation of residual solvent in vacuo. The resulting oil was then dry loaded and separated by preparative flash chromatography (CombiFlash®Rf, Teledyne ISCO, Inc, Lincoln, NE, USA) on a 5g C18 flash column (Teledyne, RediSepRf High Performance Gold). The mobile phase consisted of solvent A [acetonitrile (ACN):H_2O, 20:80] and solvent B (ACN: 100) with A held constant the 5 min followed by a linear ramp to 100% B at 60 min using a flow rate of 18 ml min^{-1} and resulted in enriched mixtures containing distinct related ZX classes. Carboxylic acids present in fraction aliquots were derivatized with trimethylsilyldiazomethane and screened using GC/EI-MS analyses. Simple sesquiterpene acids lacking further oxygenation (ZA1, ZB1, ZD1, ZD2), those with an additional ketone or alcohol (ZA2-5, ZB3, ZC3), and those with two additional sites of oxygenation (ZA6-9) separated into 3 distinct fractions based on polarity. Each enriched flash fraction was further separated to yield pure compounds by preparative HPLC on a Dionex Ultimate 3000 instrument equipped with a YMC-Pack OD-AQ column (250 x 20mm, s-10 µm,

12 nm). Enriched flash fractions were dried under a N_2 stream (20 mg) dissolved in 200 μl methanol and re-chromatographed using a H_2O:ACN gradient and flow rate of 25 ml min^{-1}. The steepness of the gradient employed varied dependent on the target compounds polarity. More polar compounds, such ZA6-ZA9, had shallow gradients from 100% H_2O to 30% ACN over 45 min; whereas, the less polar ZD-series metabolites required a gradient of 30% ACN to 70% ACN over 45 min. Manual monitoring of ultraviolet (UV; 210 nm) signals and corresponding collection of narrow fractions enabled the final purification of previously unidentified ZXs. Structures were elucidated using ^{1}H and ^{13}C APT 1D NMR experiments, as well as correlated spectroscopy (COSY), heteronuclear single quantum correlation (HSQC) and heteronuclear multiple bond correlation (HMBC) 2D experiments. Additional 2D experiments were performed to help resolve overlaying signals such as nuclear Overhauser effect spectroscopy (NOESY), total correlation spectroscopy (TOCSY), HSQC-TOCSY and heteronuclear 2 bond correlation (H2BC). NMR experiments were performed in the McKnight Brain Institute at the National High Magnetic Field Laboratory's AMRIS Facility, which is supported by National Science Foundation Cooperative Agreement No. DMR-1157490 and the State of Florida. Purified ZXs were dissolved in chloroform-d ($CDCl_3$; Cambridge Isotope Laboratories) and NMR spectra were collected on a Bruker Avance II 600-MHz cryoprobe as well as an Agilent 600-MHz ^{13}C direct detect cryoprobe. Data was analyzed using Mnova (MestreLab) software. Chemical shifts were calculated by reference to chemical shifts as follows: ^{1}H 7.26 ppm and ^{13}C 77.4 ppm for $CDCl_3$; ^{1}H 7.16 ppm and ^{13}C 128.1 ppm for benzene-d_6; and ^{1}H 1.94 ppm and ^{13}C 1.4 ppm for acetonitrile-d_3 (Supplementary Table 8). Assignments were made directly from ^{1}H and ^{13}C APT data when possible, or inferred through 2D experiments such as HSQC or HMBC.

Sequence analysis and phylogenetic tree construction. Protein sequence alignments derived from UniProtKB and Genbank IDs (Supplementary Table 3) were performed using Clustal W as implemented in the BioEdit software package

(http://www.mbio.ncsu.edu/BioEdit/bioedit.html). The maximum-likelihood phylogenetic trees were constructed using MEGA7 (http://www.megasoftware.net/megabeta.php) with bootstrap values based on 1,000 iterations.

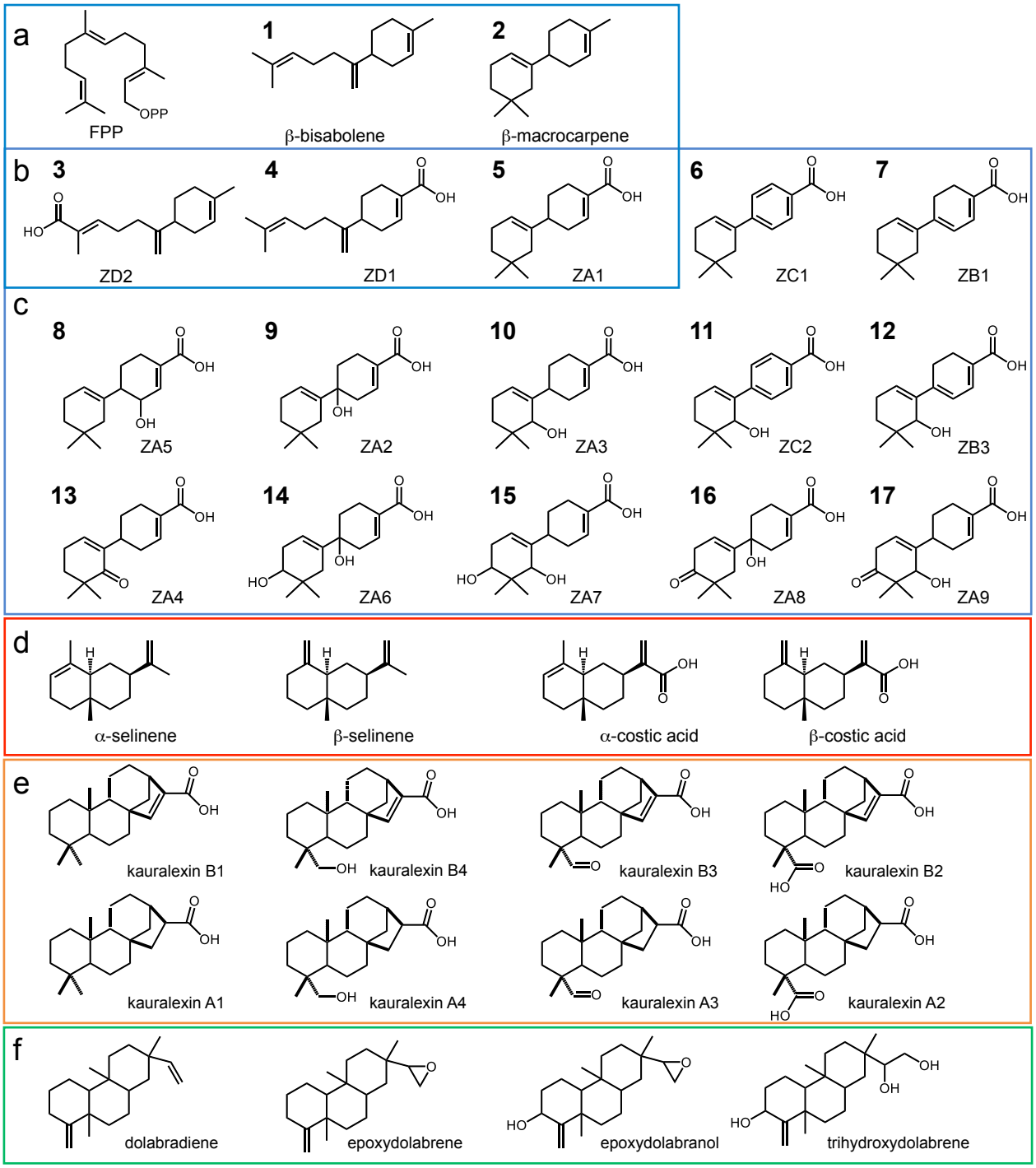
Creation of *zx1 zx2 zx3* and *zx1 zx2 zx3 zx4* mutants using CRISPR/Cas9. *Zx3* guide RNA (gRNA) target site selection was based on the B73 reference genome sequence and criteria as described[43]. Flanking regions with the target site at the middle were PCR-amplified from the maize genotype Hi-II and Sanger sequenced for accuracy of genomic sequence including the gRNA complementary sequence. The gRNA gene was constructed in the intermediate vector and the expression cassette was mobilized through a gateway reaction into the Cas9-expressing binary vector for maize Hi-II transformation at the Iowa State University Plant Transformation Facility as previously described[44]. A total of ten independent T0 transgenic plants were obtained. To examine if the target gene sequence was edited, the PCR amplicons encompassing the gRNA target site (Supplementary Fig. 28) from each plant were sequenced. Early in this effort, it was revealed that *ZmTPS6/11* were part of a 4 gene cluster evident in the B73 V4 genome[45] which reduced the frequency of complete null mutants. Ultimately, one *zx1 zx2 zx3* triple mutant and one *zx1 zx2 zx3 zx4* quadruple mutant were obtained. The homozygous mutant plants were outcrossed with B73 and the resulting F1 plants were self-pollinated to generate F2 progenies. Following genotyping, homozygous mutant plants without the CRISPR transgene were selected and backcrossed to B73.Two homozygous mutant plants, *zx1 zx2 zx3* and *zx1 zx2 zx3 zx4*, and two wild-type siblings were selected for bioassays by genotyping from self-pollinated plants after B73 backcrossing two additional times.

Root microbiome profiling. To investigate bacterial microbiomes associated with maize ZX knock out lines (*zx1 zx2 zx3* and *zx1 zx2 zx3 zx4*) and their corresponding wild type lines greenhouse grown plants were germinated in

individual 1 G pots mixed 1:1 with commercial potting soil (BM2; American Horticultural Supply, Inc) and field soil from the UCSD Biology Field Station (La Jolla, CA) where maize has been planted each year for 3 decades. After 8 weeks, all soil was gently shaken from the roots of mature plants, and sequentially rinsed with water, 70% ethanol (30 seconds) and distilled water to preferentially remove the external rhizosphere communities. Cleaned root tissues were then frozen in liquid N_2, ground to a fine powder and stored at -80°C. DNA was isolated from 100 mg aliquots of freeze-dried root samples using the PureLink™ Plant Total DNA Purification Kit (Invitrogen) as per kit protocols. DNA quality from *zx1 zx2 zx3 zx4* roots was highly compromised limiting replicates that could ultimately be obtained after 6 attempts. Final analyses combined ZX deficient lines (*zx1 zx2 zx3* and *zx1 zx2 zx3 zx4*) and also combined the respective wild type parents. Microbiome profiling was accomplished via amplicon sequencing, using primers 515F and 806R[46] to amplify the v4 region of bacterial 16S rRNA genes. Primers were modified with 5' overhangs for compatibility with the MiSeq workflow, and to create a frameshifted mixture of oligos to provide signal diversity when sequencing through the primer regions. Each PCR reaction mix consisted of 0.5 U Phusion High-Fidelity DNA polymerase with associated Phusion Green HF reaction buffer (Thermo Fisher), dNTPs at 200 µM final concentration, forward and reverse primers at 0.5 µM each, peptide nucleic acid blockers (PNA, Bio Inc) at 1 µM to prevent amplification of plastid and mitochondrial templates, 1-10 µl template DNA (depending on measured concentration) and nuclease free water to a total volume of 25 µl per reaction. Thermocycling consisted of 98 °C for 60 s, 25 cycles of (98 °C for 10 s, 75 °C for 10 s, 57 °C for 20 s, 72 °C for 15 s), final extension at 72 °C for 5 min. PCR products were cleaned using the SequalPrep Normalization Plate Kit (Thermo Fisher). An 8 cycle second round PCR was used to add sample-specific barcode indices, using the Nextera XT Index Kit (Illumina). The manufacturer's protocol was followed, except that we substituted Phusion High-Fidelity DNA polymerase for the suggested polymerase. The sequencing library also included negative control samples (i.e., DNA extractions

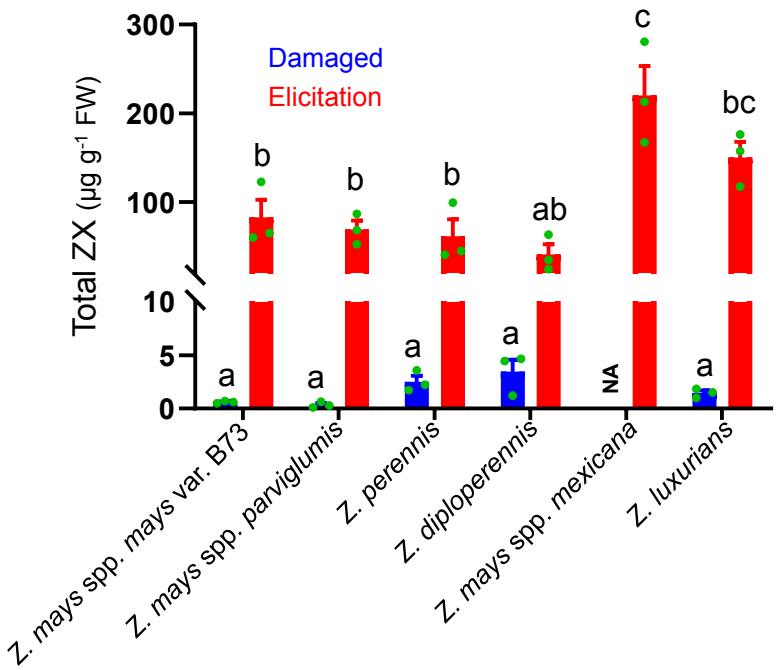
performed without any plant tissue, and PCRs run without any template DNA), and mock community control samples of known composition (20 Strain Staggered Mix Genomic Material; ATCC® MSA-1003™, American Type Culture Collection). Indexed amplicons were cleaned and normalized with the SequalPrep kit (ThermoFisher Scientific) ahead of sample pooling. Library quality and concentration were assessed with the TapeStation instrument (Agilent) and with the Library Quantification Kit for Illumina Platforms (Kapa Biosystems). Sequencing was performed with a MiSeq instrument (Illumina), using a version 2 (500 cycle) sequencing kit. Amplicon sequences were processed with the DADA2 pipeline in R v.3.5[47,48]. Briefly, primer sequences were located and trimmed using the tool Cutadapt[49], permitting a single mismatch. Reads were culled if no primer sequence was found, when lengths <50, or when they contained ambiguous base calls. Reads were trimmed at the trailing end, where quality tends to drop (20 bases for R1, 50 bases for R2), and filtered to permit a maximum of 2 expected errors[50]. True sequence variants were inferred from the observed sequences with the DADA2 algorithm [51]. Forward and reverse reads were merged, permitting one mismatch in the overlapping region. Chimeras were detected and removed using the DADA2 method. Sequence variants were assigned to taxonomic bins using a naïve Bayesian classifier[52], with the Silva reference alignment v. 132[53]. Reads were culled if they could not be classified below the rank of domain, or were classified as chloroplast or mitochondria. For assessment of phylogenetic diversity, a phylogenetic tree was constructed using the package phangorn [54], with a neighbor-joining tree as the starting point for a maximum likelihood tree (generalized time-reversible with Gamma rate variation). Further manipulations, visualization, and analyses used the package phyloseq [55]. Differential abundance of sequence variants was tested using the DESeq2 package for R [56], with taxon counts modeled on genotype [i.e., wild type vs. knock out] + locus [i.e., *zx1 zx2 zx3* vs. *zx1 zx2 zx3 zx4*). The packages ggplot2 [57] and pheatmap [58] were used for visualizations.

2. Supplementary Figures 1 to 32



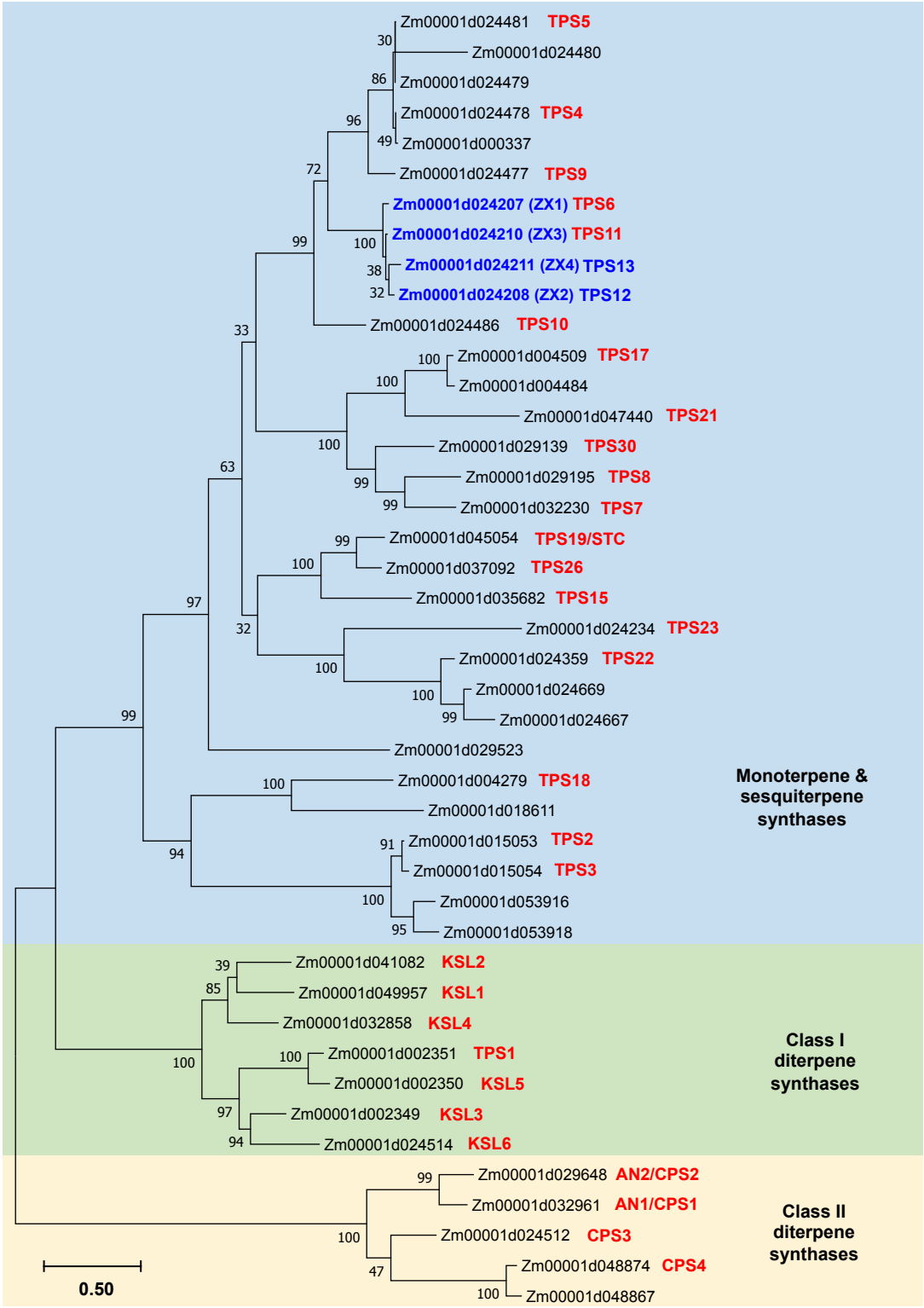
Supplementary Fig. 1 | Maize isoprenoid precursors and products in the current study relevant to zealexin pathway elucidation. a, Precursors for the ZX pathway, including farnesyl diphosphate (FPP) and products of *Zx* gene cluster I enzymes encoded by β-macrocarpene synthase genes *Zx1* to *Zx4* are (1) β-bisabolene and (2) β-macrocarpene. **b,** *Zx* gene cluster II enzymes encoded by *ZmCYP71Z19* (*Zx5*), *ZmCYP71Z18* (*Zx6*) and *ZmCYP71Z16* (*Zx7*) act on ZX1 to ZX4-derived products to produce (3) ZD2, (4) ZD1 and (5) ZA1. **c,** *Zx* gene cluster III enzymes encoded by *ZmCYP81A37* (*Zx8*), *ZmCYP81A38* (*Zx9*), *ZmCYP81A39* (*Zx10*) act on ZA1 to produce (6) ZC1, (7) ZB1, (8) ZA5, (9) ZA2, (10) ZA3, (11) ZC2 and (12) ZB3. Further modified *Zx* gene cluster III products include (13) ZA4, (14) ZA6, (15) ZA7, (16) ZA8 and (17) ZA9. **d,** Products of the β-selinene synthase ZmTPS21 include α and β-selinene with further catalysis by ZmCYP71Z19 (ZX5) yielding α and β-costic acid. **e,** Representative maize *ent*-kaurene related diterpenoids utilizing the pathway genes *ZmCYP71Z18* (*Zx6*) and *ZmCYP71Z16* (*Zx7*) include B-series kauralexins (KB; KB1, KB4, KB3, KB2) and A-series kauralexins (KA; KA1, KA4, KA3, KA2). **f,** Known dolabralexin precursor dolabradiene and pathway products derived from ZmCYP71Z18 (ZX6) and ZmCYP71Z16 (ZX7) include epoxydolabrene, epoxydolabranol, and trihydroxydolabrene. **Note:** Numbered compounds (1-17) in approximate order of occurrence represent those directly considered in the present study, unnumbered compounds represent intermediates or enzyme products that interact with the ZX biosynthetic pathway.

Supplementary Fig. 2



Supplementary Fig. 2 | Fungal-elicited ZX accumulation occurs in all species of the genus *Zea* examined. Average total zealexins in stems of 4-week old *Zea mays* spp. *mays* var. B73, *Zea mays* spp. *parviglumis* (Ames21889), *Zea diploperennis* (PI462368), *Zea luxurians* (PI422162), *Zea perennis* (Ames21874), and *Zea mays* spp. *mexicana* (Ames21851), treated with a heat-killed *F. venenatum* hyphae preparation. All stem tissues were harvested 5 days after treatment and analyzed by GC-MS. Total ZXs include ZA1, ZB1, ZD1 and ZD2. Error bars in indicate mean ± s.e.m. (n = 3 biologically independent replicates) and different letters (a–c) represent significant differences (one-way ANOVA $P < 0.05$; Tukey's test corrections for multiple comparisons, $P < 0.05$). NA (not available) represents missing samples/data due to low seed germination rates.

Supplementary Fig. 3



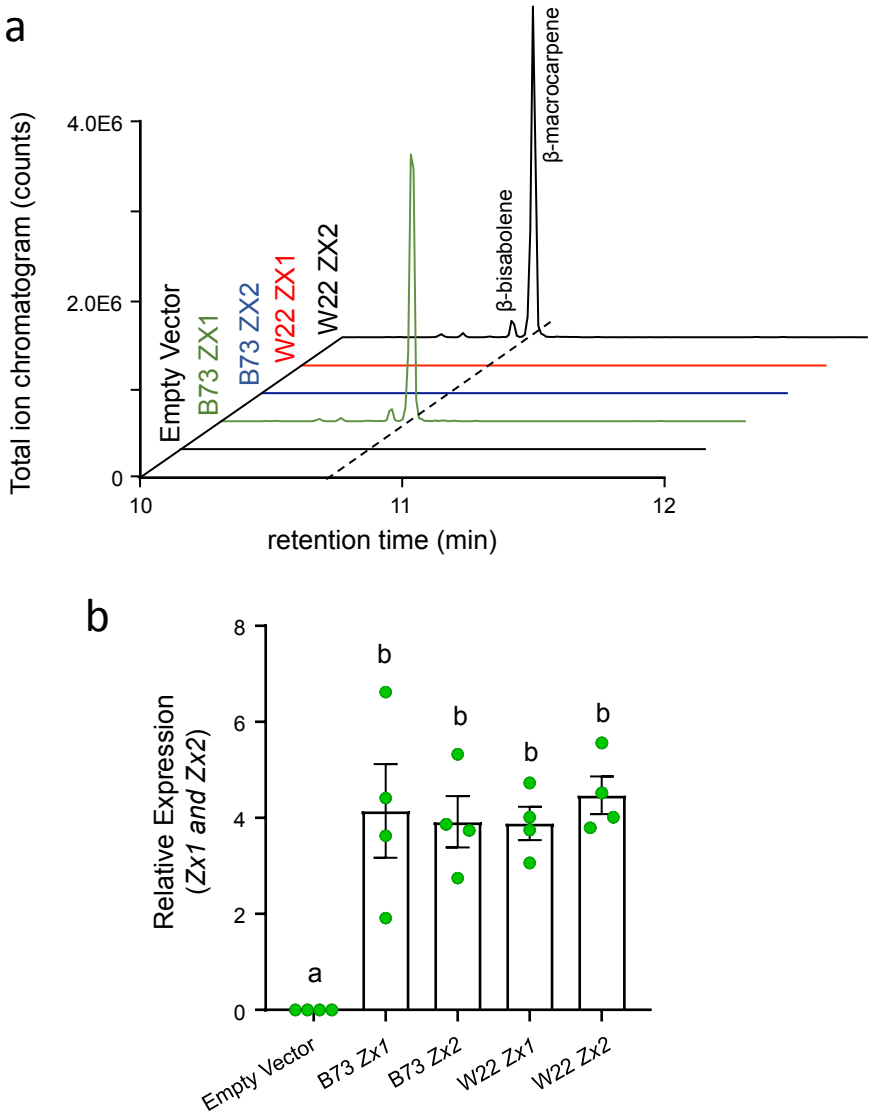
Supplementary Fig. 3 | Maximum likelihood phylogenetic tree of 43 maize (*Zea mays*) monoterpene, sesquiterpene and diterpene synthases (B73 RefGen_V4). Tree reconstruction was performed with the maximum likelihood algorithm using MEGA 7 program (www.megasoftware.net). Bootstrap values calculated from 1,000 iterations are indicated at the nodes. Protein sequences are listed in Supplementary Table 3

Supplementary Fig. 4

```
B73 ZX1 MAAPTLTADGPRLGQQEMKKMSP-SFHPTLWGDFFLSYEAPTEAQEAQMREKAAVLKEEVRNMIKGSHDVPEIVDLIITLQRLNLDYHYEDEINEKLTVV 99
B73 ZX2 MAAQTLTADGPRLGQQEMKKMSP-SFHPTLWGDFFLSYEAPTEAQEAEMRERAGVLREKVRSMIKGSHDVPEIVDLIITLQRLNLDYHYEDEINEKLTVV 99
B73 ZX3 MAAPTLTTDGPRLGQQEMKKMSP-SFHPTLWGDFFLSYEAPTEAQEAEMRQRAEVLREEVRNMIKGSHDVPEIVDLIITLQRLNLDYHYEDEINEKLAVV 99
B73 ZX4 MAALTQTTDGPRLGEQEMKKMSPASFHPTLWGDFFLSYEAPTETQEAQMREKAGVLKEEVRNMIKGSHDMPEIVDLIITLQRLNLDYHYEDEITEKLTVV 100
W22 ZX1 MAAPTLTADGPRLGQQEMKKMSP-SFHPTLWGDFFLSYEAPTEAQEAQMREKAGVLKEEVRNMIKGSHDVPEIVDLIITLQRLNLDYHYEDEINEKLTVV 99
W22 ZX2 MAAQTLTADGPRLGQQE-MKMSP-SFHPTLWGDFFLSYEAPTEAQEAEMRQRAEVLREEVRNMIKGSHDVPEIVDLIITLQRLNLDYHYEDEINEKLTVV 98
W22 ZX3 MAAPTLTTDGPRLGQQETKKMSP-SFHPTLWGDFFLSYEAPTEAQEAEMRQRAEVLREEVRNMIKGSHDVPEIVDLIITLQRLNLDYHYEDEINEKLTVV 99
W22 ZX4 MAALTQTTDGPRLGEQEMKKMSP-SFHPTLWGDFFLSYEAPTEAQEAQMREKAGVLKEEVRNMIKGSHDVPEIVDLIITLQRLNLDYHYEDEITEKLTVV 99

B73 ZX1 YKSNYDGGNLDLVSRRFYLLRKCGYDVSSDVFLKFKDQLGNFVEADTRSLLSLYNAAFLRIHGETVLDEAISFTMRVLQDRLEHLESPLAEEVSSALDTP 199
B73 ZX2 YNSNYDGGNLVLVSRRFYLLRKCGYHVSS---------------DTRTLLSLYNAAYLRIHGETMLDEAISFTTRCLQDRLEHLESPIAEEVSSALDTP 183
B73 ZX3 YNSNYDGGNLDLVSRRFYLLRKCGYHVSSDVFLNFKDQYGNFIEVDTRSLLSLYNAAYLRIHGETVLDEAISFTTRCLQDRLEHLESPIAEEVSSALDTP 199
B73 ZX4 YNSNYDGGNLDLVSRQFYLLRKCGYHVSSDVFLNFKDQYGNFIGADTRSLLSLYNAAYLRIHGETVLDEAISFTTRCLQDRLEHLESPIAEEVSSALDTP 200
W22 ZX1 YKSNYDGGNLDLVSRRFYLLRKCGYDVSIDVFLKFKDQLGNFVEADTRSLLSLYNAAFLRIHGETVLDEAISFTTRVLQDRLEHLESPLAEEVSSALDTP 199
W22 ZX2 YNSNYDGGNLDLVSRRFYLLRKCGYHVSSDVFLNFKDQYGNFIEADTRTLLSLYNAAYLRIHGETVLDEAISFTTRCLQDRLEHLESPIAEEVSSALDTP 198
W22 ZX3 YNSNYDGGNLDLVSRRFYLLRKCGYHVSSDVFLNFKDQYGNFIEADTRTLLSLYNAAYLRIHGETVLDEAISFTTRCLQDRLEHLESPIAEEVSSALDTP 199
W22 ZX4 YNSNYDGGNLDLVSRRFYLLRKCGYHVSSDVFLNFKDQYGNFIEADTRSLLSLYNAAYLRIHDETVLDEAISFTTRCLQDRLEHLESPIAEEVSSALDTP 199

                                                                                   274
B73 ZX1 LFRRVGTLEMKDYIPIYEKDAKQNKSILEFAKLNFNLLQLRYSSELKECTTWWKELRVESNLSFVRDRIVEVYFWMSGGCYDPQYSHSRIILTKIVAFIT 299
B73 ZX2 LFRRVGTLEMKDYIPIYEKDAKQNKSILEFAKLNFNLLQLLYSSELKECTAWWKELCVESNLSFVRDRIVEVYFWMSGGCYDPQYSHSRIILTKIVAFIT 283
B73 ZX3 LFRRVGTLEMKDYIPIYEKDAKQNKSILEFAKLNFNLLQLLYSSELKECTTWWKELRVESNLSFVRDRIVEVYFWMSGGCYDPQYSHSRIILTKIVAFIT 299
B73 ZX4 LFRRVGTLEIKDYIPIYEKDAKQNKSILEFAKLNFNLLQLLYSSELKECTSWWKELRVESNLSFVRDRIVEVYFWMSGGCYDPQYSHSRIILTKIVAFIT 300
W22 ZX1 LFRRVGTLEMKDYIPIYEKDAKQNKSILEFAKLNFNLLQLRYSSELKECTTWWKELRVESNLSFVRDRIVEVYFRMSGGCYDPQYSHSRIILTKIVAFIT 299
W22 ZX2 LFRRVGTLEMKDYIPIYEKDAKQNKSILEFAKLNFNLLQLLYSSELKECTAWWKELRVESNLSFVRDRIVEVYFWMSGGCYDPQYSHSRIILTKIVAFIT 298
W22 ZX3 LFRRVGTLEMKDYIPIYEKDAKQNKLILEFAKLNFNLLQLLYSSELKECTAWWKELRVESNLSFVRDRIVEVYFWMSGGCYDPQYSHSRIILTKIVAFIT 299
W22 ZX4 LFRRVGTLEMKDYIPIYEKDAKQNKSILEFAKLNFNLLQLLYSSELKECTAWWKELRVESNLSFVRDRIVEVYFWMSGGCYDPQYSHSRIILTKIVAFIT 299
                                                                          RXR

B73 ZX1 ILDDTLDSHATSCESMQLAEAIERWDESAVSLLPEYMKDFYMYLLKTFSSFENELGPDKSYRVFYLKEAVKELVREYTKEIKWRDEDYVPKTLKEHLKVS 399
B73 ZX2 ILDDTLDSLANSYESMQLAEAVERIHEGFL-------------------------------------HVKELVREYTKEIKWRDEDYVPKTLKEHLKVS 345
B73 ZX3 ILDDTLDSHANSYESMQLAEAVERWDESAVSLLPEYMKDFYMYLLKTFSSFENELGPDKSYRVFYLKEAVKELVREYTKEIKWRDEDYVPKTLKEHLKVS 399
B73 ZX4 ILDDTLDSHANSYESMQLAKAVERWDETAVSLLPEYMKDFYMYLLKTFSSFENELGPDKSYRVFYLKEAVKELVREYTKEIKWRDEDYVPKTLKEHLKVS 400
W22 ZX1 ILDDTLDSHATSCESMQLAEAIERWDESAVSLLPEYMKDFYMYLLKTFSSFENELGPDKSYRVFYLKEAVKELVREYTKEIKWRDEDYVPKTLKEHLKVS 399
W22 ZX2 ILDDTLDSHANSYESMQLAEAVERWDESAISLLPEYMKDFYMYLLKTFSSFENELGPDKSYRVFYLKEAVKELVREYTKEIKWRDEDYVPKTLKEHLKVS 398
W22 ZX3 ILDDTLDSHANSYESMQLAEAVERWDESAISLLPEYMKDFYMYLLKTFSSFENELGPDKSYRVFYLKEAVKELVREYTKEIKWRDEDYVPKTLEEHLKVS 399
W22 ZX4 ILDDTLDSHANSYESMQLAEAVERWDETAVSLLPEYMKDFYMYLLKTFSSFENELGPDKSYRVFYLKEAVKELVREYTKEIKWRDEDYVPKTLKEHLKVS 399
         DDXXD

B73 ZX1 LISIGGTLVLCSAFVGMGDVVTKKIMKWVMSDAELVKSFGIFVRLSNDIVSTKREQREKHCVSTVQCYMKQHEITMDEACEQIKELTEDSWKFMIEQGLA 499
B73 ZX2 LISIGGTLVLCSAFVGMGDVVTKKIMEWVMSDAELVKSFGIFVRLSNDIVSTKREQREKHCVSTVQCYMKQHELTMDEACEQIKELTEDSWKFMIEQGLA 445
B73 ZX3 LISIGGTLVLCSAFVGMGDVVTKKIMEWVMSDAELVKSFGIFVRLSNDIVSTKREQREKHCVSTVQCYMKQHELTMDEACEQIKELTEDSWKFMIEQGLA 499
B73 ZX4 LISIGGTLVLCSAFVGMGDVVTKKIMEWVMSDAELVKSFGIFVQLSNDIVSTKREQREKHCVSTVQCYMKQHELTMDEACEQIKELIEDSWKFMIEQGLA 500
W22 ZX1 LISIGGTLVLCSAFVGMGDVVTKKIMKWVMSDAELVKSFGIFVRLSNDIVSTKREQREKHCVSTVQCYMKQHELTMDEACEQIKELTEDSWKFMIEQGLA 499
W22 ZX2 LISIGGTLVLCSAFVGMGDVVTKKIMEWVMSDAELVKSFGIFVRLSNDIVSTKREQREKHCVSTVQCYMKQHELTMDEACELIKELTEDSWKFMIEQGLA 498
W22 ZX3 LISIGGTLVLCSAFVGMGDVVTKKIMEWVMSDAELVKSFGIFVRLSNDIVSTKREQREKHCVSTVQCYMKQHELTMDEACEQIKELTEDSWKFMIEQGLA 499
W22 ZX4 LISIGGTLVLCSAFVGMGDVVTKKIMEWVMSDAELVKSFGIFVRLSNDIVSTKREQREKHCVSTVQCYMKQHELTMDEACEQIKELTEDSWKFMIEQGLA 499
                                                 (N,D)DXX(S,T)XXXE

B73 ZX1 LKEYPIIVPRTVLEFARTVDYMYKEADKYTVSHTIKDMLTSLYVKPVLM 548
B73 ZX2 LKEYPIIVPRTVLEFARTVDYMYKEADKYTVSHTIKDMLTSLYVKPVLM 494
B73 ZX3 LKEYPIIVPRTVLEFARTVDYMYKEADKYTVSHTIKDMLTSLYVKPVLM 548
B73 ZX4 LKEYPIIVPRTVLEFARTVDYMYKEADKYTVSHTIKDMLTSLYVKPVLM 549
W22 ZX1 LKEYPIIVPRTVLEFARTVDYMYKEADKYTVSHTIKDMLTSLYVKPVLM 548
W22 ZX2 LKEYPIIVPRTVLEFARTVDYMYKEADKYTVSHTIKDMLTSLYVKPVLM 547
W22 ZX3 LKEYPIIVPRTVLEFARTVDYMYKEADKYTVSHTIKDMLTSLYVKPVLM 548
W22 ZX4 LKEYSIIVPRTVLEFARTVDYMYKEADKYTVSHTIKDMLTSLYVKPVLM 548
```

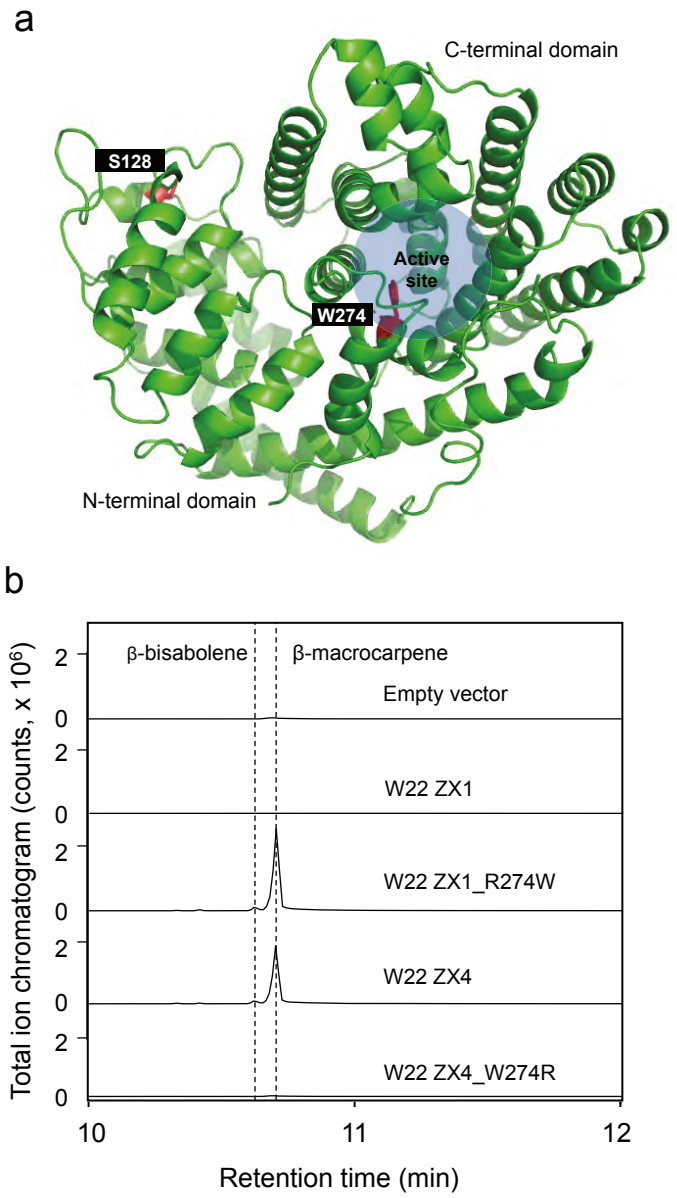
Supplementary Fig. 4 | Encoded amino acid sequence comparison of four B73 and four W22 β-macrocarpene synthases (ZX1 to ZX4) present in *Zx* gene cluster I. The alignment of predicted amino acid sequences encoded by B73 V4 genes *Zm00001d024207* (*Zx1*), *Zm00001d024208* (*Zx2*) *Zm00001d024210* (*Zx3*) and *Zm00001d024211* (*Zx4*) and by the W22 genes *Zm00004b038503* (*Zx1*), *Zm00004b038504* (*Zx2*), *Zm00004b038505* (*Zx3*), and *Zm00004b038506* (*Zx4*), was constructed with the program MEGA7 (www.megasoftware.net) and the MUSCLE (codon) algorithm. The visualization was done with the program BIOEDIT (http://www.mbio.ncsu.edu/BioEdit). Conserved amino acid motifs are indicated as RxR, DDxxD and (N,D)DXX(S,T)XXXE. Unlike W22 ZX2 to ZX4, W22 ZX1 contains a R instead of a W at position 274 as marked with a red dashed square. **Note:** B73 ZX2 contains 2 deletion sites that together are predicted to negatively impact activity.

Supplementary Fig. 5



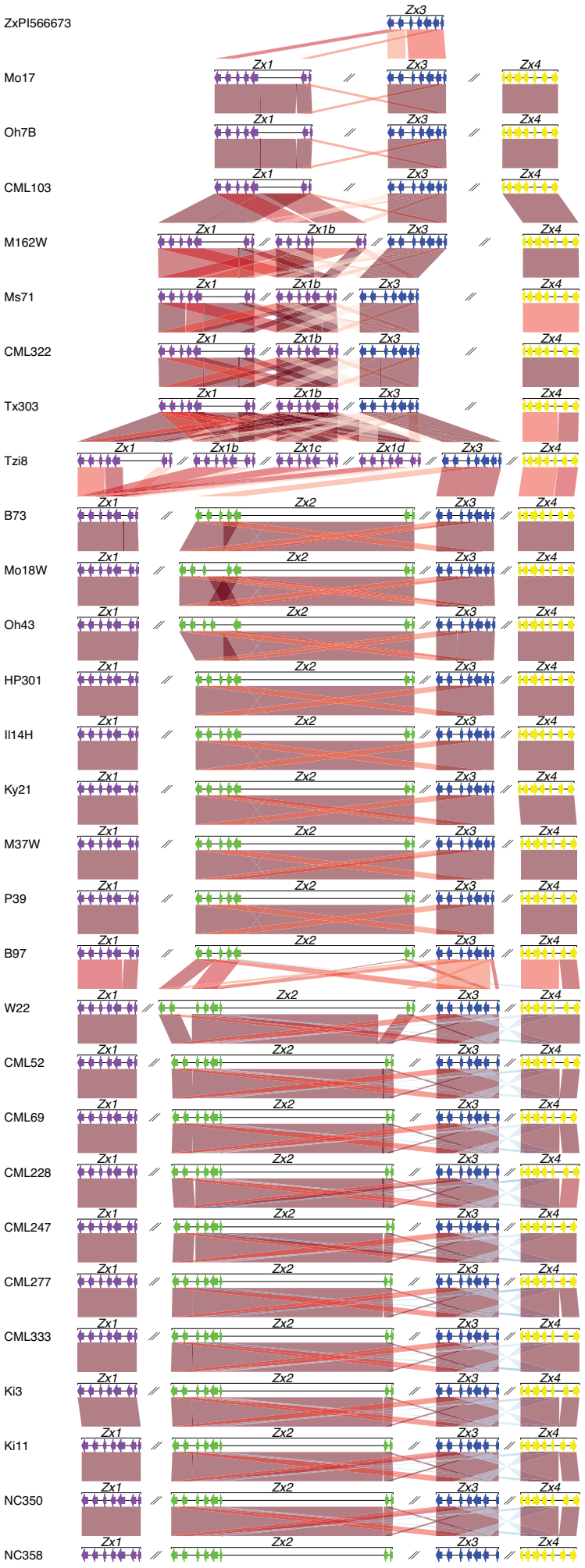
Supplementary Fig. 5 | Sequence variations in B73 *Zx2* and W22 *Zx1* are readily expressed in *N. benthamiana* yet fail to produce functional β-macrocarpene synthases. **a**, GC-MS total ion chromatograms are shown for leaf volatiles emitted following *Agrobacterium*-mediated transient *N. benthamiana* expression assays of B73 and W22 encoded β-macrocarpene synthases ZX1 and ZX2. An empty vector was used for the *Agrobacterium*-infiltrated control. Four biological repeats were performed and showed similar results. **b,** qrtPCR showing *Zx1 or Zx2* expression from transient assays in **a**. Error bars indicate mean ± s.e.m. (n = 4 biologically independent replicates). Within plots, different letters (a–b) represent significant differences (ANOVA $P < 0.05$; Tukey's test corrections for multiple comparisons, $P < 0.05$).

Supplementary Fig. 6

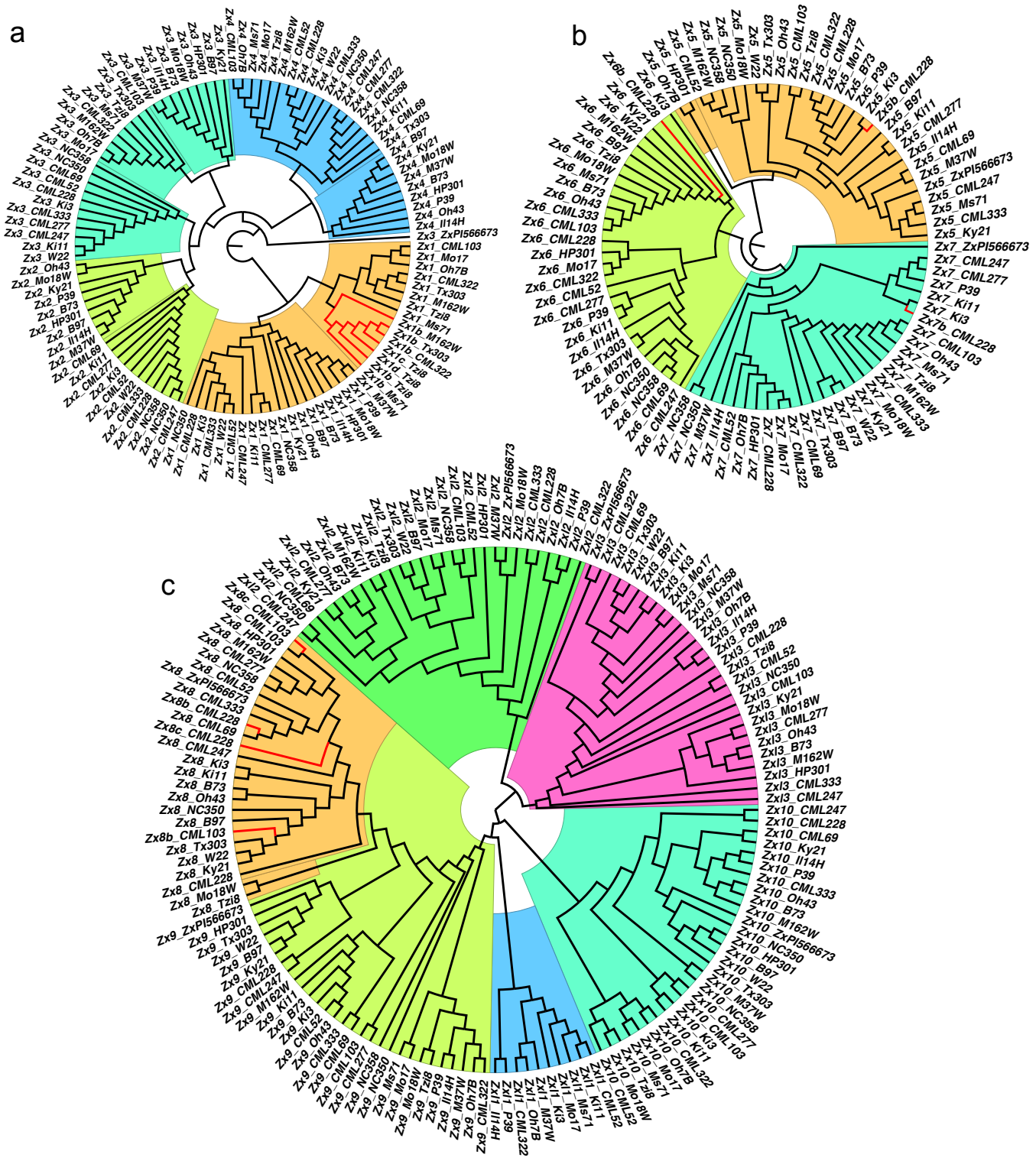


Supplementary Fig. 6 | A single nucleotide polymorphism (SNP) commonly present in maize germplasm results in a loss of function *zx1* mutation. a, Homology modeling of B73 β-macrocarpene synthase ZX1 based on the template 5eat.1.A (5-*epi*-aristolochene synthase from *N. tabacum*). **b**, GC-MS total ion chromatograms are shown for the volatiles emitted from *Agrobacterium*-mediated transient expression in *N. benthamiana* leaves of the naturally occurring W22 ZX1 (Zm00004b038503) mutant, wild type W22 ZX4 (Zm00004b038506), site directed repair (R274W) of W22 ZX1 restores catalytic activity while site directed mutagenesis (W274R) of W22 ZX4 destroys catalytic activity. An empty pLife33 vector was used for the negative control. Four biological repeats were performed and showed similar results. Of the analyzed maize inbreds, 35 of 237 harbor the null mutation in zx1, namely SNP_10_56448050 (A to G) (Supplementary Table 5).

Supplementary Fig. 7



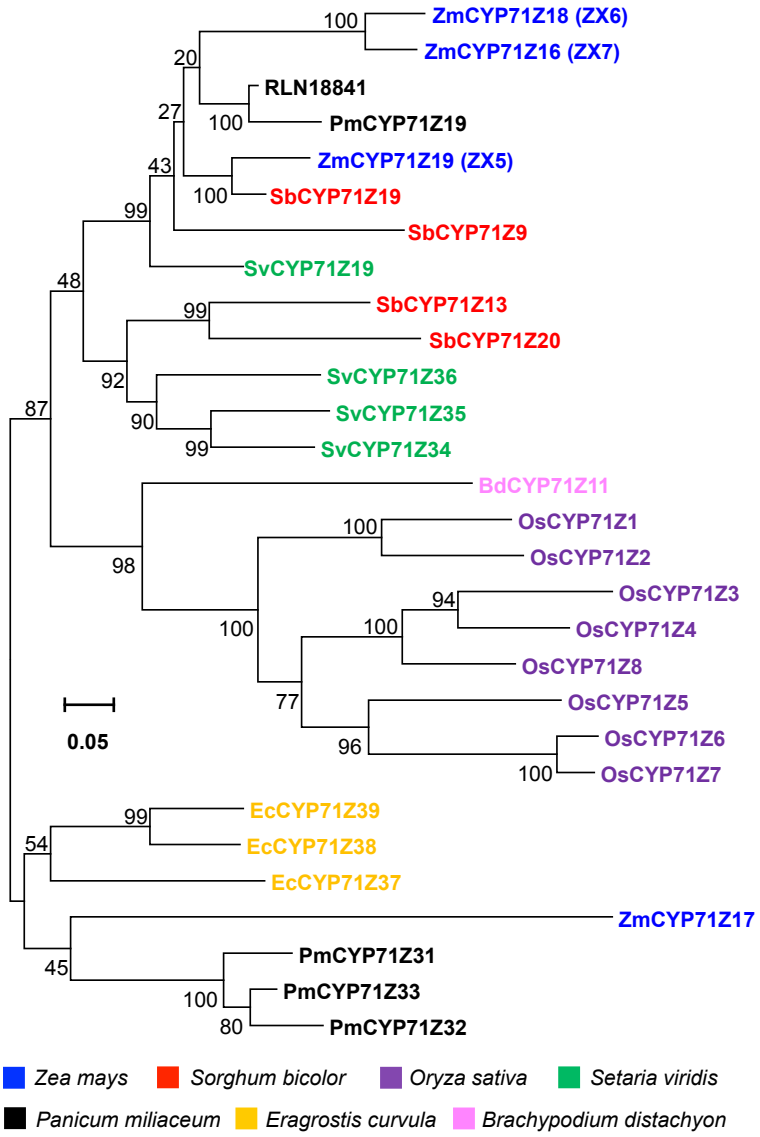
Supplementary Fig. 7 | Maize inbred lines commonly contain 4 *TPS* gene copies at *Zx* gene cluster I but can vary between 3 to 6 copies. Synteny plots of *TPS* genes in zealexin gene cluster I were examined in the reference genomes of PI566673, B73, W22, Mo17(CAU) and all other Nested Association Mapping (NAM) parents. B73 exon sequences of *Zx1* were used as a reference to query | to generate putative *Zx* gene models (Supplementary Table 6). The genomic sequences of each cluster were extracted and compared using the standalone Blastn software and the comparison results were visualized as synteny plots using genoPlotR. Less common variants in copy number and synteny were tentatively assigned as *Zx1b*, *Zx1c*, *Zx1d* to denote sequence relationships to B73 *Zx1* and comparative dissimilarity to B73 *Zx2*.



Supplementary Fig. 8 | Phylogenic analyses of blast search results for *ZmTPS*, *ZmCYP71Z* and *ZmCYP81* genes identified in *Zx* gene cluster I, II and III from highland teosinte (Zx-PI566673) and 28 maize inbred lines. Individual B73 exon sequences of *Zx1*, *Zx5* and *Zx8* were used as a reference to query the MaizeGDB BLAST database and generate a list of putative exon sequences for *Zx* gene clusters I, II and III, respectively. Adjacent and consecutive sequences were combined and used to generate putative *Zx* gene models in each cluster (Supplementary Table 6), the nucleotide sequences within each cluster were aligned with MAFFT and Maximum Likelihood (ML) phylogenetic trees were constructed using IQ-TREE with default parameters and 1,000 bootstrap replications.

Nguyen, Lam-Tung, Heiko A. Schmidt, Arndt von Haeseler, and Bui Quang Minh. (2015) IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular Biology and Evolution 32:268–74.

Supplementary Fig. 9



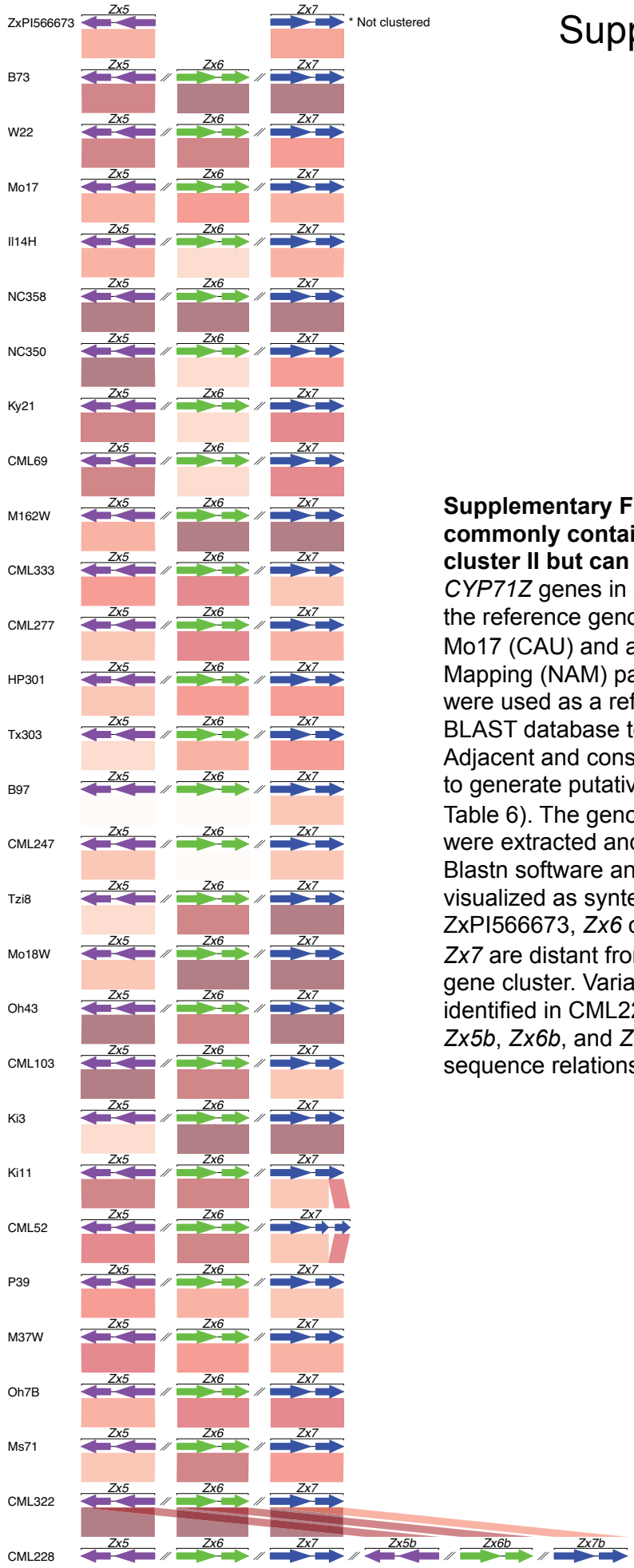
Supplementary Fig. 9 | Phylogenetic tree of the CYP71Z subfamily. Cytochrome P450s from grass species *Zea mays (Zm), Sorghum bicolor (Sb), Oryza sativa (Os), Setaria viridis (Sv), Panicum miliaceum (Pm), Eragrostis curvula (Ec),* and *Brachypodium distachyon (Bd)* with amino acid sequence identity >55% were included (ZmCYP71Z19 as a seed query). Tree reconstruction was performed with the maximum likelihood algorithm using MEGA 7 program (www.megasoftware.net). Bootstrap values calculated from 1000 iterations are indicated at the nodes. The protein accession numbers are listed in the Supplementary Table 3.

Supplementary Fig. 10

```
CYP71Z19(ZX5)  MEQKVLVAVGVAVLLVVVLSKLKSVLVTKPKLNLPPGPWTLPLIGSTHHLVTSPSIYRAMRDLAQKYGPLMMLRLGEVPT 80
CYP71Z18(ZX6)  MEDKVLIAVG-TVAVVAVLSKLKS-AVTKPKLNLPPGPWTLPLIGSIHHIVSNPLPYRAMRELAHKHGPLMMLWLGEVPT 78
CYP71Z16(ZX7)  MEDKVLLAVA-MVALIAVLSKLKSLLETKPKLNLPPGPWTLPLIGSIHHLVSSPLPYRAMRELAHKHGPLMMLWLGEVPT 79

CYP71Z19(ZX5)  LVVSSPEAAQAITKTHDIAFADRHMNTTIGVLTFNGTDLVFGPYGERWRQLRKICVLELFSVARVQSFQRIREEEVARFM 160
CYP71Z18(ZX6)  LVVSSPEAAQAITKTHDVSFADRHINSTVDILTFNGMDMVFGSYGEQWRQLRKLSVLELLSAARVQSFQRIREEEVARFM 158
CYP71Z16(ZX7)  LVVSSPEAAQAITKTHDVTFADRHMNSTVDILTFNGNDIVFGTYGEQWRQLRKLSVLELLSVARVQSFQRIREEEVARFM 159

CYP71Z19(ZX5)  QSLAASAG---TVNLSKMISRFINDTFVRECIGSRCKYQDEYLDAFDTAVRQTSVLTVADLFPSSRLMQAVGTAPRNALK 237
CYP71Z18(ZX6)  RSLAASASAGATVDLSKMISSFINDTFVRESIGSRCKYQDEYLAALDTAIRVAAELSVGNIFPSSRVLQSLSTARRKAIA 238
CYP71Z16(ZX7)  RNLAASAGAGATVDLSKMISSFINDTFVRESIGSRCKHQDEYLDALHTGIRVAAELSVANLFPSSRLLQSLSTARRKAVA 239

CYP71Z19(ZX5)  CRNRITRILEQIIREKVEAMGRGEKTAHEGLIGVLLRLQKEANLPTLLTNDTIVALMFDLFGAGSDTSSTTLNWCITELI 317
CYP71Z18(ZX6)  SRDEMARILGQIIRETKESMDQGDKTSNESMISVLLRLQKDAGLPIELTDNVVMALMFDLFGAGSDTSSTTLTWCMTELV 318
CYP71Z16(ZX7)  ARDEMARILGQIIRETKEAMDWGDKASNESMISVLLRLQKEAGLPIELTDDIVMALMFDLFGAGSDTSSTTLTWCMTEMI 319

CYP71Z19(ZX5)  RHPAAMAKAQAEVREAFKGKARIISEDDLAGAGLSYLKLVIKEALRMHCPLPLLLPRLCRETCQVMGYDIPKGTAVFINV 397
CYP71Z18(ZX6)  RYPATMAKAQAEVREAFKGKT-TITEDDLSTANLRYLKLVVKEALRLHCPVPLLLPRKCREACQVMGYDIPKGTCVFVNV 397
CYP71Z16(ZX7)  RYPATMAKAQAEVREAFKGKT-TITEDDLSRANLSYLKLVVKEALRLHCPVPLLIPRKCRETCQIMGYDIPKDTCVLVNV 398

CYP71Z19(ZX5)  WAVCRDAKYWEDPEEFRPERFEDTNLEYNYKGTNYEFLPFGSGRRMCPGANLGLGNIELALASLLYHYDWKLPDGVKPQD 477
CYP71Z18(ZX6)  WAICRDPRYWEDAEEFKPERFENSNLDY--KGTYYEYLPFGSGRRMCPGANLGVANLELALASLLYHFDWKLPSGQEPKD 475
CYP71Z16(ZX7)  WAICRDSRYWEDADEFKPERFENSSLDY--KGTSHEYLPFGSGRRMCPGGNLGVANMELALASLLYHFDWKLPSGQEPKD 476

CYP71Z19(ZX5)  VQVWEGPGLIAKKKTGLLLRPVTCIAFACSSG 509
CYP71Z18(ZX6)  VDVWEAAGLVAKKNIGLVLHPVSHIAPVNA-- 505
CYP71Z16(ZX7)  VDVWEAAGLVGRKNAGLVLHPVSRFAPVNA-- 506
```

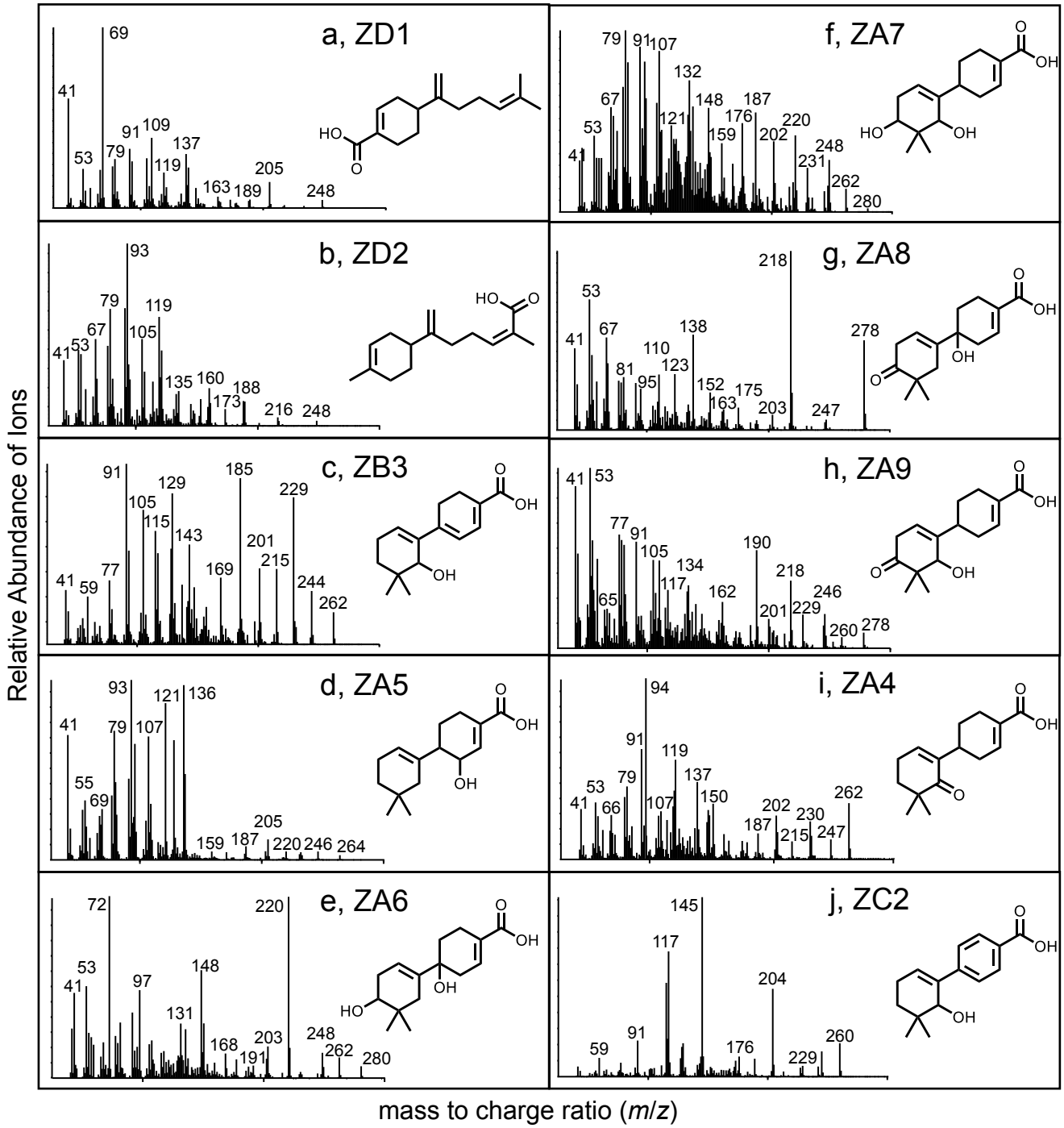
Supplementary Fig. 10 | Encoded amino acid sequence comparison of three B73 CYP71Z subfamily P450s present in *Zx* gene cluster II. The alignment of predicted amino acid sequences encoded by the B73 *ZmCYP71Z19* (*Zx5, Zm00001d014121*), *ZmCYP71Z18* (*Zx6*, *Zm00001d014134*), and *ZmCYP71Z16* (*Zx7, Zm00001d014136*) with the program MEGA7 (www.megasoftware.net) and the MUSCLE (codon) algorithm. The visualization was performed with the program BIOEDIT (http://www.mbio.ncsu.edu/BioEdit). B73 ZX5 shares 71.5-72.1% protein sequence identity to both ZX6 and ZX7, while ZX6 and ZX7 have 89.3% protein sequence identity to each other. The protein accession numbers are listed in the Supplementary Table 3.

Supplementary Fig. 11



Supplementary Fig. 11 | Maize inbred lines commonly contain 3 *CYP71Z* genes at *Zx* gene cluster II but can also contain 6. Synteny plots of *CYP71Z* genes in *Zx* gene cluster II were examined in the reference genomes of ZxPI566673, B73, W22, Mo17 (CAU) and all other Nested Association Mapping (NAM) parents. B73 exon sequences of *Zx5* were used as a reference to query the MaizeGDB BLAST database to obtain putative exon sequences. Adjacent and consecutive sequences were combined to generate putative *Zx* gene models (Supplementary Table 6). The genomic sequences of each cluster were extracted and compared using the standalone Blastn software and the comparison results were visualized as synteny plots using genoPlotR. In ZxPI566673, *Zx6* could not be found whereas *Zx5* and *Zx7* are distant from each other and do not form a gene cluster. Variations in gene copy number identified in CML228 were tentatively assigned as *Zx5b*, *Zx6b*, and *Zx7b* to denote duplication and sequence relationships to B73 *Zx5*, *Zx6* and *Zx7*.

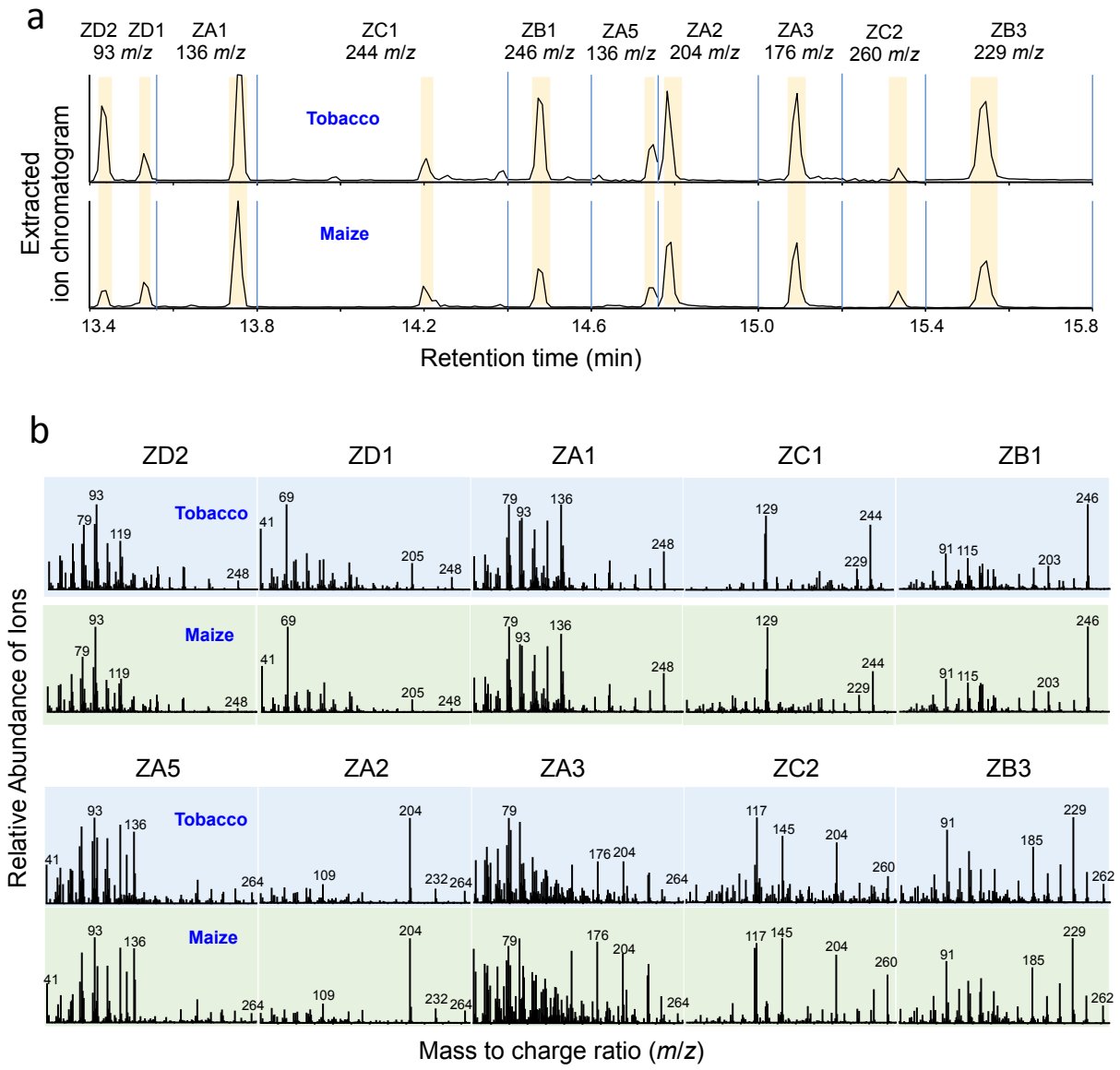
Supplementary Fig. 12



Supplementary Fig. 12 | Reference electron ionization (EI) spectra of zealexins identified in the current study following methyl ester derivatization. Maize metabolites purified, identified by NMR, derivatized with trimethylsilyldiazomethane and analyzed by electron ionization (70 eV). Two β-bisabolene derivatives include: **a,** ZD1 [4-(6-methylhepta-1,5-dien-2-yl)cyclohex-1-ene-1-carboxylic acid] and **b**, ZD2 [(E)-2-methyl-6-(4-methylcyclohex-3-en-1-yl)hepta-2,6-dienoic acid]. Six β-macrocarpene derivatives include **c**, ZB3 [6'-hydroxy-5', 5'-dimethyl-[1,1'-bi(cyclohexane)]-1,1',3-triene-4-carboxylic acid]; **d**, ZA5 [2-hydroxy-5',5'-dimethyl-[1,1'-bi(cyclohexane)]-1',3-diene-4-carboxylic acid]; **e**, ZA6 [1,4'-dihydroxy-5',5'-dimethyl-[1,1'-bi(cyclohexane)]-1',3-diene-4-carboxylic acid]; **f**, ZA7 [4',6'-dihydroxy-5',5'-dimethyl-[1,1'-bi(cyclohexane)]-1',3-diene-4-carboxylic acid]; **g**, ZA8 [1-hydroxy-5',5'-dimethyl-4'-oxo-[1,1'-bi(cyclohexane)]-1',3-diene-4-carboxylic acid]; **h**, ZA9 [6'-hydroxy-5',5'-dimethyl-4'-oxo-[1,1'-bi(cyclohexane)]-1',3-diene-4-carboxylic acid]. **i**, ZA4 and **j**, ZC2 follow from Christensen *et al*. (2018)[1] and Suzuki *et al.* (2007)[2], respectively. Reference spectra for β-macrocarpene, ZA1, ZA2, ZA3 and ZB1 follow from the original description of ZmTPS6/11[3] and zealexins[4]. Combined spectra were used to identify enzyme products of *Agrobacterium*-mediated transient *N. benthamiana* co-expression assays.

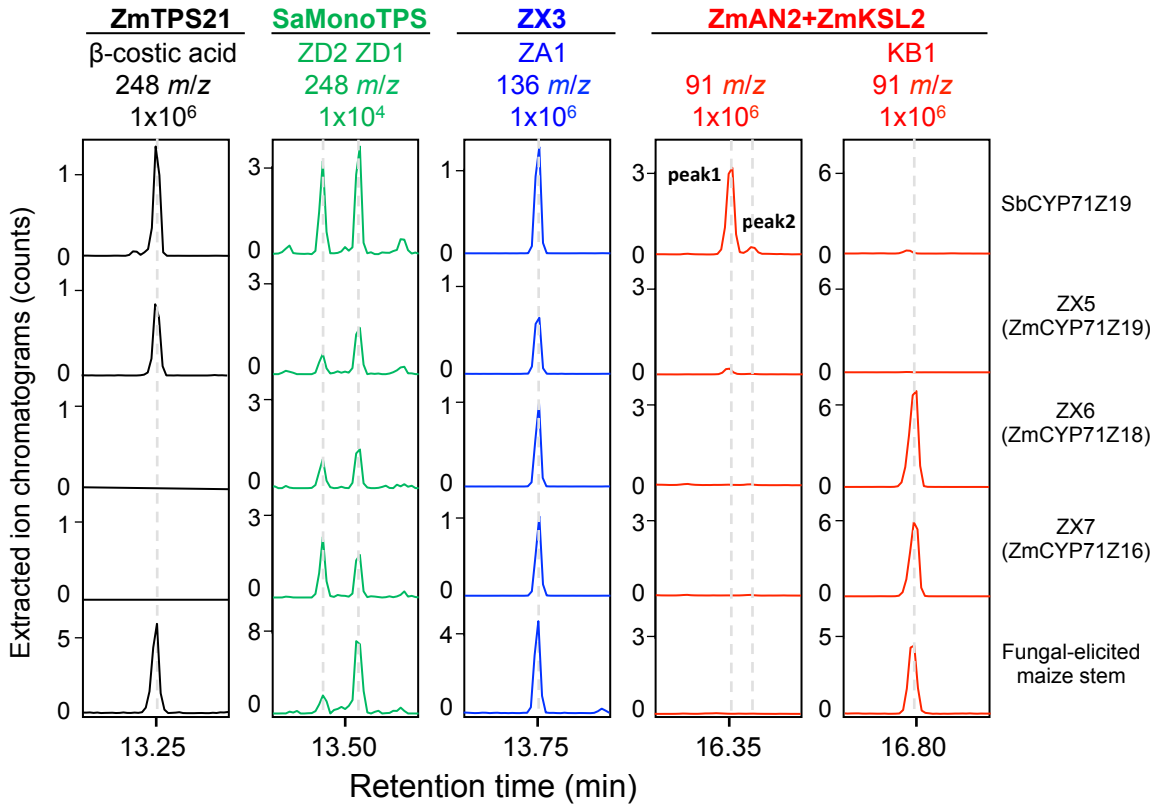
References: [1]Christensen *et al*. (2018) *Planta*. 247(4):863-873; [2]Suzuki R, Iijima M, Okada Y, Okuyama T (2007) *Chem Pharm Bull* (Tokyo) 55: 153–155; [3]Kollner *et al*. (2008). *J Biol Chem* 283: 20779–20788; [4]Huffaker *et al*. (2011) *Plant Physiol*. 156(4):2082-2097

Supplementary Fig. 13



Supplementary Fig. 13 | Representative GC-MS retention times, EI diagnostic (*m/z*) ions, and full EI spectra of ZXs present in plant extracts from transient expression assays in *N. benthamiana* and fungal-elicited maize stems. a, Representative GC-MS extracted ion chromatograms showing diagnostic EI (*m/z*) ions and relative retention times of ZX methyl ester derivatives detected in extracts *Agrobacterium*-mediated transient *N. benthamiana* (tobacco) co-expression assays and fungal-elicited maize stems. **b**, EI mass spectra of ZXs from (**a**) as methyl esters. Mass spectra highlighted with blue are from transient expression assays in *N. benthamiana* (tobacco) and those with green are from elicited maize stems. Reference metabolite extracted ion chromatograms and EI spectra were derived from the following enzyme pairs expressed in *N. benthamiana*: ZD1 and ZD2 (SaMonoTPS + ZX6); ZA1(ZX3 + ZX6); ZB1 and ZC1 (ZX3 + ZX6 + ZX8); ZA5, ZA2, ZA3, ZC2, and ZB3 (ZX3 + ZX6 + ZX8 + ZX10).

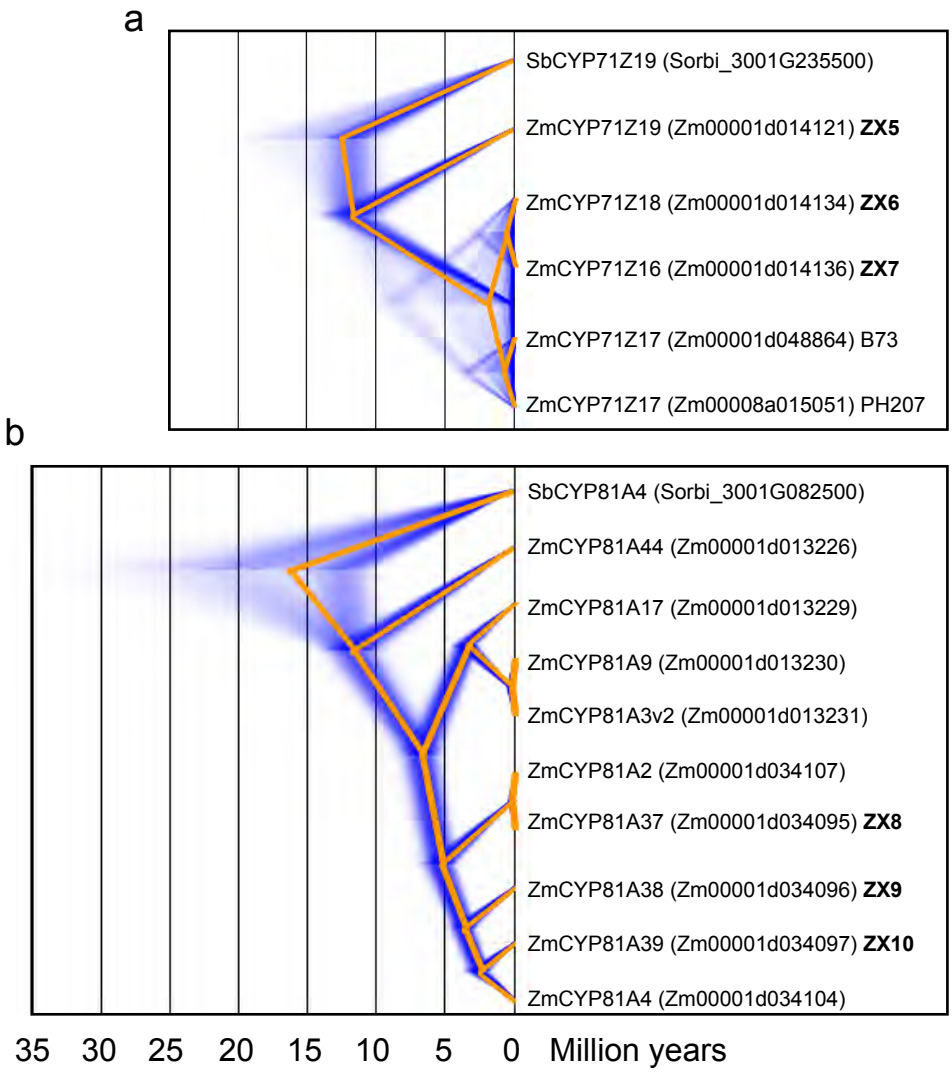
Supplementary Fig. 14



Supplementary Fig. 14 | The sorghum homolog of ZmCYP71Z19 (ZX5) oxidizes diverse terpene olefins. GC-MS extracted ion chromatograms of preparations derived from *Agrobacterium*-mediated transient *N. benthamiana* co-expression assays of Mo17 TPS21, SaMonoTPS, ZX3 or ZmAN2 + ZmKSL2 with SbCYP71Z19, ZmCYP71Z19, ZmCYP71Z16 or ZmCYP71Z18 showing production of β-costic acid (*m*/*z* 248), ZD2 (*m*/*z* 248), ZD1 (*m*/*z* 248), ZA1 (*m*/*z* 136), *ent*-kaur-15-17-ol (*m*/*z* 91, peak 1)[1], *ent*-kaur-15-17-al (*m*/*z* 91, peak 2)[2], and KB1 (*m*/*z* 91) as methyl ester derivatives. Four independent experiments were preformed and showed similar results. References: [1]Zhang *et al.* Chemical constituents of *Aristolochia constricta*: antispasmodic effects of its constituents in guinea-pig ileum and isolation of a diterpeno-lignan hybrid. *J Nat Prod*. 2008, 71:1167-72. [2]Su WC, Fang JM, Cheng YS. Abietanes and kauranes from leaves of *Cryptomeria japonica*. *Phytochemistry*. 1994, 35:1279–1284.

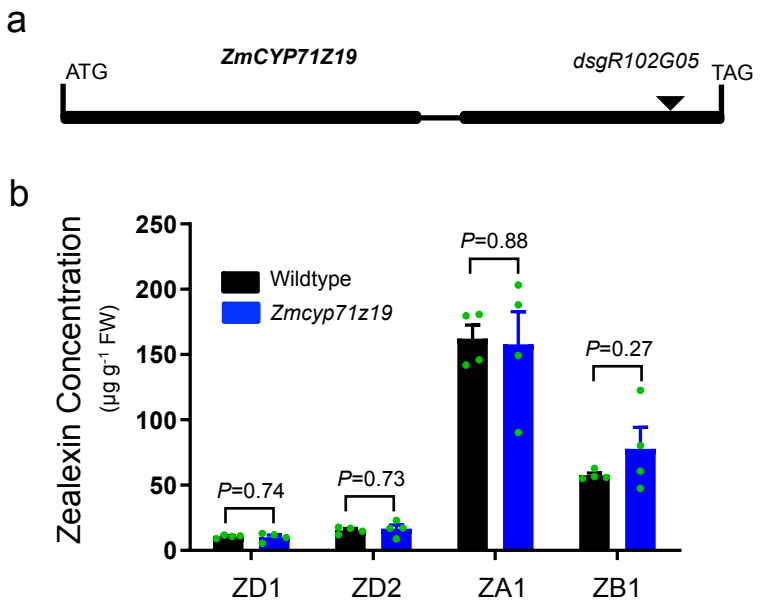
.

Supplementary Fig. 15



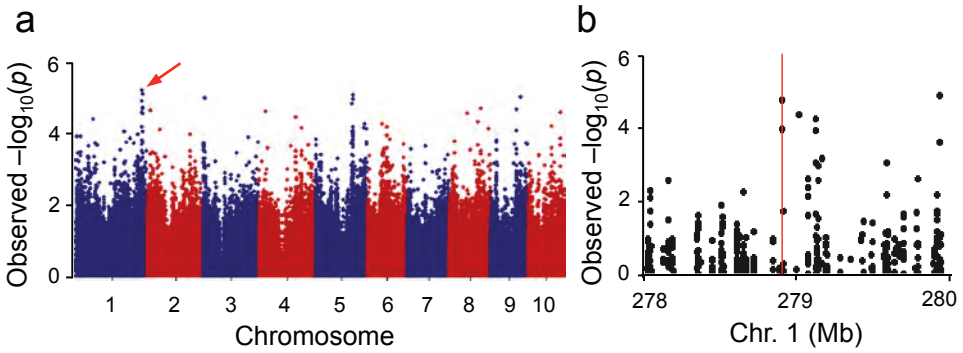
Supplementary Fig. 15 | DensiTree visualizations of the relationships among the maize *CYP71Z* gene subfamily (A) and the *CYP81A* gene subfamily (B) in the B73 RefGen_V4 reference genome. The *ZmCYP71Z* gene subfamily appears to have arisen from a recent duplication of the *ZmCYP71Z19* (*Zx5*: *Zm00001d014121*) gene. The *ZmCYP81A* gene subfamily appears to have arisen from a more ancient duplication of the *ZmCYP81A4* (*Zm00001d034104*) gene. Orange lines show the consensus tree and blue shading shows uncertainty in the age estimates of nodes in the tree.

Supplementary Fig. 16



Supplementary Fig. 16 | Characterization of the W22 *zx5* (*Zmcyp71z19*) *dsg*-transposon mutant suggests an endogenous role of ZX6 and ZX7 in zealexin pathway redundancy. **a**, The *Dsg* is inserted in exon 2 of the W22 *Zmcyp71z19* gene. **b**, Average quantity of ZD1, ZD2, ZA1 and ZB1 from W22 wild type siblings and *Zmcyp71z19* plants treated with heat-killed *F. venenatum* hyphae for 3 days. Error bars indicate mean ± s.e.m. (n = 4 biologically independent replicates). P values represent Student's t test, two-tailed distribution, equal variance.

Supplementary Fig. 17



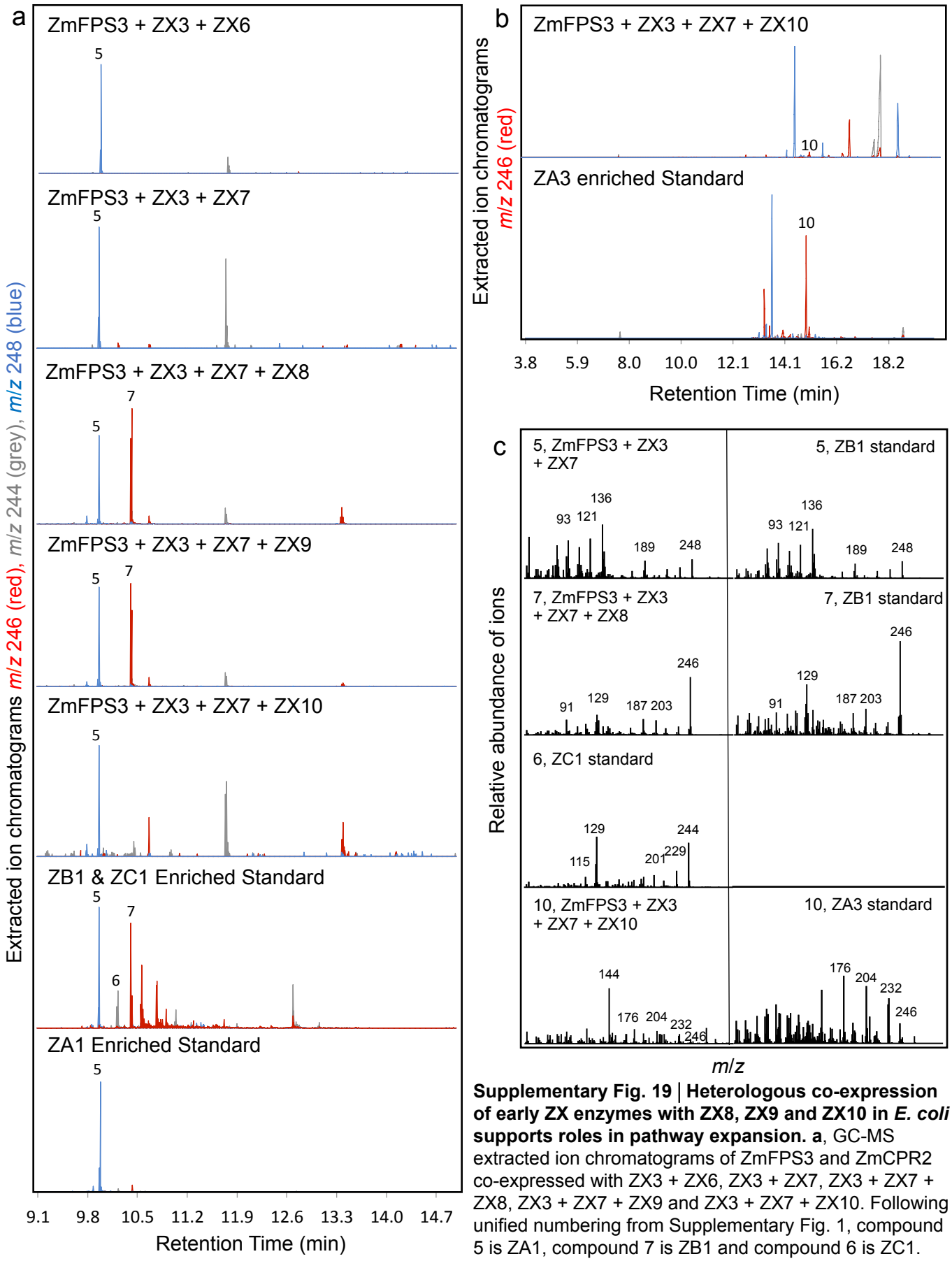
Supplementary Fig. 17 | An association analysis using the Goodman diversity panel provides additional support for *Zx* gene cluster III. a, Manhattan plot of the Goodman diversity panel association analysis using ratio of ZB1 to ZA1 as a mapping trait. 35 days after planting 258 Goodman diversity lines were wounded, stem elicited with heat-killed *F. venenatum* hyphae, harvested after 3 days and analyzed by LC-MS. Negative log10-transformed *P* values from the compressed mixed linear model are plotted on the y axis. 246,477 SNP markers identified (red line arrow) the largest log10-transformed *P* located on Chr. 1 (B73 RefGen_v2). **b,** Local Manhattan plot surrounding the peak on Chr. 1.

Supplementary Fig. 18

```
                                                   Proline-rich region
B73 CYP81A37_ZX8     METANVVILSFFFIFAIHRLFNRLHSRSTKMKPLPPGPLAIPVLGHLYLLLEKNMHHLFT 60
Mo17 CYP81A37_ZX8    METAYVVILSFFFTFAIHRLLNRRHSRNMNTKQLPPGPLAIPVLGHLH-LLEKKIHDSFT 59
B73 CYP81A38_ZX9     METANVVILS-FFIFAIHRLFNRLHSRSTKMKPLPPGPLAIPVLGHLYLLLEKNIHHMFT 59
Mo17 CYP81A38_ZX9    METANVTILSFFFIFAIHRLLNRLHNRSTKMKQLPPGPLAIPVLGHLYLLLEKNIHHLFT 60
B73 CYP81A39_ZX10    METAYVVVLSFFFIVAIHRLLNRGESKNMSKQPLPPGPRAIPVLGHLH-LLEKPYHLAFM 59
Mo17 CYP81A39_ZX10   METAFVVVLSFLFIVAIHRLLNRGESKNMSKQPLPPGPRAIPVLGHLH-LLEKPYHLAFM 59

B73 CYP81A37_ZX8     RLAARYGPVLYLRLGSRNAVVVSSVDCARECFTEHDVTFANRPTFPTLHLMTYGGTTVGT 120
Mo17 CYP81A37_ZX8    RLAARYGPVLYVRLGSRDAVVVSSVDCARECFTEHDVTFANRPSFPTLHLMTYGGTTVGT 119
B73 CYP81A38_ZX9     RLAARYGPVLYLRLGSRNAVIVSSVDCARECFTEHDVTFANRPTFPTLHLMTYGGTTVGT 119
Mo17 CYP81A38_ZX9    RLAARYGPVLYLRLGSRNAVVVSSVDCARECFTEHDVTFANRPTFPTLHLMTYGGTTVGT 120
B73 CYP81A39_ZX10    RLAARYGPVFSLQLGSRAAVVVSTADCARECLAEHDATFANRPTYPTLHLMTYGGATIGH 119
Mo17 CYP81A39_ZX10   RLAARYGPVFSLQLGSRAAVVVSTADCARECLAEHDATFANRPTYPTLHLMTYGGATIGH 119

B73 CYP81A37_ZX8     CAYGPYWRHIRRVITVHLLSALRVR-SMVPAIEAEVRAMARSMYRAAAAAPCGTGAAKVE 179
Mo17 CYP81A37_ZX8    CAYGPYWRHIRRVITVHLLSAQRVR-AMVPAIEAEVRAMARSMYRAAVTAPGGAGAAKVE 178
B73 CYP81A38_ZX9     CAYGPYWRHIRRVITVHLLSALRVR-SMVPAIEAEVRAMARSMYRAAAAAPCGTGAAKVE 178
Mo17 CYP81A38_ZX9    CAYGPYWRHIRRVITVHLLSALRVR-SMVPAIEAEVRAMARSMYRAAAAAPCGTGAAKVE 179
B73 CYP81A39_ZX10    SAYGAHWRHVRRVVSLHLLSPQRVLSSMVPTITAEVRAMARRMYRAAAG-----GAARIE 174
Mo17 CYP81A39_ZX10   SAYGAHWRHVRRVVSLHLLSPQRVLSSMVPTITAEVRAMARRMYRAAAG-----GAARIE 174

B73 CYP81A37_ZX8     LRGRLFEVALSALMETVAQSKTSRSSMGGAADTGMSPEAQEFKESMDVMIPLLGTANMWD 239
Mo17 CYP81A37_ZX8    LRGMLFELALSALMETVAKTKTSRT--VDAADTGLSPEAREFKESMDVMVPLLGTANMWD 236
B73 CYP81A38_ZX9     LRGRLFEVALSALMETVAQSKTSRSSMGGAADTGMSPEAQEFKESMDVMVPLLGTANMWD 238
Mo17 CYP81A38_ZX9    LRGRLFEVALSALMETVAQSKTSRSSMGGAADTGMSPEAQEFKESMDVMVPLLGTANMWD 239
B73 CYP81A39_ZX10    LKGRLLELSLSTLMETIAQTKTYRP--ADDADTDMSPETQEFKQLLDAIISLLGLTNVLE 232
Mo17 CYP81A39_ZX10   LKGRLLELSLSTLMETIAQTKTYRP--ADDADTDMSPETQEFKQLLDAIISLLGLTNVLE 232

B73 CYP81A37_ZX8     FLPVLQRFDVFGVKNKMAAAVSGRDAFFRRLIDEERRRLDDGVESENKSMMAVLLTLQKS 299
Mo17 CYP81A37_ZX8    FLPVLQRFDVFGVKNKIAAAVSSRDAFFRRLIDAERRRLDGGVKSEDKSMMAVLLTLQKS 296
B73 CYP81A38_ZX9     FLPVLQRFDVFGVKNKMAAAVSGRDAFFRRLIDEERRRLDDGVESENKSMMAVLLTLQKS 298
Mo17 CYP81A38_ZX9    FLPVLQRFDVFGVKNKMAAAVSGRDAFFRRLIDEERRRLDDGVQSENKSMMAVLLTLQKS 299
B73 CYP81A39_ZX10    FLPVLQTFDVLGLKKKIAAAAGRRDAFFQRLIDAERRRLDDGVEGQKKSMLAVLLALQKS 292
Mo17 CYP81A39_ZX10   FLPVLQTFDVLGLKKKIAAAAGRRDAFFQRLIDAERRRLDDGVEGQKKSMLAVLLALQKS 292

B73 CYP81A37_ZX8     EPENYSDDMIMSLCFSMFSAGTETTATTAEWAMSLLLNNPEVLKKAQGEIDAHVGNSRLL 359
Mo17 CYP81A37_ZX8    EPENYNDDMIMSLCFSMFSAGTETTATTAEWAMSLLLNHPEVLKKAQGQIDAYVGNSRLL 356
B73 CYP81A38_ZX9     EPENYSDDMIMSLCFSMFSAGTETTATTAEWAMSLLLNNPEVLKKAQGEIDAHVGNSRLL 358
Mo17 CYP81A38_ZX9    EPENYSDDMIMSLCFSMFSAGTETTATTAEWAMSLLLNNPEVLKKAQGEIDAHVGNSRLL 359
B73 CYP81A39_ZX10    EPENYTDEKIMALCFSMFIAGTDSTATAMEWAMSLLLNHPEVLKKAREEIDAHVGSSRLL 352
Mo17 CYP81A39_ZX10   EPENYTDEKIMALCFSMFIAGTDSTATAMEWAMSLLLNHPEVLKKAREEIDAHVGSSRLL 352
                                O2-binding region (A/G)GX(D/E)T(T/S)
B73 CYP81A37_ZX8     GADDMPHLPYLQCVLTETLRLYPVFPMLIAHESTADCKVGGHHVPSGTMLLTNAYAIHRD 419
Mo17 CYP81A37_ZX8    CADDMPHLPYLQCILTETLRLYPIIPMLIAHESTADCKVGGHHVPSGTMLLVNAYAIHRD 416
B73 CYP81A38_ZX9     GADDMPHLPYLQCVLTETLRLYPVFPMLIAHESTADCKVGGHHVPSGTMLLTNAYAIHRD 418
Mo17 CYP81A38_ZX9    GADDMPHLPYLQCVLTETLRLYPVFPMLIAHESTADCKVGGHHVPSGTMHLTNVYAIHRD 419
B73 CYP81A39_ZX10    GADDVPSLGYLHCVLNETLRLYPVGPTLIPHESTADCTVGGYRVPSGTMLLVNVYAIHRD 412
Mo17 CYP81A39_ZX10   GADDVPSLGYLHCVLNETLRLYPVGPTLIPHESTADCTVGGYRVPSGTMLLVNVYAIHRD 412

B73 CYP81A37_ZX8     PAAWTEPDAFRPERFEDG--SAEGKLLIPFGMGRRKCPGETMALRTLGLVLGTLIQCFDW 477
Mo17 CYP81A37_ZX8    PAAWMDSTVFRPERFEDGSASAEGRLLIPFGMGRRKCPGETMALRTLGLVLGTLIQCFDW 476
B73 CYP81A38_ZX9     PAAWTEPDAFRPERFEDG--SAEGKLLIPFGMGRRKCPGETMALRTLGLVLGTLIQCFDW 476
Mo17 CYP81A38_ZX9    PAAWTEPDAFRPERFEDG--SAEGKLLIPFGMGRRKCPGETMALRTLGLVLGTLIQCFDW 477
B73 CYP81A39_ZX10    PATWPDPDVFRPERFEDGGGSAEGRLLIPFGMGRRKCPGETMALQIMGLALGTMIQCFDW 472
Mo17 CYP81A39_ZX10   PATWPDPDVFRPERFEDGGGSAEGRLLIPFGMGRRKCPGETMALQTMGLALGTMIQCFDW 472
                                                Heme binding PFGXGRRXCXG
B73 CYP81A37_ZX8     ATVGG--VPKVDMTEASGLTLPRAVPLEAMCKPRQAMLDVLQNL 519
Mo17 CYP81A37_ZX8    ATVGG--VPKVDMTEASGLTLPRAVPLEAMCKPRQAMLDVLQNL 518
B73 CYP81A38_ZX9     ATVGG--VPKVDMTEASGLTLPRAVPLEAMCKPRQAMLDVLQKL 518
Mo17 CYP81A38_ZX9    ATVGG--VPKVDMTEASGLTLPRAVPLEAMCKPRQAMLDVLQKL 519
B73 CYP81A39_ZX10    GAVGGGGAPKVDMTQGGGLTLPRAVPLEAMCKPRQVMLDVLQKL 516
Mo17 CYP81A39_ZX10   GAVGGGEAPKVDMTQGGGLTLPRAVPLEAMCKPRQVMLDVLEKL 516
```

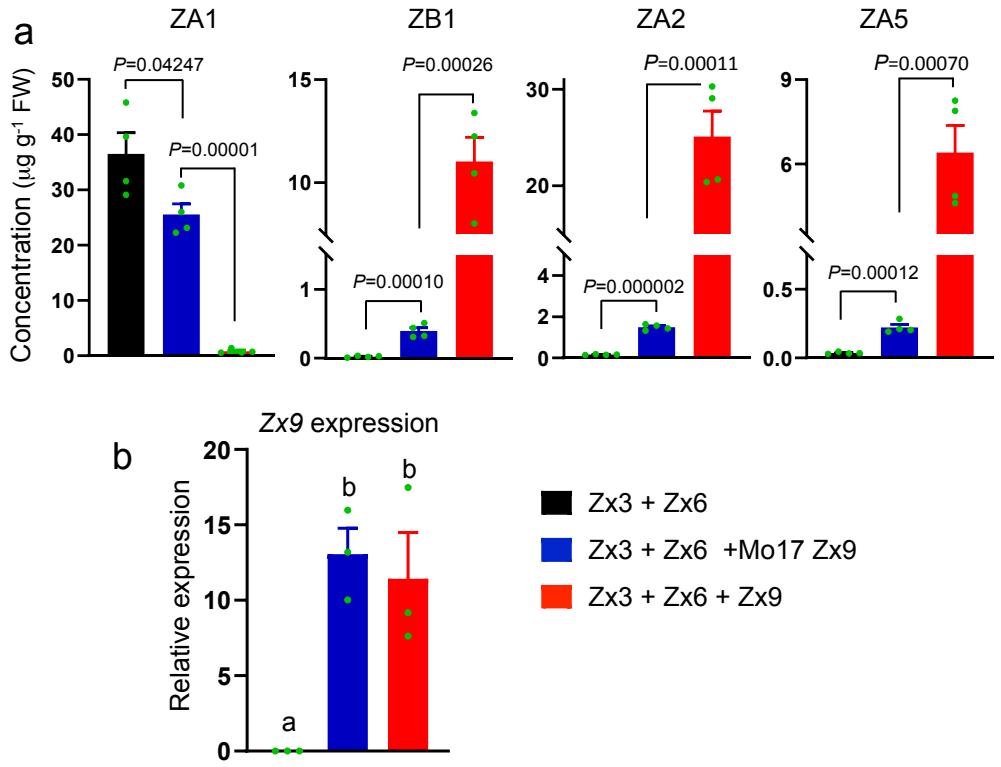
Supplementary Fig. 18 | Encoded amino acid sequence comparison of *Zx* gene cluster III CYP81A subfamily P450s in B73 and Mo17 inbred lines. The alignment was constructed based on predicted amino acid sequences of ZmCYP81A37, ZX8 (B73, Zm00001d034095; Mo17, deduced from genomic sequence no assigned ID), ZmCYP81A38, ZX9 (B73, Zm00001d034096; Mo17, Zm00014a026557) and ZmCYP81A39, ZX10 (B73, Zm00001d034097; Mo17, Zm00014a026556) with the program MEGA7 (www.megasoftware.net) and the MUSCLE (codon) algorithm. The visualization was done with the program BIOEDIT (http://www.mbio.ncsu.edu/BioEdit). B73 ZX8 shares 99% and 72% amino acid identity with B73 ZX9 and B73 ZX10, respectively.

Supplementary Fig. 19



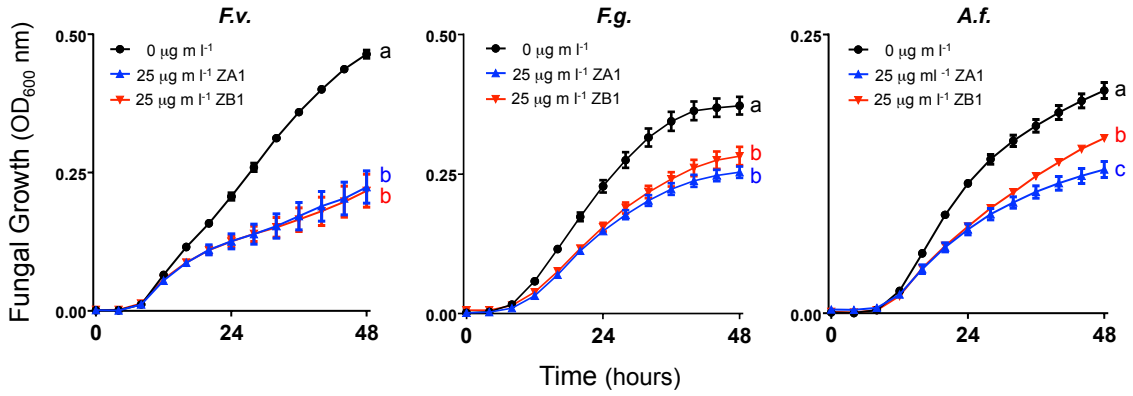
Supplementary Fig. 19 | Heterologous co-expression of early ZX enzymes with ZX8, ZX9 and ZX10 in *E. coli* supports roles in pathway expansion. a, GC-MS extracted ion chromatograms of ZmFPS3 and ZmCPR2 co-expressed with ZX3 + ZX6, ZX3 + ZX7, ZX3 + ZX7 + ZX8, ZX3 + ZX7 + ZX9 and ZX3 + ZX7 + ZX10. Following unified numbering from Supplementary Fig. 1, compound 5 is ZA1, compound 7 is ZB1 and compound 6 is ZC1. Enriched ZX standards are derived from root extracts and all were runs analyzed on a HP5-MS column. **b**, GC-MS extracted ion chromatograms of ZmFPS3 and ZmCPR2 co-expressed ZX3, ZX7, and ZX10 run on a separate GC column (DB-35) and MS instrument. The ZA3 enriched standard was a Mo17 root extract. Compound 10 is ZA3. **c**, All compounds in (**a**) and (**b**) were analyzed as methyl ester derivatives.

Supplementary Fig. 20



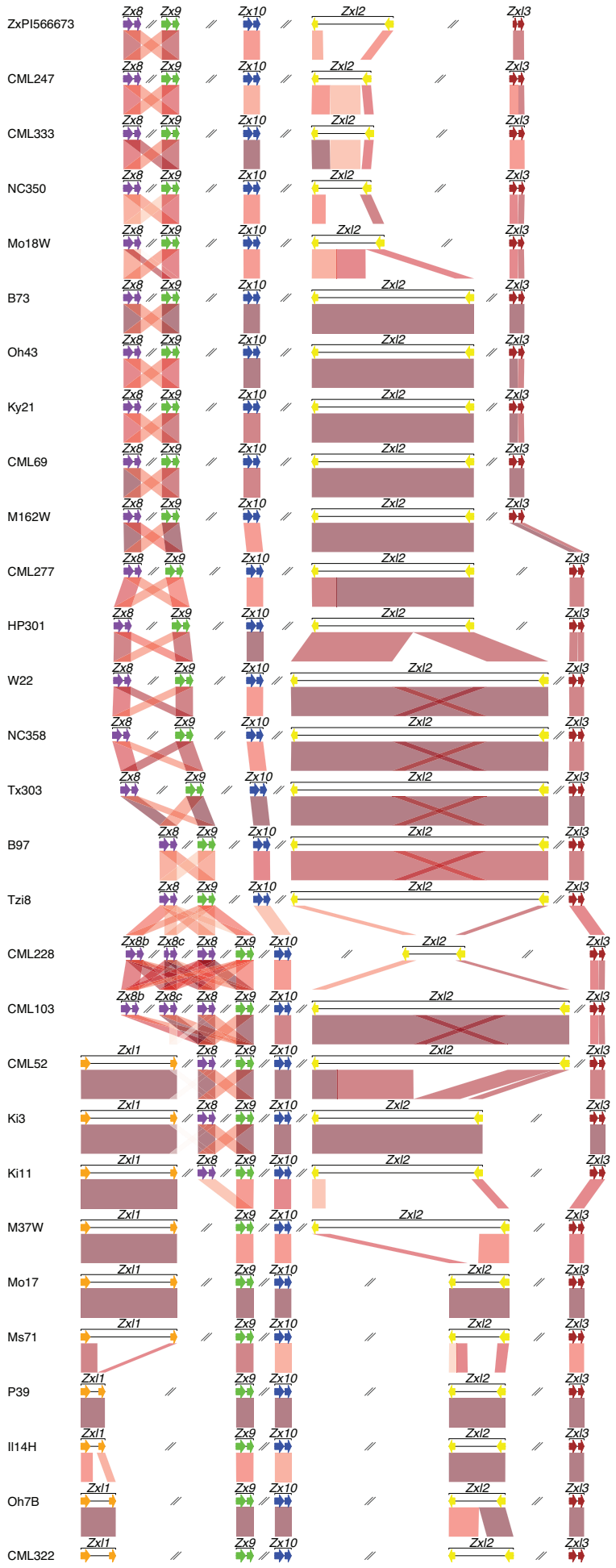
Supplementary Fig. 20 | Mo17 ZX9 (ZmCYP81A38) retains residual catalytic activity towards ZA1. a,Transient expression of ZX3 + ZX6 with ZX9, Mo17 ZX9, or empty vector in *N. benthamiana* demonstrates that Mo17 ZX9 has residual enzymatic activity in converting ZA1 to ZB1, ZA2 and ZA5. Zealexins extracted from 5-day *Agrobacterium*-inoculated leaves were measured with GC-MS. Error bars indicate mean ± s.e.m. (n = 4 biologically independent replicates). Two-tailed P-values from Student's t test (unpaired) are shown above bars to indicate significant differences. **b,** qrtPCR showing *ZX9* expression from transient assays of ZX3 + ZX6 with ZX9, Mo17 ZX9, or empty vector in *N. benthamiana.* Error bars indicate mean ± s.e.m. (n = 3 biologically independent replicates). Within plots, different letters (a–b) represent significant differences (ANOVA $P < 0.05$; Tukey's test corrections for multiple comparisons, $P < 0.05$).

Supplementary Fig. 21



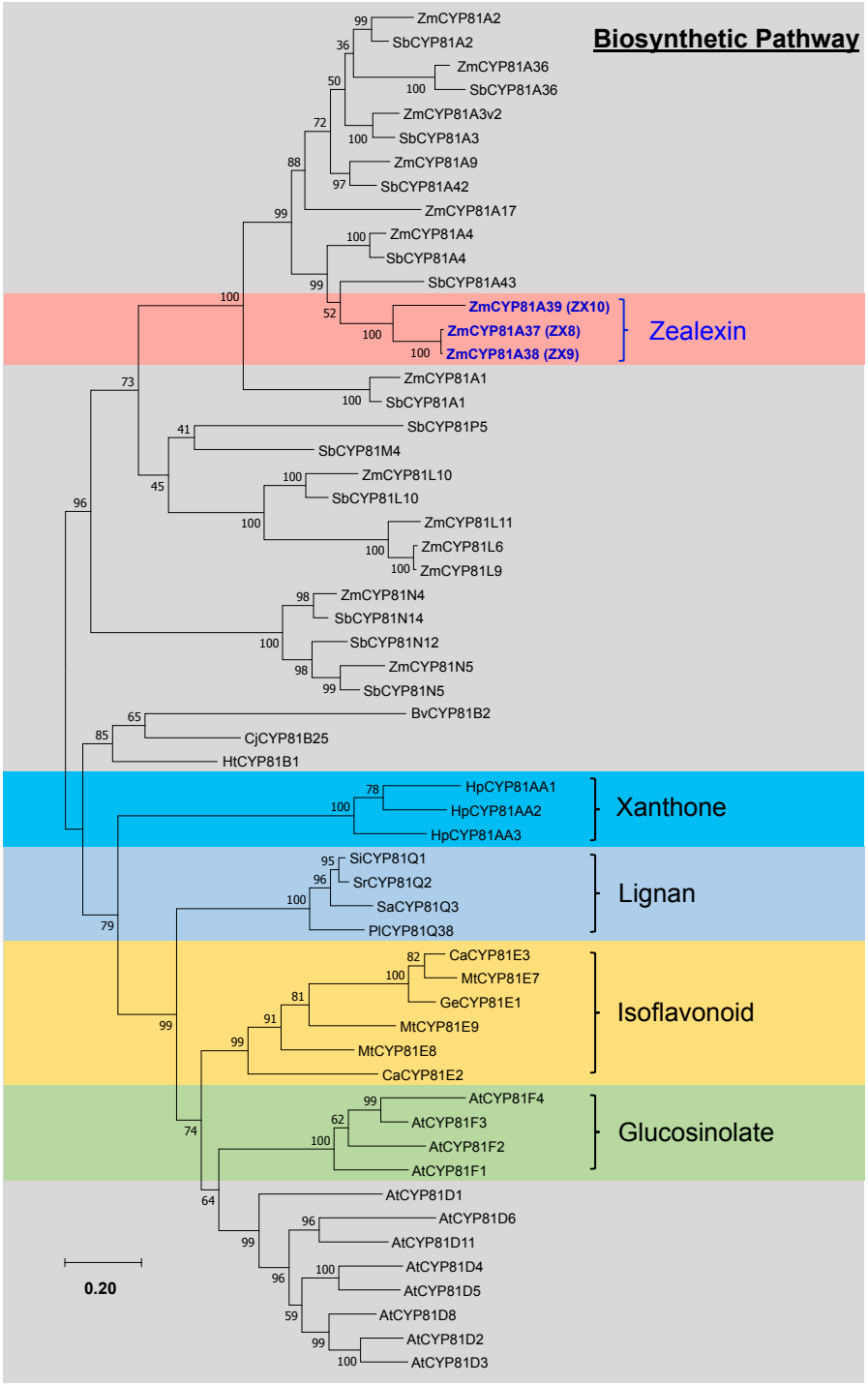
Supplementary Fig. 21 | *Zx* gene cluster III product ZB1 displays antibiotic activity against maize fungal pathogens. Fungal growth estimates measured at 600 nm (optical density) of *Fusarium verticillioides* (*F.v.*), *Fusarium graminearum* (*F.g.*) and *Aspergillus flavus* (*A.f.*) in liquid medium in the presence of a dimethylsulfoxide (DMSO) solvent control (0 μg ml^{-1}; black circles), ZA1 (25 μg ml^{-1}; blue triangles), or ZB1 (25 μg ml^{-1}; red arrows). Error bars indicate mean ± s.e.m. (n = 6 biologically independent replicates) and different letters (a–c) represent significant differences (one-way ANOVA $P < 0.05$; Tukey's test corrections for multiple comparisons, $P < 0.05$).

Supplementary Fig. 22



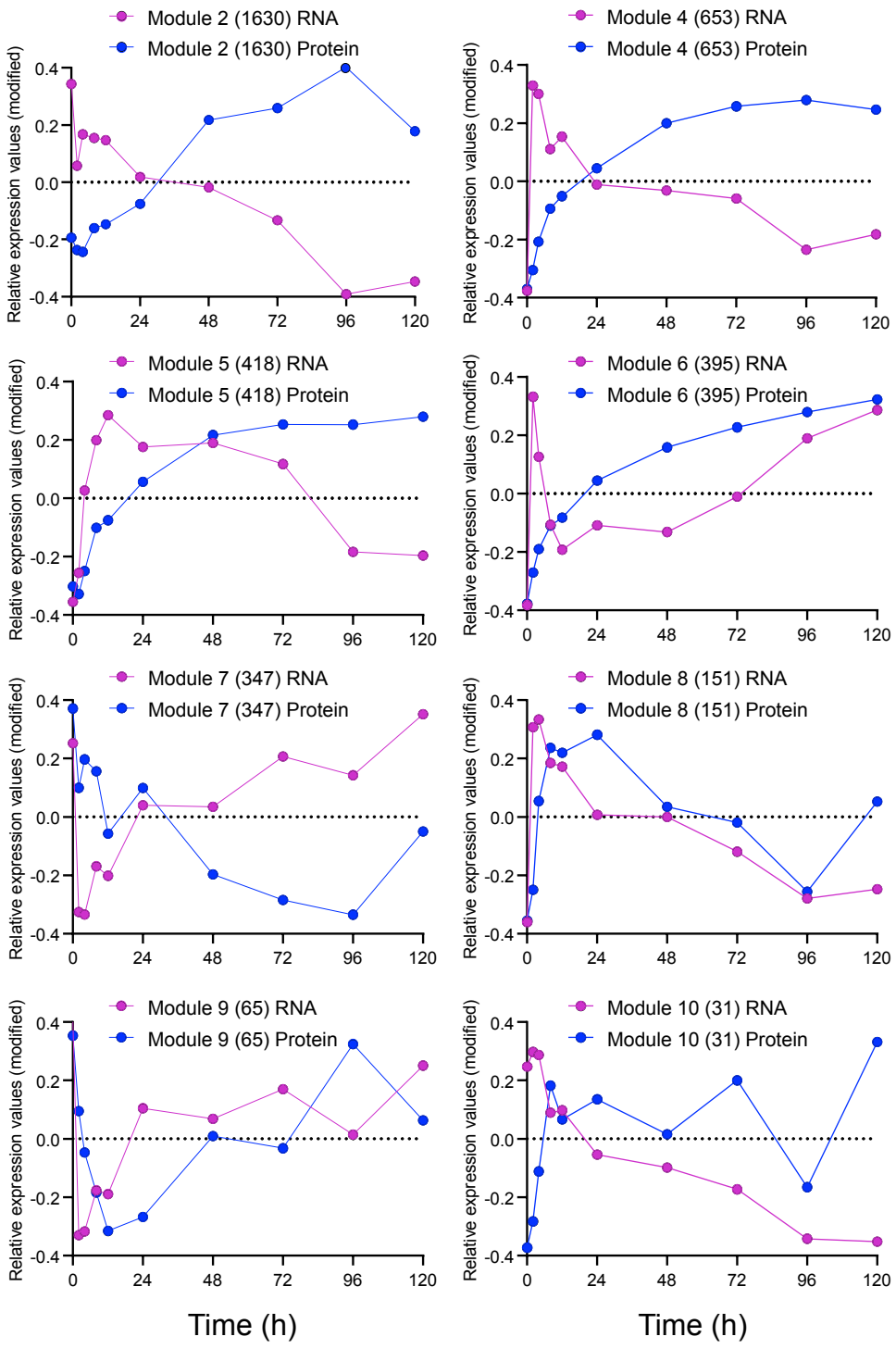
Supplementary Fig. 22 | *Zx* gene cluster III co-occurs within a larger variable cluster of *CYP81* genes and contains duplications of *Zx8* in select inbred lines. Synteny plots of *CYP81A* genes in *Zx* gene cluster III were examined in the reference genomes of Zx-PI566673, B73, W22, Mo17(CAU) and all other Nested Association Mapping (NAM) parents. B73 exon sequences of *Zx8* were used as a reference to query the MaizeGDB BLAST database to obtain putative exon sequences. Adjacent and consecutive sequences were combined to generate putative *Zx* gene models (Supplementary Table 6). The genomic sequences of each cluster were extracted and compared using the standalone Blastn software and the comparison results were visualized as synteny plots using genoPlotR. B73 *Zx8*, *Zx9* and *Zx10* occur as part of a larger cluster *CYP81A* genes that include *ZmCYP81A4* (*Zxl2*) and *ZmCYP81A2* (*Zxl3*) of unknown function. In 10 inbred lines, an additional related *ZmCYP81* sequence can be found nearby *Zx8* termed zealexin-like 1 (*Zxl1*). Duplicate copies of *Zx8* occur in CML228 and CML103 termed *Zx8b* and *Zx8c*.

Supplementary Fig. 23



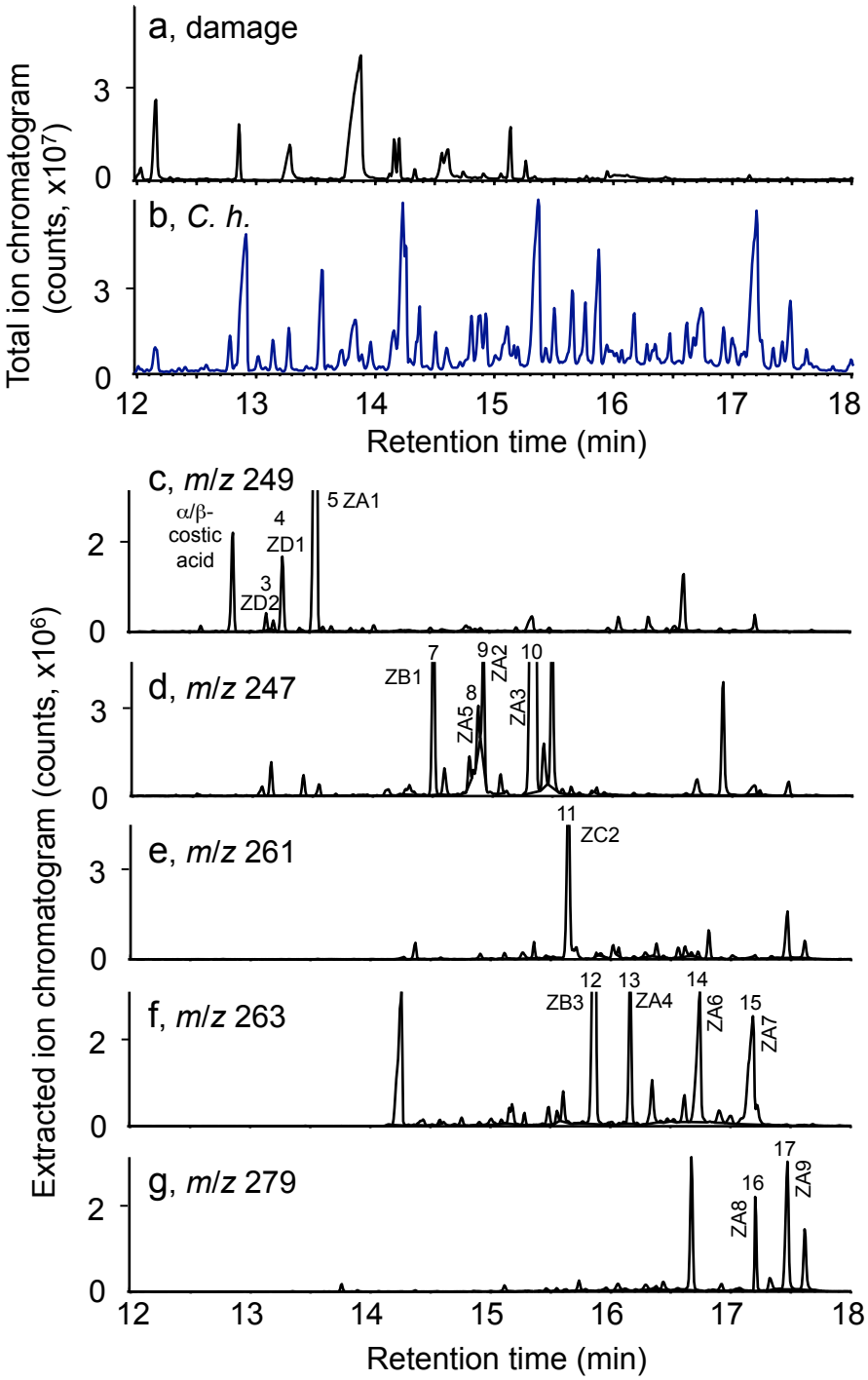
Supplementary Fig. 23 | *Zx* gene cluster III expands the known roles of CYP81 enzymes to include sesquiterpenoid defenses. Phylogenetic analysis of *Zea mays* (Zm) CYP81 subfamily of P450s together with representative members of the CYP81 subfamily based on amino acid sequences from *Arabidopsis thaliana* (At), *Beta vulgaris* (Bv), *Cicer arietinum* (Ca), *Coptis japonica* (Cj), *Glycyrrhiza echinate* (Ge), *Helianthus tuberosus* (Ht), *Hypericum perforatum* (Hp), *Medicago truncatula* (Mt), *Phryma leptostachya* (Pl), *Sesamum alatum* (Sa), *Sesamum indicum* (Si), *Sesamum radiatum* (Sr), and *Sorghum bicolor* (Sb). Tree reconstruction was performed with the maximum likelihood algorithm using MEGA 7 program. Bootstrap values calculated from 1000 iterations are indicated at the nodes. The protein accession numbers and references to specific roles in glucosinolate, isoflavonoid, lignin, and xanthone are listed in the Supplementary Table 3.

Supplementary Fig. 24



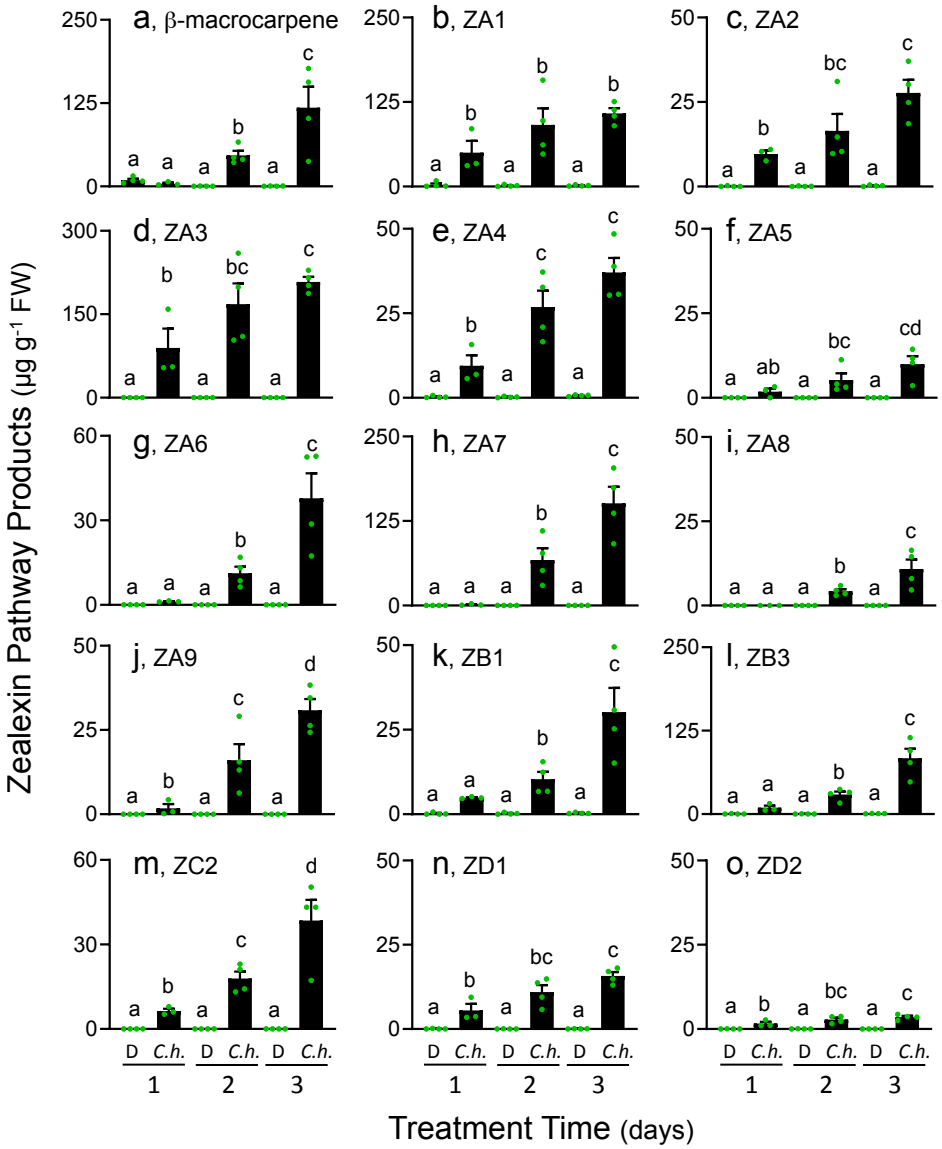
Supplementary Fig. 24 | Weighted Gene Co-Expression Network Analysis (WGCNA) of W22 transcriptome and proteome changes following *Fusarium* elicitation yields 10 module eigengenes dominated by co-suppression, co-activation and dysregulated patterns. A ten point 120 h *F. venenatum* elicitation time course in W22 stems utilized 38 day old plants. Graphs represent eight ordered RNA and protein module eigengenes (modules 2,4-10) which are the first principal components and summarize expression patterns for each module. Modules 1 and 3 are shown in Fig. 5 and not redrawn here. The vertical axes indicate expression values relative to the mean expression across all time points, which are indicated in the horizontal axes. RNA and protein fold changes compared to the 0 h time point were scaled and modified using a rank-order normalization. Numbers in parentheses denote the number of genes in each module (purple lines, RNA; blue lines, protein).

Supplementary Fig. 25



Supplementary Fig. 25 | A complex blend of ZX accumulate following stem infection with a necrotrophic fungal pathogen. a, GC/(+)CI-MS total ion chromatograms of extracts from representative maize stems 3 days after wounding or **b**, inoculation with *Cochliobolus heterostrophus* (*C. h.* 100 μl of 10^7 spores ml^{-1} ; SLB) analyzed as methyl ester derivatives. **c-g**, Representative extracted ion chromatograms of complex oxygenated ZXs following SLB infection. With reference to compound numbering in Supplementary Fig. 1, representative retention times (RT) in order and GC/(+)CI-MS *m/z* ions are as follows (1) β-bisabolene $[M+H]^+$ *m/z* 205, RT 9.81 min (not shown); (2) β-macrocarpene $[M+H]^+$ *m/z* 205; RT, 9.85 min (not shown) **c**, α/β-costic acids *m/z* 249, RT, 12.86 min; (3) ZD2 $[M+H]^+$ *m/z* 249, RT 13.14 min; (4) ZD1 $[M+H]^+$ *m/z* 249, RT 13.27 min; (5) ZA1 $[M+H]^+$ *m/z* 249, RT 13.55 min; (6) ZC1 $[M+H]^+$ *m/z* 245, RT 14.12 (not shown) **d**, (7) ZB1 $[M+H]^+$ *m/z* 247, RT 14.51; (8) ZA5 $[M+H]^+$ *m/z* 265 (fragment $[M-H_2O]^+$ *m/z* 247 shown) RT 14.88; (9) ZA2 (fragment $[M-H_2O]^+$ *m/z* 247 shown), RT 14.93; (10) ZA3 (fragment $[M-H_2O]^+$ *m/z* 247 shown), RT 15.32. **e**, (11) ZC2 $[M+H]^+$ *m/z* 261, RT 15.65. **f**, (12) ZB3 $[M+H]^+$ *m/z* 263, RT 15.87; (13) ZA4 $[M+H]^+$ *m/z* 263, RT 16.17; (14) ZA6 $[M-2H_2O]^+$ *m/z* 245 (fragment $[M-H_2O]^+$ *m/z* 263 shown), RT 16.75; (15) ZA7 $[M-2H_2O]^+$ *m/z* 245 (fragment $[M-H_2O]^+$ *m/z* 263 shown), RT 17.20. **g**, (16) ZA8 $[M+H]^+$ *m/z* 279, RT 17.21; (17) ZA9 $[M+H]^+$ *m/z* 279, RT 17.48. In each section (**a-g**), the y axis denotes relative abundance of ions.

Supplementary Fig. 26



Supplementary Fig. 26 | Quantification of ZX accumulation following stem infection with *C. heterostrophus* supports multiple predominant endproducts. Time course of ZX accumulation 1, 2 and 3 d after maize stems were either damaged and treated with 100 μl H_2O (D) or inoculated with *C. heterostrophus* (*C. h.* 100 μl of 10^7 spores ml^{-1}) analyzed as GC/(+)CI-MS. **a,** β-macrocarpene, **b**, ZA1; **c**, ZA2; **d**, ZA3; **e**, ZA4; **f**, ZA5; **g**, ZA6; **h**, ZA7; **i**, ZA8; **j**, ZA9; **k**, ZB1; **l**, ZB3; **m**, ZC2; **n**, ZD1; **o**, ZD2. Data are presented as mean ± s.e.m. (*n*=4 biologically independent samples except for the day 1 *C.h.*-treatment where the data is *n*=3). Within plots, different letters (a–d) represent significant differences (one-way ANOVA $P < 0.05$; Tukey's test corrections for multiple comparisons, $P < 0.05$).

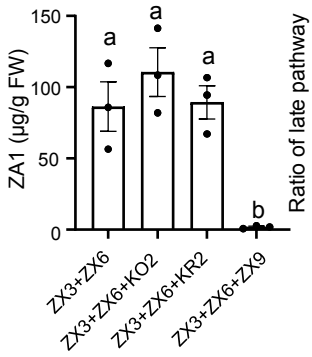
Supplementary Fig. 27

a

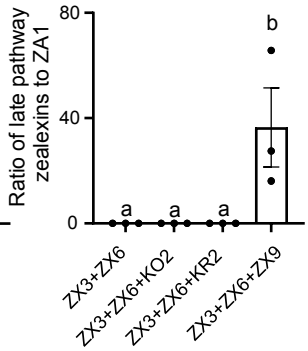
B73 RefGen V4	Annotation	Sum of Zx5 to Zx7	Zm00001d014134	Zm00001d029648	Zm00001d046342	Zm00001d041082	Zm00001d014121	Zm00001d032858	Zm00001d044147	Zm00001d024943	Zm00001d024207	Zm00001d024412	Zm00001d034096	Zm00001d014136	Zm00001d042814	Zm00001d025200	Zm00001d008837	Zm00001d016471	Zm00001d003283	Zm00001d029139	Zm00001d024210	Zm00001d011420	Zm00001d024208	Zm00001d024211
Sum of Zx5 to Zx7	Sum of Zx5 to Zx7																							
Zm00001d014134	Zx6 (CYP71Z18)	1																						
Zm00001d029648	AN2 (ZmCPS2)	1	2																					
Zm00001d046342	CYP701A43 (ZmKO2)	2	3	4																				
Zm00001d041082	ZmKSL2	2	2	3	4																			
Zm00001d014121	Zx5 (CYP71Z19)	3	7	6	11	10																		
Zm00001d032858	ZmKSL4	16	19	13	27	35	42																	
Zm00001d044147	CYP72A358	21	25	5	39	55	15	11																
Zm00001d024943	CYP93G6	32	35	34	46	39	174	108	34															
Zm00001d024207	Zx1	34	41	40	12	25	12	38	72															
Zm00001d024412	unknown P450	35	34	40	171	42	279	413	339	148	206													
Zm00001d034096	Zx9 (CYP81A38)	46	62	57	60	51	3	46	33		8	215												
Zm00001d014136	Zx7 (CYP71Z16)	52	90	100	160	145	125	59	43	129			374											
Zm00001d042814	CYP86A36	64	51	72	68	148		172	367	133		57												
Zm00001d025200	CYP71C60	66	74	75	91	53	87				35	401	95	581										
Zm00001d008837	CYP96E1	75	72	99	65	109	257	34	90	60	417	233	398		5									
Zm00001d016471	CYP73A8	124	122	27	293	295	303	185	71	282		390			145									
Zm00001d003283	CYP71S3	133	139	84	143	122	254					296	383											
Zm00001d029139	ZmTPS30	133	157	68	129	268	294	140	91	220		487	423	12										
Zm00001d024210	Zx3	137	160	182	102	182	46	194			1		19			18								
Zm00001d011420	CYP72A348	187	169	70	120	259			436						410			72	483					
Zm00001d024208	Zx2	346	362	465	242	44										29								
Zm00001d024211	Zx4	365	408		377	79	349				216		429			8					298		1	

1 500

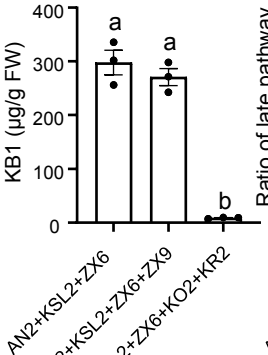
b



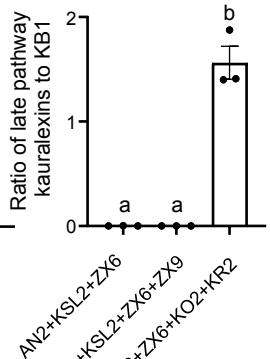
c



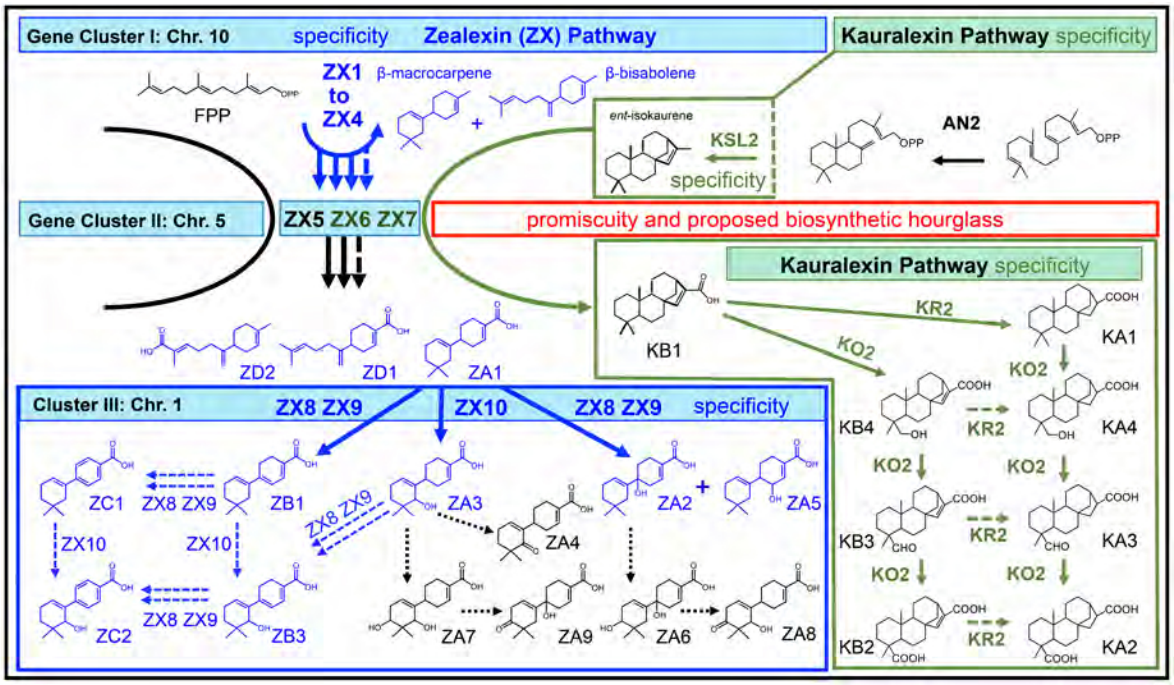
d



e



f

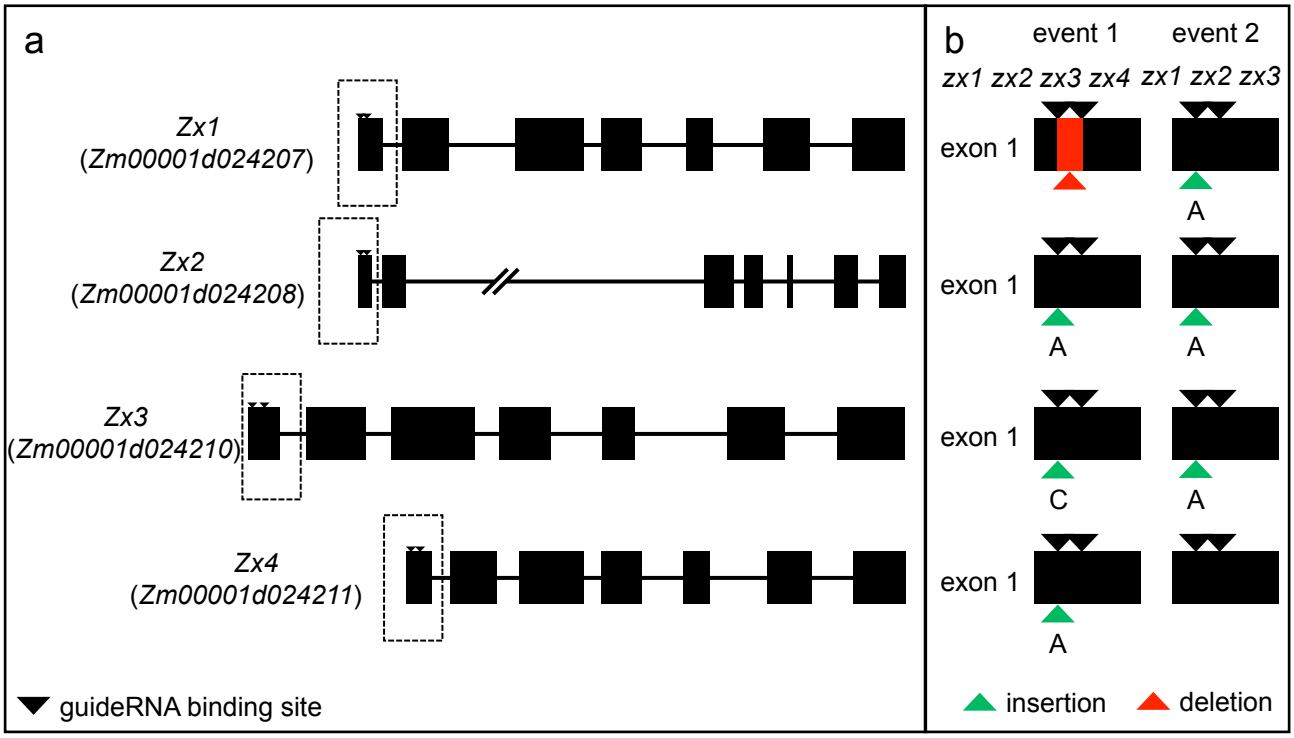


Supplementary Fig. 27

Fig. 27 | Promiscuous enzymes at an intermediate node in ZX and kauralexin biosynthesis support partially overlapping activities and are consistent with a biosynthetic hourglass harboring genetic redundancy to ensure production of protective antibiotic cocktails. a, Mutual Rank (MR) analysis showing co-expression of the sum of *Zx5* to *Zx7* with maize all *P450s* and all terpene synthase (*TPS*) genes in a dataset of 1960 RNA-seq samples. Collectively multiple individual kauralexin pathway genes (*ZmCYP71Z18/Zx6*, *ZmCYP71Z16/Zx7, ZmAN2*, *ZmKO2*, *ZmKSL2*)[1] are highly co-expressed with zealexin pathway genes (*Zx1*, *Zx2*, *Zx3*, *Zx4*, *Zx5*, *Zx6*, *Zx7*, *Zx9*). MR correlations > 500 were omitted. **b**, *Agrobacterium*-mediated transient expression of enzymes impacting ZA1 production in *N. benthamiana* leaves. Unlike late kauralexin pathway enzymes (ZmKO2 and ZmKR2), only ZX9 significantly reduces the accumulation of ZA1 consistent with precursor product relationships. **c**, The ratio of late pathway zealexins (sum of ZB1, ZA5, ZA2, ZC1) to ZA1. **d**, *Agrobacterium*-mediated transient expression of enzymes impacting KB1 production in *N. benthamiana* leaves. Unlike the late zealexin pathway enzyme ZX9, only late kauralexin pathway [kaurene oxidase 2 (ZmKO2) and kauralexin reductase 2 (ZmKR2)] enzymes significantly reduce the accumulation of KB1 consistent with precursor product relationships. **e**, The ratio of late pathway kauralexins (sum of KA1, KA2, KA3, KB2, and KB3) to KB1 from *Agrobacterium*-mediated transient expression in *N. benthamiana* leaves. **b-e**; Expressed protein combinations are indicated on the x-axis. No evidence for alternative or unknown reaction products of ZA1 or KB1 could be obtained. Error bars indicate mean ± s.e.m. (n = 3 biologically independent replicates) and different letters (a–b) represent significant differences (one-way ANOVA $P < 0.05$; Tukey's test corrections for multiple comparisons, $P < 0.05$). This experiment was replicated twice. **e**, Combined analyses of the ZX and kauralexin pathways[1] support a biosynthetic hourglass model based on both enzyme promiscuity and strong co-expression of *ZmCYP71Z18/Zx6* and *ZmCYP71Z16/Zx7* that connect with genes encoding the comparatively product specific early pathway enzymes (ZX1 to ZX4, ZmKSL2) and late pathway enzymes (ZX8, ZX9, ZX10, ZmKO2, ZmKR2).

[1]Ding, Y. *et al.* Multiple genes recruited from hormone pathways partition maize diterpenoid defences. *Nature Plants*, doi:10.1038/s41477-019-0509-6 (2019)

Supplementary Fig. 28



c

Zx1

WT	ACTGCAGATGGTCCTCGCCTCGGCCAGCAGGAGATGAAGAAGATGTC
zx1 zx2 zx3 zx4	ACTGCAGATGGTCCTCG----------
zx1 zx2 zx3	ACTGCAGATGGTCCTCGaCCTCGGCCAGCAGGAGATGAAGAAGATGTC

Zx2

WT	ACTGCAGATGGTCCTCGCCTCGGCCAGCAGGAGATGAAGAAGATGT
zx1 zx2 zx3 zx4	ACTGCAGATGGTCCTCGaCCTCGGCCAGCAGGAGATGAAGAAGATGT
zx1 zx2 zx3	ACTGCAGATGGTCCTCGaCCTCGGCCAGCAGGAGATGAAGAAGATGT

Zx3

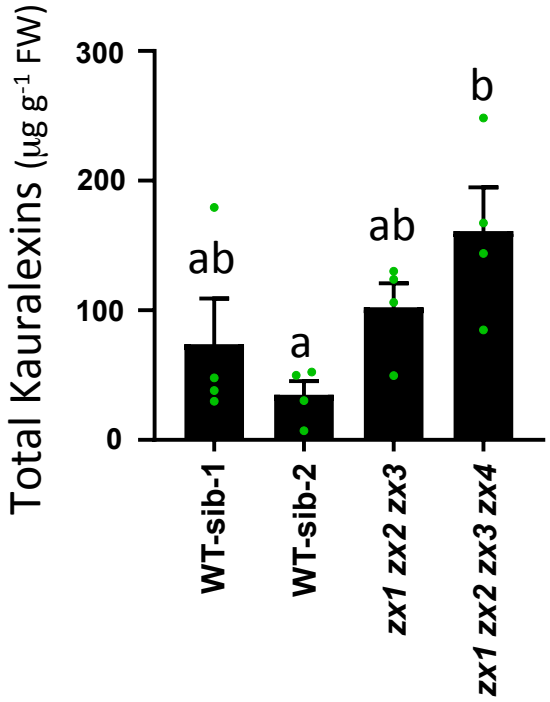
WT	ACTACAGATGGTCCTCGCCTCGGCCAGCAGGAGATGAAGAAGATGT
zx1 zx2 zx3 zx4	ACTACAGATGGTCCTCGcCCTCGGCCAGCAGGAGATGAAGAAGATGT
zx1 zx2 zx3	ACTACAGATGGTCCTCGaCCTCGGCCAGCAGGAGATGAAGAAGATGT

Zx4

WT	ACTACAGATGGTCCTCGCCTCGGCGAGCAGGAGATGAAGAAGATGT
zx1 zx2 zx3 zx4	ACTACAGATGGTCCTCGaCCTCGGCGAGCAGGAGATGAAGAAGATGT
zx1 zx2 zx3	ACTACAGATGGTCCTCGCCTCGGCGAGCAGGAGATGAAGAAGATGT

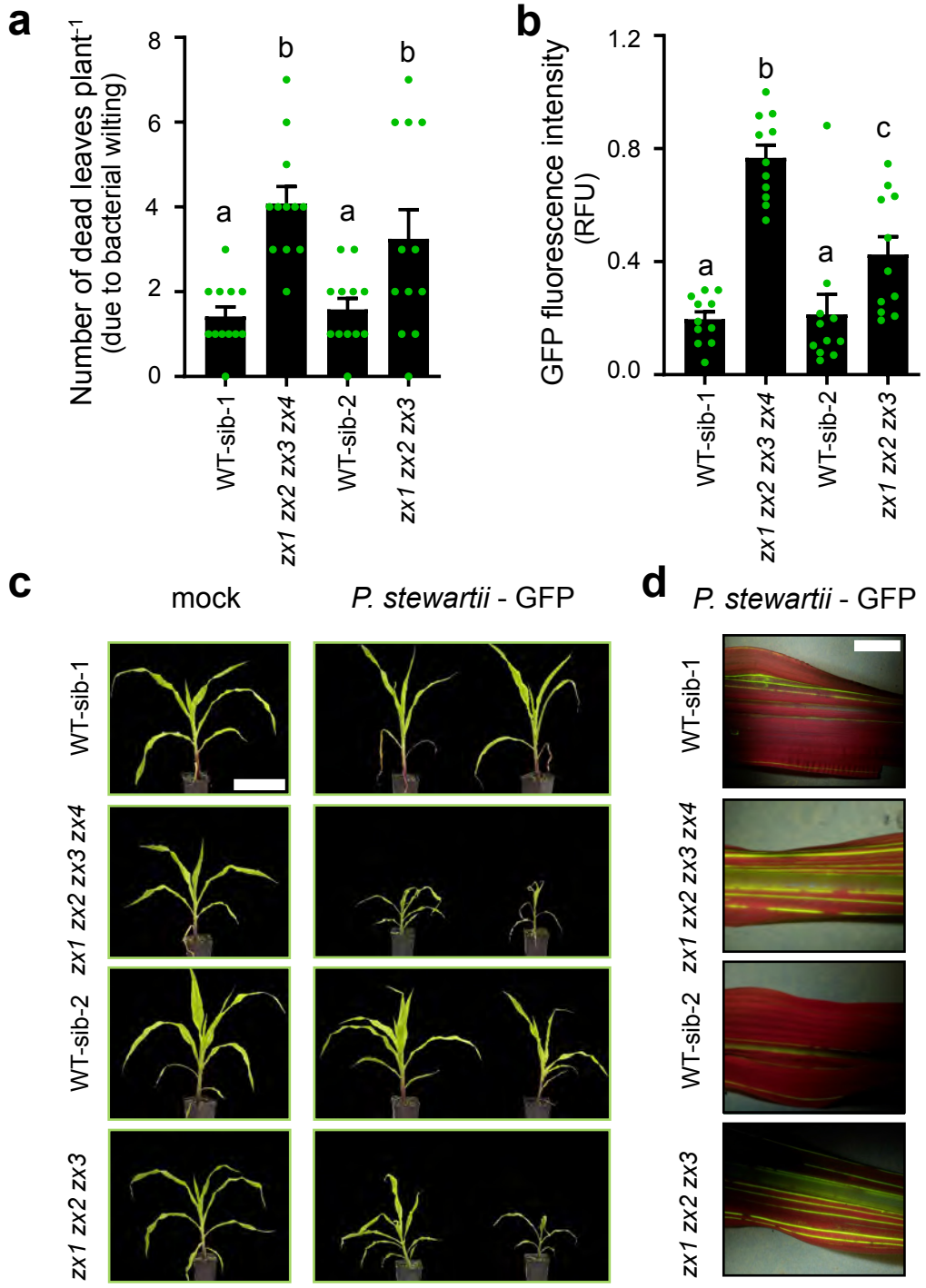
Supplementary Fig. 28 | Graphical representation of the zealexin biosynthetic pathway genes *Zx1* to *Zx4* and corresponding triple (*zx1 zx2 zx3*) and quadruple (*zx1 zx2 zx3 zx4*) mutants using CRISPR–Cas9. a, Gene structure of maize β-macrocarpene synthase genes *Zx1*, *Zx2*, *Zx3* and *Zx4*. Small arrows on top of exon one denote gRNA 1 and gRNA 2-binding sites, respectively. **b**, Graphical representation of the first exon of *Zx1* to *Zx4* and their respective mutations in triple and quadruple mutants created by CRISPR–Cas9. **c**, Partial sequence of *Zx1*, *Zx2*, *Zx3* and *Zx4* showing gRNA-binding site (underlined) and the respective mutations (--- deletion, green insertion of adenine or cytosine).

Supplementary Fig. 29



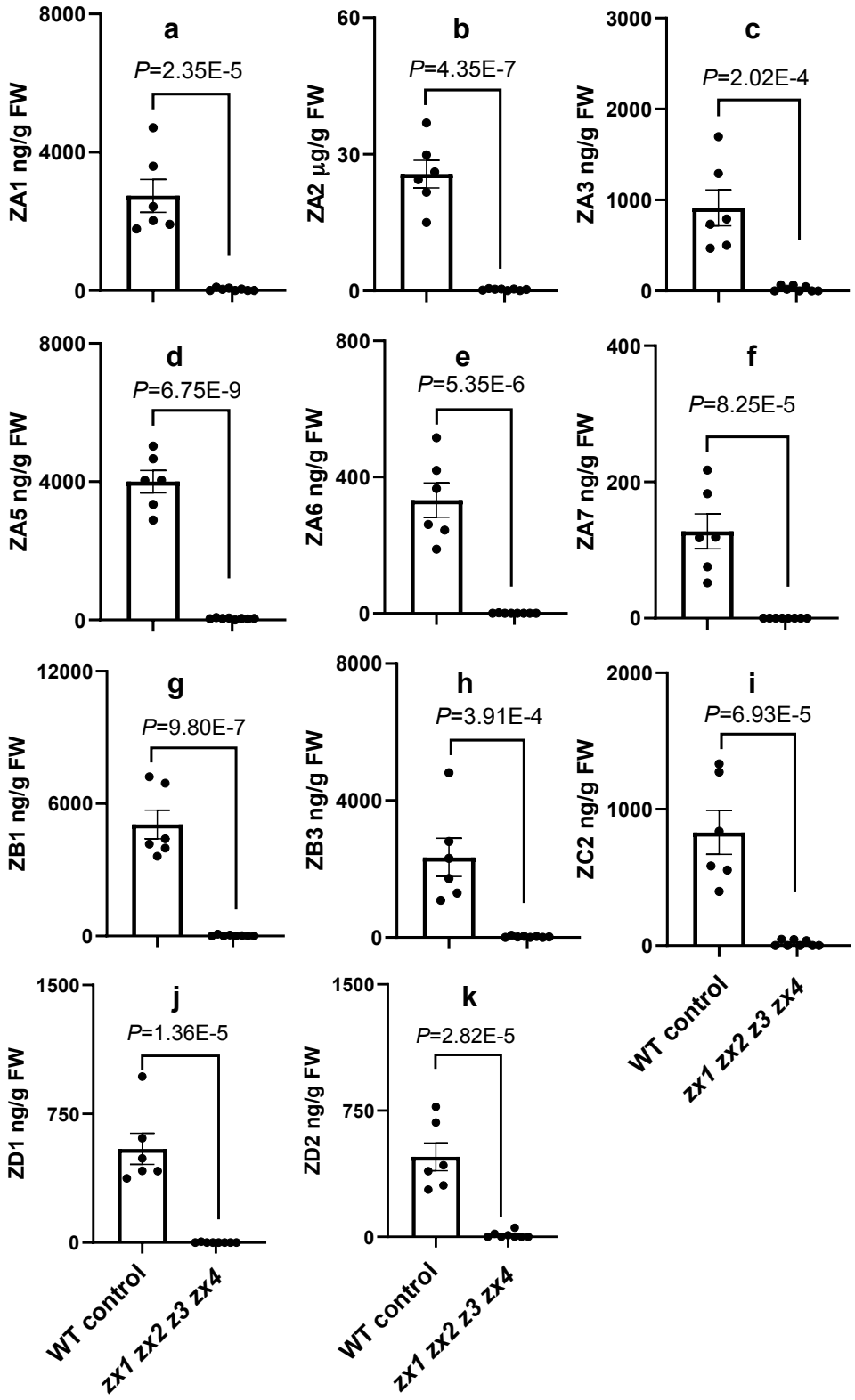
Supplementary Fig. 29 | Levels of total kauralexins present in zealexin triple (*zx1 zx2 zx3*) and quadruple (*zx1 zx2 zx3 zx4*) mutant plants in response to *Fusarium graminearum*. CRISPR/Cas9 derived triple and quadruple mutants and the respective wildtype siblings (WT-sib) were grown for 25 days and stem-inoculated with *Fusarium graminearum* (10 µL of 1.5 x 10^5 conidia ml^{-1}) for 10 days. Kauralexins were measured by GC-MS. Error bars in the bar chart indicate mean ± s.e.m (n = 4 biologically independent replicates). Within plots, different letters (a–b) represent significant differences (one-way ANOVA $P < 0.05$; Tukey's test corrections for multiple comparisons, $P < 0.05$).

Supplementary Fig. 30



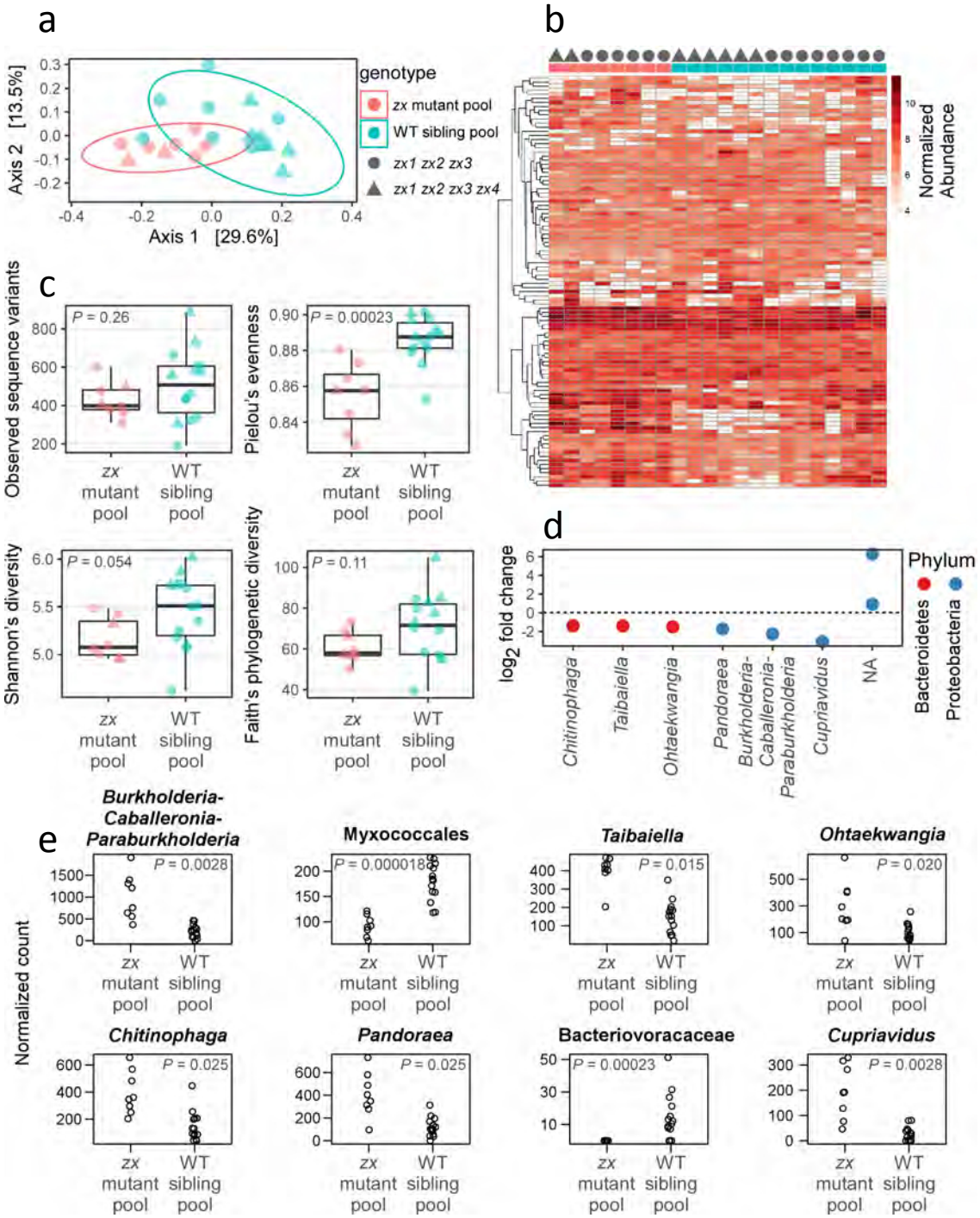
Supplementary Fig. 30 | ZX deficient triple and quadruple mutant plants display increased susceptibility to Stewarts wilt (*Pantoea stewartii*), a xylem-dwelling bacterium. **a**, Quantification of wilting symptoms in *P. stewartii* infected plants of triple (*zx1 zx2 zx3*), quadruple (*zx1 zx2 zx3 zx4*) and the respective wild type siblings (WT-sib). Data are presented as mean ± s.e.m. (n=12 biologically independent samples). **b**, Quantification of relative GFP fluorescence in *P. stewartii* infected plants. GFP fluorescence intensity was normalized to the highest fluorescence value and is shown as relative fluorescence units (RFU). Maize seedlings were infected with *P. stewartii*-GFP (DC283-GFP) or mock inoculated (controls, shown in **c**); GFP fluorescence was examined after 5 days post inoculation (dpi) while wilting was examined at 16 dpi. Data are presented as mean ± s.e.m. (n=11 biologically independent samples). Within plots, different letters (a–c) represent significant differences (one-way ANOVA $P < 0.05$; Tukey's test corrections for multiple comparisons; $P < 0.05$). **c**, Representative mock-inoculated or *P. stewartii* infected WT-sib-1, *zx1 zx2 zx3 zx4*, WT-sib-2, and *zx1 zx2 zx3* plants. Bar = 20 cm. **d**, Representative blue light illuminated leaf sections of *P. stewartii* infected WT-sib-1, *zx1 zx2 zx3 zx4*, WT-sib-2, and *zx1 zx2 zx3* plants. Bar = 1 cm.

Supplementary Fig. 31



Supplementary Fig. 31 | Quantification of ZX accumulation in roots of wild type (WT) and *zx1 zx2 zx3 zx4* quadruple mutant plants. ZX from roots of 8 week old field-soil grown plants were analyzed with GC-MS. **a,** ZA1, **b**, ZA2; **c**, ZA3; **d**, ZA5; **e**, ZA6; **f**, ZA7; **g**, ZB1; **h**, ZB3; **i**, ZC2; **j**, ZD1; **k**, ZD2. Data are presented as mean ± s.e.m. (WT control: *n* = 6 biologically independent samples; *zx1 zx2 zx3 zx4*: *n* = 8 biologically independent samples). *P* values represent Student's *t* test, two-tailed distribution, equal variance.

Supplementary Fig. 32



Supplementary Fig. 32 | Reduced ZX production substantially alters the bacterial microbiome associated with maize roots, lowering overall evenness and altering the abundances of particular taxa. a, Principal coordinates plot of bacterial communities associated with maize roots. Parenthetical percentages indicate the percentage of variation explained by each axis. **b**, Heatmap showing normalized counts in each sample for the 100 most abundant bacterial sequence variants associated with maize roots. Rows correspond to sequence variants, and were hierarchically clustered. **c**, diversity indices for bacterial communities associated with corn roots. *P* values represent Student's *t* test, two-tailed distribution. Box-plots show median with upper and lower quartiles; whiskers extend to 1.5x the interquartile range (*zx* mutant pool, *zx1 zx2 zx3 zx4* + *zx1 zx2 zx3*: $n = 8$ biologically independent samples; WT sibling pool: $n = 14$ biologically independent samples). **d**, Log2 fold-changes for sequence variants, identified to genus and phylum, that differed significantly in abundance between wild type plants vs. *zx* mutant pool (DESeq with Wald test, $P < 0.05$ with False Discovery Rate multiple testing correction; see panel **e** for adjusted *P*-values)). **e**, normalized counts across samples, for those sequence variants shown in **d**.

3. Supplementary Table 8

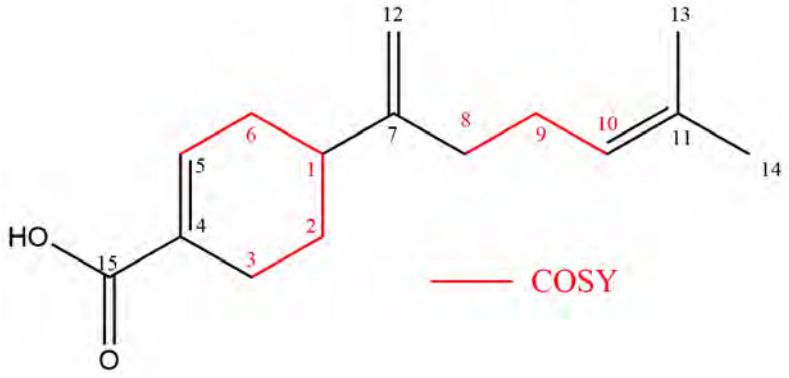
Supplementary Table 8: β-macrocarpene synthase derived ZXs purified and elucidated from fungal elicited maize tissue. All NMR data was acquired in 2.5-mm NMR tubes (Norell) at 22 °C using a 5-mm TXI cryoprobe (Bruker Corporation) and a Bruker Avance II 600 console (600 MHz for 1H and 151 MHz for 13C). One- and two-dimensional ^{1}H and ^{13}C NMR spectroscopy with heteronuclear multiple-bond correlations (HMBC) and assigned carbon numbers (C) were used for structural elucidation. In all figures, correlated spectroscopy (COSY) correlations are shown in red bonds. Zealexin NMR data was collected in $CDCl_3$ (ZA5, ZB3 methyl ester), acetonitrile-(d_3) (ZA8, ZA9) and benzene (d_6) (ZD1, ZD2, ZA6 methyl ester, ZA7 methyl ester) and the spectra were referenced to the following chemical shifts ^{1}H 7.26 ppm and ^{13}C 77.4 ppm for $CDCl_3$; ^{1}H 7.16 ppm and ^{13}C 128.1 ppm for benzene-d_6; and ^{1}H 1.94 ppm and ^{13}C 1.4 ppm for acetonitrile-d_3. NMR spectra were processed using Bruker Topspin 2.0 and MestReNova (Mestrelab Research) software packages. For each drawn compound structure, all carbons (C) are numbered. Coupling constants are given in Hertz [Hz]. In plant tissues all detected zealexins occur as free carboxylic acids. In select cases, for example ZB3, ZA6 and ZA7, final purification prior to NMR utilized methyl ester derivatives.

Supplementary Table: **8a.** Summarized NMR spectral data for ZD1
8b. Summarized NMR spectral data for ZD2
8c. Summarized NMR spectral data for ZA5
8d. Summarized NMR spectral data for ZB3 (methyl ester)
8e. Summarized NMR spectral data for ZA6 (methyl ester)
8f. Summarized NMR spectral data for ZA7 (methyl ester)
8g. Summarized NMR spectral data for ZA8
8h. Summarized NMR spectral data for ZA9

Supplementary Table: 8a. Summarized NMR spectral data for ZD1

4-(6-methylhepta-1,5-dien-2-yl)cyclohex-1-ene-1-carboxylic acid

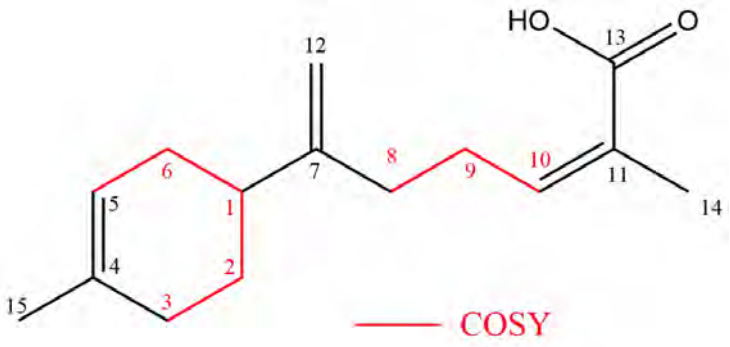
Position	δ ^{13}C [ppm]	δ ^{1}H [ppm]	J coupling constants [Hz]	HMBC correlations
1	72.2	-		
2	31.5	2H 1.82, 1.71	1.82, 1H, m 1.71, 1H, m	C4, C5
3	21.9	2H 2.32	m	
4	130.1	-		
5	137.9	1H 6.88	t J=1.54	C1, C3, C4, C15
6	37.1	2H 2.55, 2.21	2.55, 1H, q J=3.1 2.21 1H m	C7, C12
7	168.3	-		
8	51.7	2H 2.18	d *J* = 2.0	C9, C10, C11, C13, C14
9	34.4	-		
10	200.9	-		
11	40	2H 2.30	2.30, 2H, dd *J* =4.9, 1.5	
12	123	1H 6.01	t J=1.6	C1, C7, C8, C11
13	27.9	3H 1.01	s	C8, C9, C10, C11, C14
14	28.3	3H 1.02	s	C8, C9, C10, C11, C13
15	168	-		



Supplementary Table: 8b. Summarized NMR spectral data for ZD2

(E)-2-methyl-6-(4-methylcyclohex-3-en-1-yl)hepta-2,6-dienoic acid

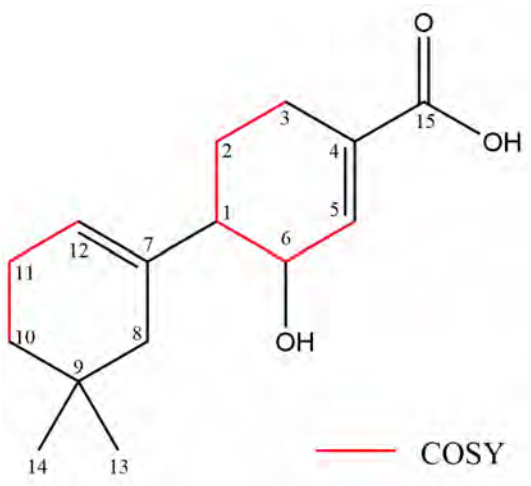
Position	δ ^{13}C [ppm]	δ ^{1}H [ppm]	J coupling constants [Hz]	HMBC correlations
1	39.6	1H 1.98 (m)	m	C3, C8
2	28.2	2H 1.67, 1.38	1.67, 1H, m 1.38, 1H, m	C1, C4
3	30.9	2H 1.91, 1.85	1.91, 1H 1.85, 1H	C1, C2, C4, C5
4	133.4	-		
5	120.8	1H 5.42	5.42, 1H, m	C1, C3, C6, C15
6	31.2	2H 2.04, 1.98	2.04, 1H, m 1.98, 1H, m	C1
7	152.9	-		
8	33.1	2H 1.89	m	C1, C7, C9, C10, C12
9	27.3	2H 2.02	m	C7, C8, C10, C11
10	144.7	1H 7.01	s	C8, C9, C11, C14, C13
11	127.6	-		
12 (14)	107.8	2H 4.83, 4.71	4.83, 1H, s 4.71, 1H, s	C1, C7, C8
13	173.8	-		
14 (12)	11.8	3H, 1.76	s	C8, C10, C11, C13
15	23.5	3H, 1.64	s	C3, C4, C5



Supplementary Table: 8c. Summarized NMR spectral data for ZA5

2-hydroxy-5',5'-dimethyl-[1,1'-bi(cyclohexane)]-1',3-diene-4-carboxylic acid

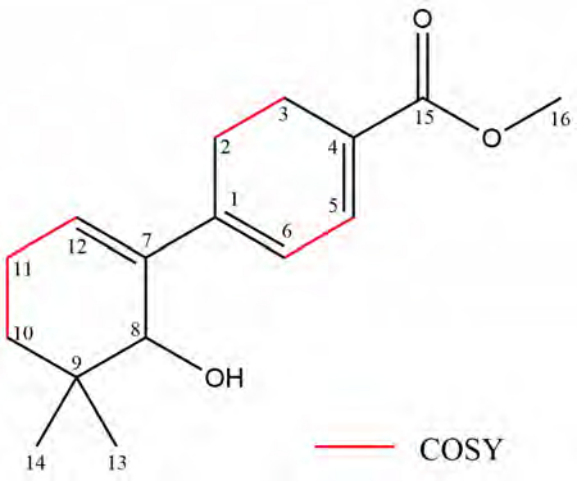
Position	δ ^{13}C [ppm]	δ ^{1}H [ppm]	J coupling constants [Hz]	HMBC correlations
1	50.2	1H 2.04	m	
2	25.7	2H 1.77, 1.60	1.77, 1H, m 1.60, 1H m	C1, C3, C4, C6
3	24.6	2H 2.41, 2.27	2.41, 1H, dd J=18.2, 5.3 2.27, 1H, m	C1, C4, C5
4	136.6	-		
5	142.8	1H 7.02	s	C1, C3, C4, C7, C15
6	68.5	1H 4.25	d J=9.3	C4, C5, C7
7	136.3	-		
8	39	2H 1.71	d J=8.4	C7, C14
9	29.1	-		
10	35.2	2H 1.35	t J=6.5	C8, C11, C12, C13, C14
11	23.3	2H 2.07	m	C7, C12
12	123.6	1H 5.57	s	C1, C8, C10, C11
13	27.9	3H 0.91	s	C7(w), C8, C9, C10, C14
14	28.7	3H 0.93	s	C7(w), C8, C9, C10, C13
15	171.8	-		



Supplementary Table: 8d. Summarized NMR spectral data for ZB3 (methyl ester)

6'-hydroxy-5',5'-dimethyl-[1,1'-bi(cyclohexane)]-1,1',3-triene-4-carboxylic acid methyl ester

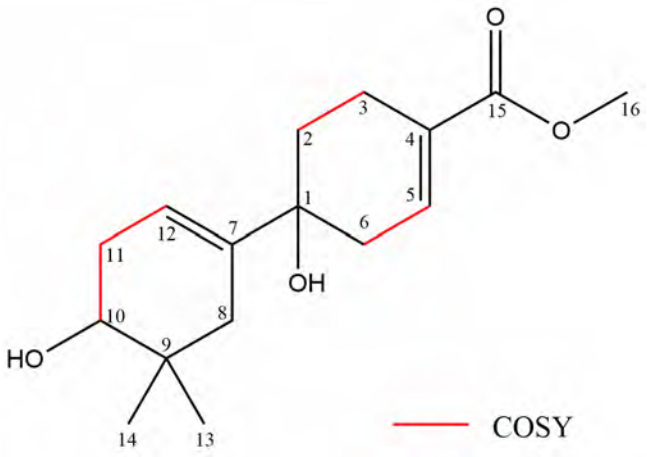
Position	δ ^{13}C [ppm]	δ ^{1}H [ppm]	J coupling constants [Hz]	HMBC correlations
1	142.6	-		
2	24.3	2H 2.52, 2.44	2.52, 1H, 2.44, 1H, m	C1, C3, C4, C6
3	22	2H 2.52	m	C1, C2, C4, C5
4	126.3	-		
5	134.6	1H 7.11	d, J=6.13	C1, C3, C4, C6, C15
6	118.4	1H 6.37	d, J=5.85	C3, C4, C7
7	137.6	-		
8	71.9	1H 3.92	br s	C1, C7, C9, C10, C11, C12
9	33.8	-		
10	28.8	2H 1.26, 1.68	1.26, 1H, 1.68, 1H, m	C1, C8, C9, C11, C12
11	24.2	2H 2.46, 2.62	2.46, 1H, 2.62, 1H, m	C8, C9, C10, C12
12	129.3	1H 6.12	t, J=4.17	C1, C10, C11, C8
13	23.7	3H 0.84	s	
14	26.4	3H 1.07	s	
15	167.9	-		
16	51.7	3H 3.76	s	C15



Supplementary Table: 8e. Summarized NMR spectral data for ZA6 (methyl ester)

1,4'-dihydroxy-5',5'-dimethyl-[1,1'-bi(cyclohexane)]-1',3-diene-4-carboxylic acid methyl ester

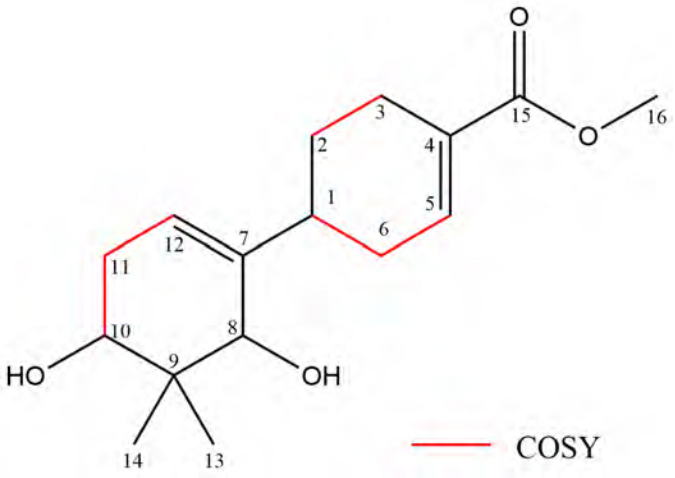
Position	δ ^{13}C [ppm]	δ ^{1}H [ppm]	J coupling constants [Hz]	HMBC correlations
1	71.2	-		
2	31.4	2H 1.42	m	C1, C3, C4, C7, C6
3	22.4	2H 2.49	m	C1, C2, C4, C5
4	130	-		
5	136.1	1H 6.92	m	C1, C15
6	37.2	2H 2.07, 1.92	2.07, 1H, m 1.92, 1H, m	C2, C4, C5, C7
7	141.3	-		
8	36.6	2H 1.81, 1.64	1.81, 1H, m 1.64, 1H, m	C7, C9, C10, C12, C13, C14
9	34	-		
10	73.1	1H 3.24	3.24, 1H, m	C8, C9, C13, C14
11	32.1	2H 2.12, 1.82	2.12, 1H, m 1.82, 1H, m	C7, C9, C10, C12
12	116.7	1H 5.33	5.33, 1H, m	C8, C10, C11
13	26.3	3H 0.84	s	C1, C7, C8, C9, C10, C14
14	21.7	3H 0.86	s	C9, C10, C13
15	170.8	-		
16	50.9	3H 3.45	s	



Supplementary Table: 8f. Summarized NMR spectral data for ZA7 (methyl ester)

4',6'-dihydroxy-5',5'-dimethyl-[1,1'-bi(cyclohexane)]-1',3-diene-4-carboxylic acid methyl ester

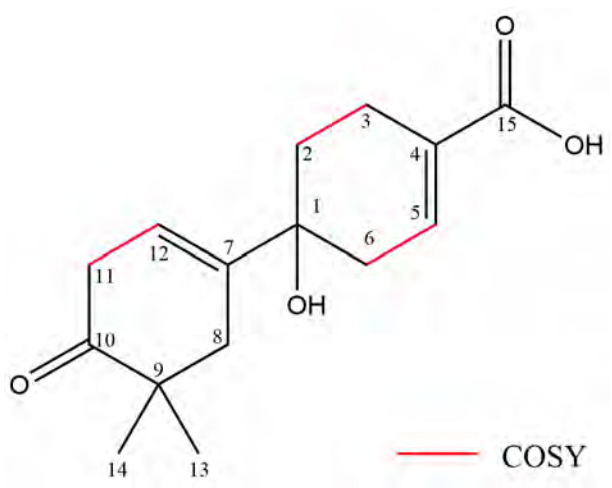
Position	δ ^{13}C [ppm]	δ ^{1}H [ppm]	J coupling constants [Hz]	HMBC correlations
1	36.7	1H 2.13	2.13, 1H, m	C2, C6
2	27.9	2H 1.65, 1.21	1.65, 1H, m 1.21, 1H, m	C1, C3, C4, C6, C7
3	24.9	2H 2.51, 2.31	2.51, 1H, m 2.31, 1H, m	C1, C2, C4, C5
4	130.3	-		
5	139.3	1H 7.06	7.06, 1H, m	C1, C2, C6, C15
6	31.3	2H 2.04, 1.90	2.04, 1H, m 1.90, 1H, m	C1, C3, C4, C5
7	142	-		
8	75.8	1H 3.38	3.38, 1H, s	C1, C7, C9, C10, C11, C12, C13, C14
9	39.1	-		
10	69.8	1H 3.63	3.63, 1H, dd J=8.8, 5.9	C8, C9, C11, C12, C13, C14
11	32.4	2H 2.12, 1.75	2.12, 1H, m 1.75, 1H, m	C10, C12, C17
12	120.4	1H 5.04	5.04, 1H, ddd J=4.7, 2.9, 1.0	C1, C7, C8, C10, C11
13	21.9	3H 1.04	s	C8, C9, C10, C14
14	17.8	3H 0.75	s	C8, C9, C10, C13
15	167.3	-		
16	50.88	3H 3.47	s	



Supplementary Table: 8g. Summarized NMR spectral data for ZA8

1-hydroxy-5',5'-dimethyl-4'-oxo-[1,1'-bi(cyclohexane)]-1',3-diene-4-carboxylic acid

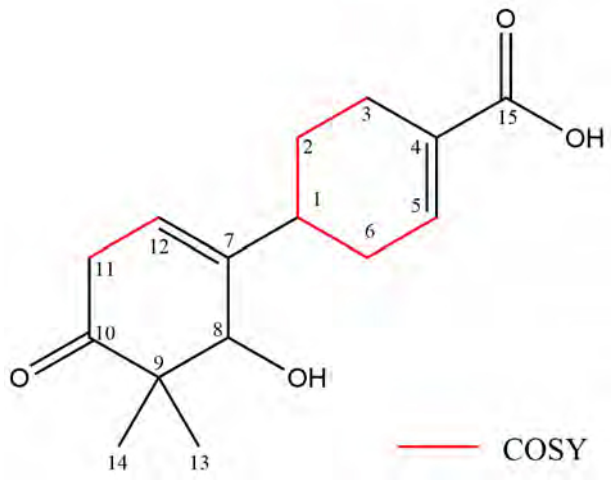
Position	δ ^{13}C [ppm]	δ ^{1}H [ppm]	J coupling constants [Hz]	HMBC correlations
1	72.2	-		
2	31.5	2H 1.82, 1.71	1.82, 1H, m 1.71, 1H, m	C4, C5
3	21.9	2H 2.32	m	
4	130.1	-		
5	137.9	1H 6.88	t J=1.54	C1, C3, C4, C15
6	37.1	2H 2.55, 2.21	2.55, 1H, q J=3.1 2.21 1H m	C7, C12
7	168.3	-		
8	51.7	2H 2.18	d *J* = 2.0	C9, C10, C11, C13, C14
9	34.4	-		
10	200.9	-		
11	40	2H 2.30	2.30, 2H, dd *J* =4.9, 1.5	
12	123	1H 6.01	t J=1.6	C1, C7, C8, C11
13	27.9	3H 1.01	s	C8, C9, C10, C11, C14
14	28.3	3H 1.02	s	C8, C9, C10, C11, C13
15	168	-		



Supplementary Table: 8h. Summarized NMR spectral data for ZA9

6'-hydroxy-5',5'-dimethyl-4'-oxo-[1,1'-bi(cyclohexane)]-1',3-diene-4-carboxylic acid

Position	δ ^{13}C [ppm]	δ ^{1}H [ppm]	J coupling constants [Hz]	HMBC correlations
1	37.4	1H 2.63	br m	C7
2	27.3	2H 1.97, 1.55	1.97, 1H, m 1.55, 1H, m	C1, CC4, C6
3	24.5	2.32 1H m, 2.26 1H m	2.32, 1H, m 2.26, 1H, m	C6
4	130.4	-		
5	139.8	1H 6.99	br m	C1, C3, C15
6	30.8	2H 2.42, 2.25	2.42, 1H, m 2.25, 1H, m	C2, C4, C5
7	167.4	-		
8	74.6	1H 4.00	d J=6.2	C7, C11, C12, C13, C14
9	38.4	-		
10	199.6	-		
11	48.3	2h 2.39, 2.10	2.39, 1H, m 2.10, 1H, m	C8, C9, C10, C13, C14
12	124.1	1H 5.69	br s	C1, C8, C11
13	26.4	3H 0.97	s	C8, C9, C10, C11, C14
14	23.3	0.98 3H s	s	C8, C9, C10, C11, C13
15	168.1	-		



4. Supplementary References

1 Sadre, R. *et al.* Cytosolic lipid droplets as engineered organelles for production and accumulation of terpenoid biomaterials in leaves. *Nat Commun* **10**, 853 (2019).

2 Bach, S. S. *et al.* High-throughput testing of terpenoid biosynthesis candidate genes using transient expression in *Nicotiana benthamiana*. *Methods Mol. Biol.* **1153**, 245-255 (2014).

3 Morrone, D. *et al.* Increasing diterpene yield with a modular metabolic engineering system in *E. coli*: comparison of MEV and MEP isoprenoid precursor pathway engineering. *Appl. Microbiol. Biotechnol.* **85**, 1893-1906 (2010).

4 Murphy, K. M. *et al.* A customizable approach for the enzymatic production and purification of diterpenoid natural products. *J Vis Exp* 152, e59992 (2019).

5 Mafu, S. *et al.* Discovery, biosynthesis and stress-related accumulation of dolabradiene-derived defenses in maize. *Plant Physiol.* **176**, 2677-2690 (2018).

6 Huffaker, A. *et al.* Novel acidic sesquiterpenoids constitute a dominant class of pathogen-induced phytoalexins in maize. *Plant Physiol.* **156**, 2082-2097 (2011).

7 Schmelz, E. A., Engelberth, J., Tumlinson, J. H., Block, A. & Alborn, H. T. The use of vapor phase extraction in metabolic profiling of phytohormones and other metabolites. *Plant J.* **39**, 790-808 (2004).

8 Starks, C. M., Back, K., Chappell, J. & Noel, J. P. Structural basis for cyclic terpene biosynthesis by tobacco 5-*epi*-aristolochene synthase. *Science* **277**, 1815-1820 (1997).

9 Kremling, K. A. G. *et al.* Dysregulation of expression correlates with rare-allele burden and fitness loss in maize. *Nature* **555**, 520-523 (2018).

10 Ding, Y. *et al.* Multiple genes recruited from hormone pathways partition maize diterpenoid defences. *Nature Plants* **10**, 1043-1056 (2019).

11 Wisecaver, J. H. *et al.* A global co-expression network approach for connecting genes to specialized metabolic pathways in plants. *The Plant Cell* **29**, 944-959 (2017).

12 Cannon, E. K. *et al.* POPcorn: An online resource providing access to distributed and diverse maize project data. *Int J Plant Genomics* **2011**, 923-935 (2011).

13 Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).

14 Guy, L., Kultima, J. R. & Andersson, S. G. genoPlotR: comparative gene and genome visualization in R. *Bioinformatics* **26**, 2334-2335 (2010).

15 Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059-3066 (2002).

16 Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268-274 (2015).
17 Lee, M. *et al.* Expanding the genetic map of maize with the intermated B73 x Mo17 (IBM) population. *Plant Mol. Biol.* **48**, 453-461 (2002).
18 Eichten, S. R. *et al.* B73-Mo17 Near-isogenic lines demonstrate dispersed structural variation in maize. *Plant Physiol.* **156**, 1679-1690 (2011).
19 Flint-Garcia, S. A. *et al.* Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant J.* **44**, 1054-1064 (2005).
20 Bradbury, P. J. *et al.* TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633-2635 (2007).
21 Yu, J. M. *et al.* A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **38**, 203-208 (2006).
22 Samayoa, L. F., Malvar, R. A., Olukolu, B. A., Holland, J. B. & Butron, A. Genome-wide association study reveals a set of genes associated with resistance to the Mediterranean corn borer (*Sesamia nonagrioides* L.) in a maize diversity panel. *BMC Plant Biol.* **15**, 35 (2015).
23 Zhang, Z. W. *et al.* Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* **42**, 355-360 (2010).
24 Lipka, A. E. *et al.* GAPIT: genome association and prediction integrated tool. *Bioinformatics* **28**, 2397-2399 (2012).
25 VanRaden, P. M. Efficient methods to compute genomic predictions. *J. Dairy Sci.* **91**, 4414-4423 (2008).
26 Turner, S. D. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv*, doi:10.1101/005165 (2014).
27 Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
28 Bouckaert, R. *et al.* BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* **15**, e1006650 (2019).
29 Drummond, A. J. & Suchard, M. A. Bayesian random local clocks, or one rate to rule them all. *BMC Biol.* **8**, 114 (2010).
30 Swigonova, Z. *et al.* Close split of sorghum and maize genome progenitors. *Genome Res.* **14**, 1916-1923 (2004).
31 Bouckaert, R. R. DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics* **26**, 1372-1373 (2010).
32 Liu, D. *et al.* Validation of reference genes for gene expression studies in virus-infected *Nicotiana benthamiana* using quantitative real-time PCR. *PLoS One* **7**, e46451 (2012).
33 Horevaj, P., Milus, E. A. & Bluhm, B. H. A real-time qPCR assay to quantify *Fusarium graminearum* biomass in wheat kernels. *J. Appl. Microbiol.* **111**, 396-406 (2011).
34 Schmelz, E. A. *et al.* Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 5455-5460 (2011).

35 Garcia, N., Li, Y., Dooner, H. K. & Messing, J. Maize defective kernel mutant generated by insertion of a Ds element in a gene encoding a highly conserved TTI2 cochaperone. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 5165-5170 (2017).
36 de Hoon, M. J., Imoto, S., Nolan, J. & Miyano, S. Open source clustering software. *Bioinformatics* **20**, 1453-1454 (2004).
37 Saldanha, A. J. Java Treeview--extensible visualization of microarray data. *Bioinformatics* **20**, 3246-3248 (2004).
38 Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
39 Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, Article 17, (2005).
40 Alexa, A. & Rahnenfuhrer, J. Gene set enrichment analysis with topGO. (https://bioconductor.org/packages/release/bioc/vignettes/topGO/inst/doc/topGO.pdf) (2016).
41 Wimalanathan, K., Friedberg, I., Andorf, C. M. & Lawrence-Dill, C. J. Maize GO Annotation-Methods, Evaluation, and Review (maize-GAMER). *Plant Direct* **2**, e00052 (2018).
42 TW, H. B. & Girke, T. systemPipeR: NGS workflow and report generation environment. *BMC Bioinformatics* **17**, 388 (2016).
43 Brazelton, V. A. *et al.* A quick guide to CRISPR sgRNA design tools. *Gm Crops & Food-Biotechnology in Agriculture and the Food Chain* **6**, 266-276 (2015).
44 Char, S. N. *et al.* An *Agrobacterium*-delivered CRISPR/Cas9 system for high-frequency targeted mutagenesis in maize. *Plant Biotechnology Journal* **15**, 257-268 (2017).
45 Jiao, Y. *et al.* Improved maize reference genome with single-molecule technologies. *Nature* **546**, 524-527 (2017).
46 Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. U. S. A.* **108 Suppl 1**, 4516-4522 (2011).
47 R Development Core Team, R. (R foundation for statistical computing Vienna, Austria, 2011).
48 Callahan, B. J., Sankaran, K., Fukuyama, J. A., McMurdie, P. J. & Holmes, S. P. Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. *F1000Res* **5**, 1492 (2016).
49 Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal* **17**, 3 (2011).
50 Edgar, R. C. & Flyvbjerg, H. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* **31**, 3476-3482 (2015).
51 Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**, 581-583 (2016).

52 Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261-5267 (2007).
53 Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590-596, doi:10.1093/nar/gks1219 (2013).
54 Schliep, K. P. Phangorn: phylogenetic analysis in R. *Bioinformatics* **27**, 592-593, doi:10.1093/bioinformatics/btq706 (2011).
55 McMurdie, P. J. & Holmes, S. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217, doi:10.1371/journal.pone.0061217 (2013).
56 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
57 Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer Publishing Company, Incorporated, 2009).
58 Kolde, R. Pheatmap: pretty heatmaps. *R package version* **61**, 915 (2012).