nature plants

ARTICLES

https://doi.org/10.1038/s41477-019-0509-6

Multiple genes recruited from hormone pathways partition maize diterpenoid defences

Yezhang Ding[1], Katherine M. Murphy[2], Elly Poretsky[1], Sibongile Mafu[2], Bing Yang[3], Si Nian Char[3], Shawn A. Christensen[4], Evan Saldivar[1], Mengxi Wu[1], Qiang Wang[5], Lexiang Ji[6], Robert J. Schmitz[7], Karl A. Kremling[8], Edward S. Buckler[8,9], Zhouxin Shen[1], Steven P. Briggs[1], Jörg Bohlmann[10], Andrew Sher[1], Gabriel Castro-Falcon[11], Chambers C. Hughes[11], Alisa Huffaker[1], Philipp Zerbe[2] and Eric A. Schmelz[1*]

Duplication and divergence of primary pathway genes underlie the evolution of plant specialized metabolism; however, mechanisms partitioning parallel hormone and defence pathways are often speculative. For example, the primary pathway intermediate *ent*-kaurene is essential for gibberellin biosynthesis and is also a proposed precursor for maize antibiotics. By integrating transcriptional coregulation patterns, genome-wide association studies, combinatorial enzyme assays, proteomics and targeted mutant analyses, we show that maize kauralexin biosynthesis proceeds via the positional isomer *ent*-isokaurene formed by a diterpene synthase pair recruited from gibberellin metabolism. The oxygenation and subsequent desaturation of *ent*-isokaurene by three promiscuous cytochrome P450s and a new steroid 5α reductase indirectly yields predominant *ent*-kaurene-associated antibiotics required for *Fusarium* stalk rot resistance. The divergence and differential expression of pathway branches derived from multiple duplicated hormone-metabolic genes minimizes dysregulation of primary metabolism via the circuitous biosynthesis of *ent*-kaurene-related antibiotics without the production of growth hormone precursors during defence.

To navigate complex interactions with their environment, plants synthesize myriad specialized small molecule metabolites[1–3]. With dynamic and tissue-specific accumulation patterns, specialized metabolites serve vital functions in plant development, defence and chemical ecology[1,2,4]. Environmental selection pressures, combined with gene duplication and functional divergence, are dominant drivers in gene recruitment enabling specialized pathway expansion[3,5,6]. Moreover, conserved primary pathways, especially those underlying hormone biosynthesis, provide genetic reservoirs for the continued derivation of bioactive metabolites[7–9]. Precisely how primary and specialized metabolic pathways are regulated directly impacts stress resistance[1].

Phytohormone-biosynthetic and signalling pathways are common targets for pathogen attack[10]. Across diverse pathosystems, numerous mechanisms driving disease susceptibility are mediated by pathogen-induced phytohormone dysregulation[10–12]. For example, plant development relies upon gibberellin (GA) biosynthesis generating a series of oxygenated diterpenoid phytohormones[13]. Prior to our awareness of plant GAs[14,15], bioactive GAs were discovered as potent virulence factors produced by the fungus *Gibberella fujikuroi*, which causes bakanae disease in rice (*Oryza sativa*)[16]. Thus, the GA pathway collectively serves not only as a source of growth hormones but also as a pathogenesis target and genetic reservoir for defence evolution[6,10,12,13,17]. Pairs of prototypical class II and class I diterpene synthases (diTPS), namely *ent*-copalyl diphosphate synthases (*ent*-CPS) and *ent*-kaurene synthases (*ent*-KS) that form the central GA precursor *ent*-kaurene, have been repeatedly duplicated and functionalized in the expansion of diterpenoid chemical diversity[18].

In Poaceous grain crops, lineage-specific diterpenoid networks mediate diverse antimicrobial and allelopathic functions[17,19]. In rice, combinations of CPS, kaurene synthase-like (KSL) diTPS and cytochrome P450 (CYP) monooxygenases collectively mediate the biosynthesis of protective momilactones, oryzalexins and phytocassanes[19,20]. More recently, maize (*Zea mays*) was also discovered to contain unique fungal-elicited pathways of *ent*-CPS-requiring antibiotic diterpenoids, termed dolabralexins and kauralexins, which are derived from dolabradiene and *ent*-kaurene, respectively[21–24]. Loss of both pathways through mutations in the maize *ent*-CPS, *Anther ear 2* (*ZmAN2*), causes increased fungal susceptibility[21,23,25]. In addition to the core enzymes of GA metabolism, namely the *ent*-CPS termed Anther Ear 1 (ZmAN1) and the *ent*-KS ZmKSL3, the maize genome also contains four *CPS* and seven *KSL* genes[19,22,26]. This includes two additional *ent*-KS, *ZmTPS1* and *ZmKSL5*, that could serve GA metabolism[22]. Furthermore, the kaurene oxidase ZmKO1 (ZmCYP701A26) is a P450 displaying substrate promiscuity prototypical for the GA pathway enzyme[27,28]. Duplicated maize GA biosynthetic genes predictably dominate fungal-inducible

[1]Section of Cell and Developmental Biology, University of California San Diego, La Jolla, CA, USA. [2]Department of Plant Biology, University of California Davis, Davis, CA, USA. [3]Department of Genetics, Development and Cell Biology, Iowa State University, Ames, IA, USA. [4]Chemistry Research Unit, Center for Medical, Agricultural, and Veterinary Entomology, US Department of Agriculture—Agricultural Research Service, Gainesville, FL, USA. [5]Institute of Ecological Agriculture, Sichuan Agricultural University, Chengdu, China. [6]Institute of Bioinformatics, University of Georgia, Athens, GA, USA. [7]Department of Genetics, University of Georgia, Athens, GA, USA. [8]Department of Plant Breeding and Genetics, Cornell University, Ithaca, NY, USA. [9]Robert W. Holley Center for Agriculture and Health, US Department of Agriculture—Agricultural Research Service, Ithaca, NY, USA. [10]Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, Canada. [11]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA. *e-mail: eschmelz@ucsd.edu

production of kauralexin antibiotics; however, an inherent challenge remains: how can phytohormone and defence-related biosynthetic steps be duplicated, while avoiding pathogen-induced dysregulation of primary pathways?

Combining transcriptional coregulation and genome-wide association studies (GWAS) with combinatorial enzyme assays, proteomics and mutant analyses, we demonstrate that kauralexin biosynthesis proceeds via the *ent*-kaurene positional isomer, *ent*-isokaurene. The pathway recruits a diTPS pair and a CYP701A-family P450 derived from duplicated GA metabolic genes that retain ancestral activity yet display regulatory functionalization for pathogen-induced expression. Two functionally redundant, promiscuous CYP71 family P450s form predominantly *ent*-isokaurene-derived kauralexins that are further converted by a new steroid 5α reductase to yield *ent*-kaurene-associated kauralexins indirectly, while avoiding *ent*-kaurene as a predominant precursor. Mutant analyses demonstrate that only *ent*-isokaurene is required for both kauralexin production and *Fusarium* resistance. Collectively, the maize kauralexin pathway highlights multiple functional mechanisms to bypass the biosynthesis of GA precursors, while relying on duplicated GA biosynthetic genes for defence.

Results

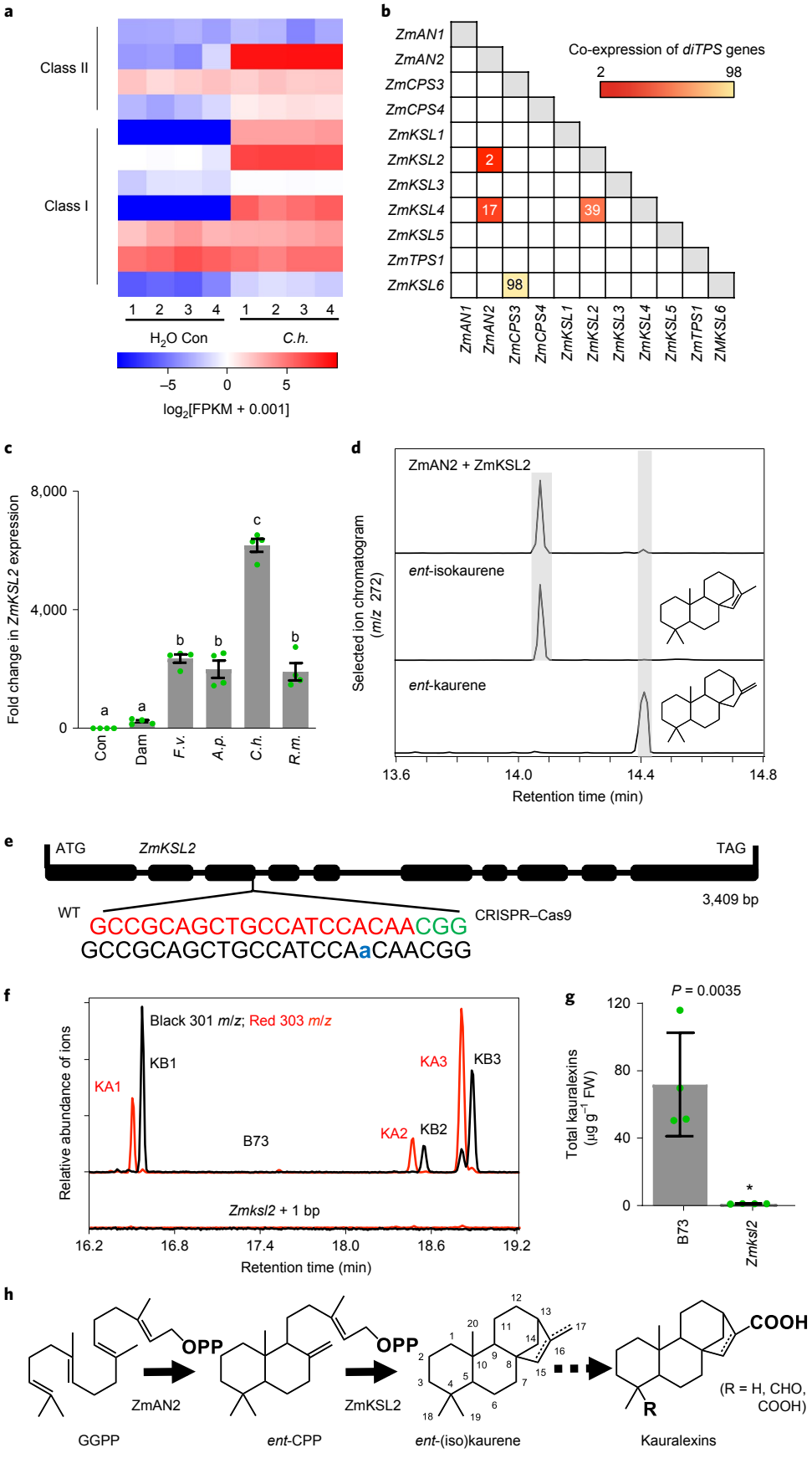
To understand the source of fungal-elicited kauralexins predicted to share the GA precursor *ent*-kaurene[22], southern leaf blight (SLB; *Cochliobolous heterostrophus*) inoculation was used to generate a replicated fungal-elicited transcriptome in maize leaves (Supplementary Table 1). Examination of resulting RNA-seq expression patterns of all diTPS genes (Fig. 1a) confirmed that *ZmAN2* was the only strongly induced class II diTPS[24,29]. Of the seven class I diTPS present in maize, *ZmKSL4* is specific for the dolabralexin pathway[23]. Transcript levels of the *ent*-kaurene synthases, *ZmTPS1* and *ZmKSL5*, were not significantly increased (Fig. 1a and Supplementary Table 2)[22]. The remaining kauralexin biosynthetic candidate genes, *ZmKSL1*, *ZmKSL2*, *ZmKSL3* and *ZmKSL6* exhibited statistically significant transcript accumulation following SLB challenge (Fig. 1a and Supplementary Table 2). To focus hypothesized endogenous relationships, mutual rank (MR)-based global gene co-expression analyses were used to associate transcript accumulation patterns of maize diTPS in an expansive combined RNA-seq dataset[30]. Select gene co-expression analyses revealed the highest degree of coregulation between *ZmAN2* and *ZmKSL2* (Fig. 1b), suggesting a role in kauralexin production. As a visual aid, all biochemicals and genes with examined relevance to the kauralexin biosynthesis are summarized (Supplementary Figs. 1 and 2 and Supplementary Table 2). In contrast to the largely constitutive GA pathway transcripts such as *ZmAN1* and *ZmKSL3*, the significant accumulation of *ZmKSL2, ZmAN2* and kauralexin pathway metabolites following attack by four different fungal pathogens demonstrate significant coregulation during elicitation (Fig. 1c and Supplementary Fig. 2). Phylogenic analyses of maize, wheat, rice, switchgrass and sorghum diTPS place ZmKSL2 and ZmKSL4 in a separate distant branch to *ent*-kaurene synthases ZmKSL3 and ZmKSL5 (Supplementary Fig. 3)[22].

ZmKSL2 catalyses production of the *ent*-kaurene positional isomer, *ent*-isokaurene. *Agrobacterium*-mediated transient co-expression of ZmAN2 and ZmKSL2 in *N. benthamiana* yielded *ent*-isokaurene as a major product with trace amounts (<5%) of *ent*-kaurene, demonstrating that ZmKSL2 predominantly functions as an *ent*-isokaurene synthase (Fig. 1d and Supplementary Fig. 4). To characterize the full range of potential endogenous ZmKSL2 products, we performed paired *Escherichia coli* co-expression assays with all CPS present in the maize genome, namely *ent*-CPS (ZmAN2), (+)-CPS (ZmCPS3) and a labda-8,13-dien-15-yl diphosphate (8,13-CPP) synthase (ZmCPS4)[31] using an engineered *E. coli* platform[32]. Consistent with an *ent*-isokaurene synthase, pairwise activity of ZmKSL2 with ZmAN2 converted geranylgeranyl-diphosphate (GGPP) into predominately *ent*-isokaurene, whereas pairing with ZmCPS3 and ZmCPS4 resulted in the production of an unknown diterpenoid hydrocarbon and manoyl oxide, respectively (Supplementary Fig. 5). ZmKSL2 substrate promiscuity is similar to ZmKSL4[23] and related class I diTPS from rice and wheat[33,34]. However, other than *ent*-isokaurene (Supplementary Fig. 2), specific products from different ZmCPS-ZmKSL2 pairings have yet to be observed as maize metabolites. To examine the endogenous role of *ZmKSL2* we generated an insertion-based frame shift mutation using clustered regularly interspaced short palindromic repeats (CRISPR)–Cas9 gene editing (Fig. 1e). *Zmksl2* plants displayed a near-complete deficiency in kauralexin accumulation following elicitation with heat-killed *Fusarium* hyphae (Fig. 1f,g). Thus, ZmKSL2 and its major respective product *ent*-isokaurene are required for significant kauralexin biosynthesis.

Tandem duplicate CYP71 P450s catalyse committed kauralexin biosynthesis. Cytochrome P450s are essential for GA biosynthesis and dramatically expand the diversity of specialized metabolic pathways by facilitating an array of functional modifications[35]. To identify P450 candidates in the kauralexin biosynthetic pathway, we analysed RNA-seq data from SLB-challenged leaves to identify

Fig. 1 | ZmKSL2 is an *ent*-isokaurene synthase required for kauralexin production. a, Heat map depicting maize leaf RNA-seq expression profiles (log_2[fragments per kilobase of transcript per million mapped reads (FPKM) + 0.1] of diTPS genes in response to H_2O controls (Con) and *C. heterostrophus* (*C.h.*) inoculation treatments after 24 h. **b**, Heat map depicting the correlation of co-expression of 11 maize diTPS genes present in a dataset of 2094 RNA-seq samples. Low numbers in squares indicate supportive MR scores (<250). **c**, qrtPCR fold change in *ZmKSL2* transcript levels from stems of 35-day-old Mo17 plants, which were damaged and treated with either 100 µl of H_2O (Dam) or spore suspensions (1×10^7 ml^{-1}) of *Fusarium verticillioides* (*F.v.*), *Aspergillus parasiticus* nor-1 (*A.p.*), *C.h.* or *Rhizopus microsporus* (*R.m.*), and collected 48 h later. Intact stems were used for controls (Con). Error bars in the bar chart indicate mean ± s.e.m. ($n = 4$ biologically independent replicates). Within plots, different letters (a–c) represent significant differences (one-way ANOVA followed by Tukey's test corrections for multiple comparisons; $P < 0.05$). **d**, GC–MS select ion chromatograms (m/z 272) of the major (~95%) and minor (~5%) products present in *Agrobacterium*-mediated transient *Nicotiana benthamiana* co-expression assays using the class II diTPS *ent*-copalyl diphosphate synthase ZmAN2 paired with the class I diTPS ZmKSL2. Included are corresponding *ent*-isokaurene and *ent*-kaurene standards. Four independent experiments were performed and showed similar results. **e**, Gene structure of *ZmKSL2* with gRNA sequence designed to target the third exon. Nucleotides in red represent target sites, green nucleotides indicate protospacer adjacent motif sequences for the gRNA and +1 bp represents a site-specific insertion mutation from selected T0 plants. **f**, Representative GC–MS extracted ion chromatograms (m/z signals: red 303 A-series kauralexins (KAs) and black 301 B-series kauralexins (KBs)) from stem tissues of B73 (wild type) and *Zmksl2* elicited with heat-killed *Fusarium venenatum* hyphae and collected at 3 d. Four biological repeats were performed and showed similar results. **g**, Quantity of total kauralexins from B73 (wild-type) and *Zmksl2* plants treated with elicitor. The asterisk denotes significant differences ($P < 0.05$) using Student t-tests (two-tailed distribution, unpaired). Error bars in the bar chart indicate mean ± s.e.m. ($n = 4$ biologically independent replicates). **h**, Pathway for *ent*-isokaurene biosynthesis from GGPP via *ent*-CPP in maize. ZmAN2 catalyses the cyclization of GGPP to *ent*-CPP and ZmKSL2 sequentially converts *ent*-CPP to the major (95%) and minor (5%) products *ent*-isokaurene and *ent*-kaurene, respectively.

genes displaying significant coregulation. Of 344 maize P450s, we identified 55 transcripts that were significantly up-regulated more than two-fold (log_2) following fungal challenge (Supplementary Table 3). To prioritize candidates for biochemical characterization, MR analyses of *ZmKSL2* with the top 15 co-expressed P450s and the entire transcriptome revealed the greatest degree of



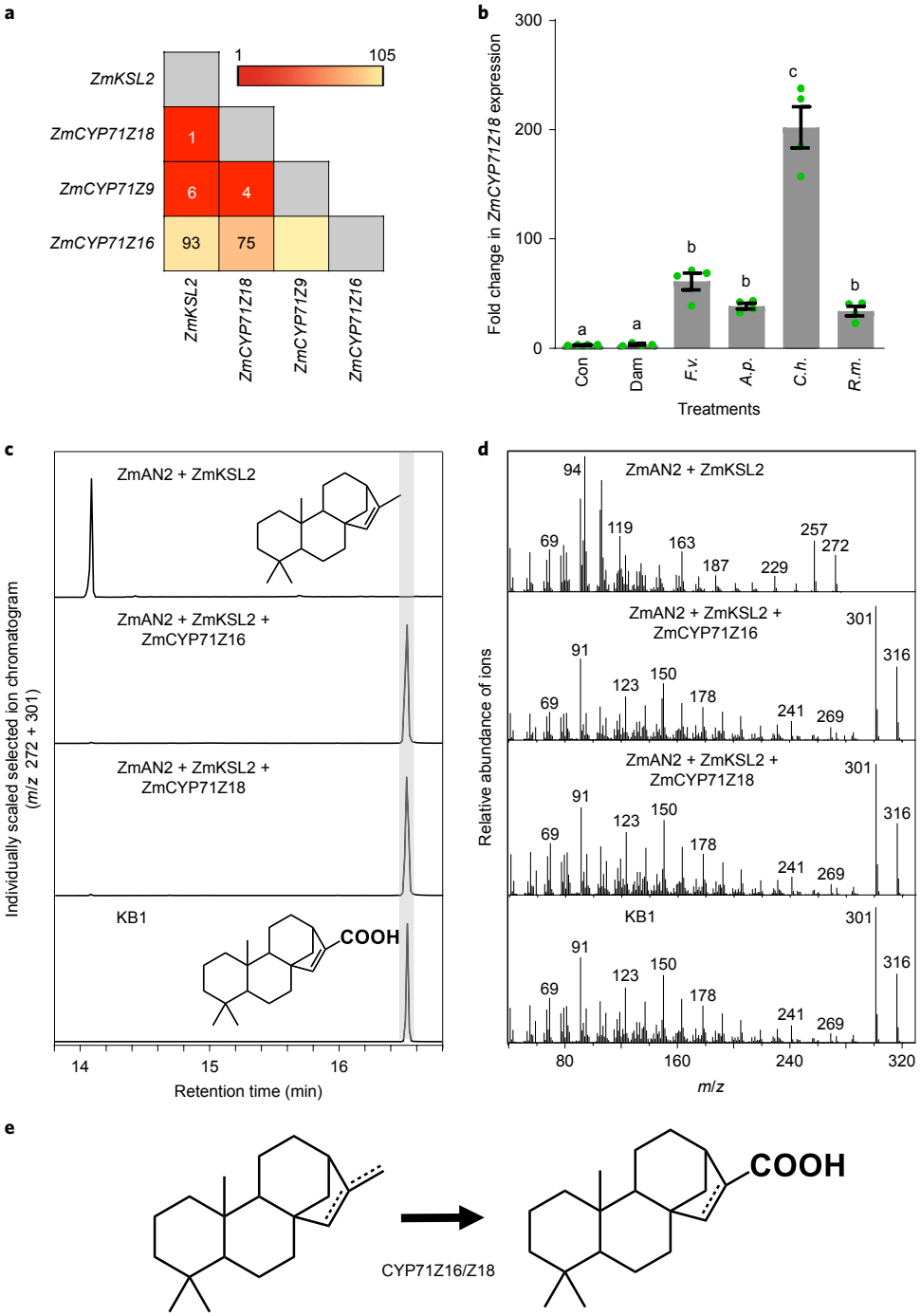


Fig. 2 | Two cytochrome P450s, ZmCYP71Z16 and ZmCYP71Z18, catalyse the production of kauralexins from *ent*-isokaurene and *ent*-kaurene. a, Heat map depicting the co-expression of *ZmKSL2* with all *ZmCYP71* family genes that contain low MR (<100) scores and are supportive of co-regulation in a dataset of 2,094 RNA-seq samples. **b**, qrtPCR fold change of *ZmCYP71Z18* transcript levels present in stems of 35-day-old Mo17 plants, which were damaged and treated with either 100 µl of H_2O (Dam) or spore suspensions (1×10^7 ml^{-1}) of *F.v.*, *A.p.*, *C.h.* or *R.m.*, and collected 48 h later for analyses. Intact stems were used for controls (Con). Error bars in the bar chart indicate mean ± s.e.m. ($n=4$ biologically independent replicates). Within plots, different letters (a–c) represent significant differences (one-way ANOVA followed by Tukey's test corrections for multiple comparisons; $P<0.05$). **c**,**d**, Select ion chromatograms (**c**) and corresponding EI-mass spectra (**d**) of hexane extracts derived from *Agrobacterium*-mediated transient *N. benthamiana* co-expression assays of ZmAN2, ZmKSL2 and ZmCYP71Z16 or ZmCYP71Z18 compared to the authentic kauralexin B1. Four independent experiments were performed and showed similar results. **e**, Schematic representation of the reactions catalysed by ZmCYP71Z16/Z18 in kauralexin biosynthesis.

coregulation with *ZmCYP71Z18* (Fig. 2a and Supplementary Tables 2 and 4). Quantitative real-time PCR (qrtPCR) transcript analysis of *ZmCYP71Z18* coupled with restriction digests to discriminate the closely related tandem duplicate gene *ZmCYP71Z16* (Supplementary Fig. 6) confirmed a broadly fungal-inducible transcript accumulation pattern consistent with *ZmAN2* and *ZmKSL2* (Figs. 1c and 2b and Supplementary Fig. 2).

To enable a comprehensive functional analysis of ZmCYP71Z18, a full-length B73 clone and a synthetic N-terminally modified and codon-optimized gene were generated for co-expression with

ZmAN2 and *ZmKSL2* using *N. benthamiana* and *E. coli*, respectively. Both in planta and microbial co-expression assays resulted in the oxidation of *ent*-isokaurene to kauralexin B1 (KB1) (Fig. 2c–e and Supplementary Figs. 7–9). In a similar combination, pairwise activity of ZmAN2 and the *ent*-KS ZmKSL3 in *N. benthamiana* and *E. coli* co-expression assays predominantly yielded kauralexin A1 (KA1) (Supplementary Figs. 7, 9). Consistent with a recent gene duplication event, *ZmCYP71Z16* occurs in a tandem array with *ZmCYP71Z18* on chromosome 5 and shares an 89% amino acid sequence identity[23]. Both P450s catalyse the oxidation of dolabradiene to 3β-hydroxy-15,16-epoxy-dolabradiene in co-expression assays with ZmAN2 and ZmKSL4 (ref. [23]). To examine potential promiscuous catalytic activity of ZmCYP71Z16 on *ent*-isokaurene and *ent*-kaurene, we performed enzyme co-expression studies with ZmAN2 and either ZmKSL2 or ZmKSL3 in both *N. benthamiana* and *E. coli* and detected respective KB1 and KA1 product formation (Fig. 2c and Supplementary Figs. 7, 9). In contrast, when co-expressing the related P450 *ZmCYP71Z9* (Fig. 2a and Supplementary Table 4) with ZmAN2 and ZmKSL2, no KB1 production was observed (Supplementary Fig. 8). Although differences in substrate affinity and pathway involvement of ZmCYP71Z16 and ZmCYP71Z18 in planta cannot be excluded, transcript co-expression analyses and heterologous enzyme co-expression assays are consistent with redundant catalytic functions as both P450 enzymes are capable of oxidizing diterpene olefins to dolabralexins and kauralexins A1/B1. ZmCYP71Z18 is part of a narrow phylogenic subclade of related sorghum and rice P450s, which includes OsCYP71Z6 and Z7 (Supplementary Fig. 10) involved in the production of rice diterpenoid defences[19,36]. Although all examined species in the genus *Zea* produce kauralexins (Supplementary Fig. 11), Poaceous P450s capable of the C-17 sequential oxidation of *ent*-(iso)kaurene to carboxylic acids have not been characterized outside maize and have probably evolved after the divergence from rice and sorghum, 70 and 26 million years ago, respectively[37].

Kaurene oxidase2, recruited from GA biosynthetic enzymes, drives kauralexin biosynthesis. To identify the endogenous P450(s) responsible for the C-19 oxygenation of kauralexins, we used the Goodman diversity panel for metabolite-based GWAS[38] and, specifically, the ratio of C-19-oxygenated to total kauralexins as a phenotypic mapping trait. Statistically significant single nucleotide polymorphisms (SNPs) were identified on chromosome 9 and spanned the genetic location of two *ent*-kaurene oxidases (KO) (Fig. 3a,b, Supplementary Fig. 5 and Supplementary Table 5). Known KOs in the CYP701 family catalyse multiple oxidations of the *ent*-kaurene C4α methyl group to yield *ent*-kaurenoic acid for GA biosynthesis[39]. Furthermore, KO enzymes have been shown to exhibit remarkable substrate promiscuity. For example, the single *Arabidopsis* KO (At5g25900) is capable of using over 20 different substrates in vitro[28]. As with gibberellins, a classical KO-like oxidation at the C-19 position is predictably required for producing highly oxidized kauralexins[24]. Phylogenic analyses (Supplementary Fig. 13) support the presence of two predicted maize KO homologues, namely ZmKO1 (ZmCYP701A26) and ZmKO2 (ZmCYP701A43), which share 81% amino acid sequence identity[27]. Consistent with a comparatively recent duplication event 3 million years ago, *ZmKO1* and *ZmKO2* are tandemly arrayed, each containing eight exons and seven introns (Supplementary Table 2)[40]. ZmKO1 exhibits prototypical KO activity[27], substrate promiscuity and parsimonious GA biosynthetic pathway assignment due to constitutive gene expression in growing tissues (Fig. 3c). In contrast, *ZmKO2* exhibits broadly inducible transcript accumulation in response to fungal infection (Fig. 3d), yet it has not been biochemically proven to use *ent*-kaurene or *ent*-isokaurene as substrates[27].

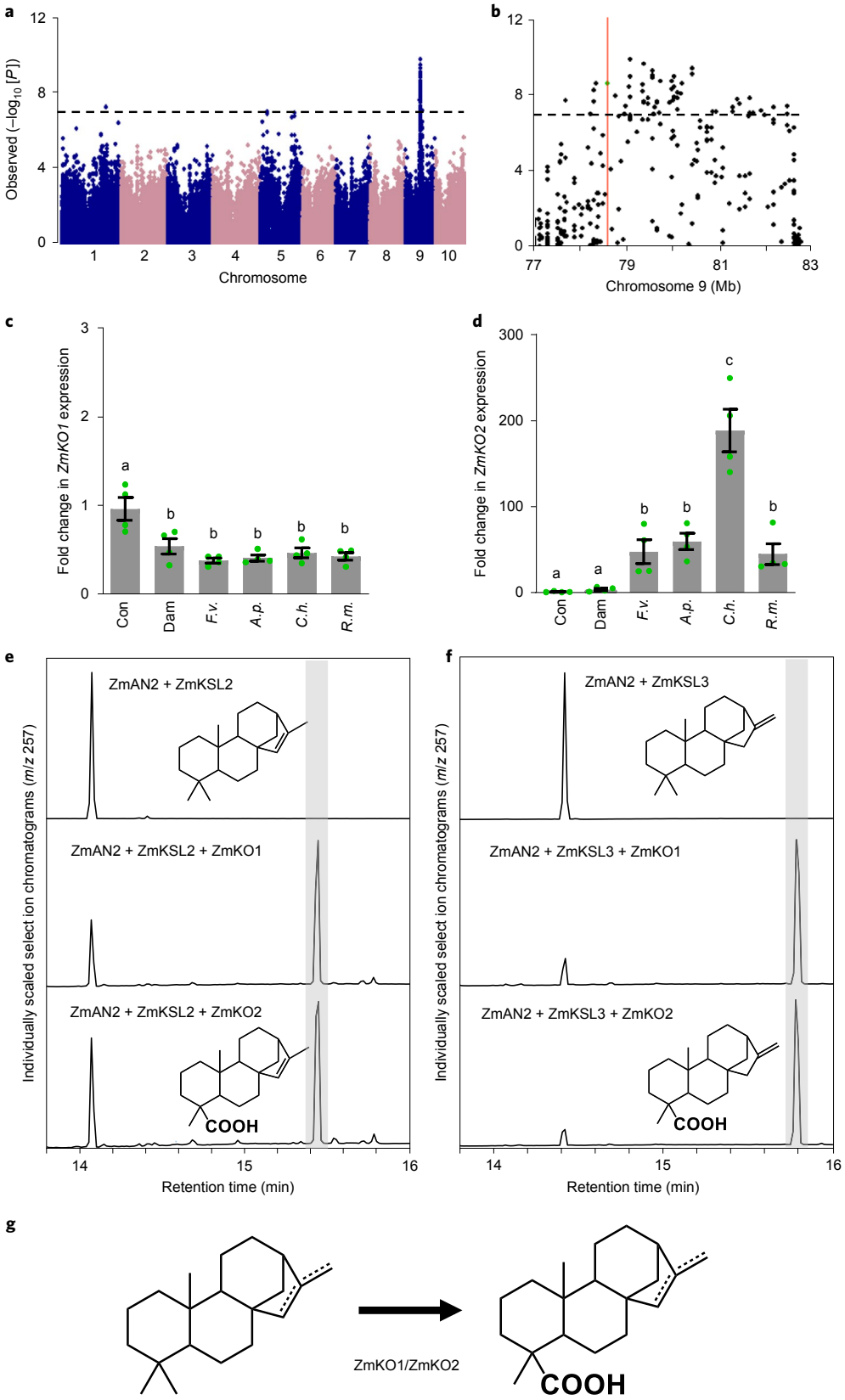
To evaluate function, transient co-expression of ZmAN2 and ZmKSL3 with either ZmKO1 or ZmKO2 in *N. benthamiana* resulted in the formation of the GA precursor *ent*-kaurenoic acid, confirming that both KO enzymes retain prototypical KO activity (Fig. 3e,f and Supplementary Fig. 14). To examine activity on *ent*-isokaurene, ZmKO1/2 enzymes were individually co-expressed with ZmAN2 and ZmKSL2 and yielded the endogenous maize metabolite *ent*-isokaurenoic acid (Fig. 3e,f and Supplementary Figs. 2 and 14). *E. coli* expression of codon-optimized and N-terminally modified ZmKO1 or ZmKO2 combined with ZmAN2 and ZmKSL3 or ZmKSL2 produced identical results (Supplementary Fig. 15). Replicated genetic and biochemical data support the shared C-19 oxidation of *ent*-kaurene and *ent*-isokaurene by ZmKO1/2; however, only *ZmKO2* transcripts are co-expressed with kauralexin genes (Supplementary Fig. 16) and accumulate following pathogen attack (Fig. 3c,d).

During the formation of *ent*-kaurenoic acid (synonym *ent*-kauren-19-oic acid), KO catalyses the sequential production of alcohol and aldehyde intermediates[41]. To identify additional ZmKO2 pathway products in planta, large-scale extraction, chromatographic separations, gas chromatography–mass spectrometry (GC–MS) and nuclear magnetic resonance (NMR) analyses were performed on field-grown maize tissues and confirmed the existence of *ent*-kaur-19-ol-17-oic acid and *ent*-kaur-15-en-19-ol-17-oic acid, termed kauralexin A4 (KA4) and B4 (KB4), respectively (Supplementary Fig. 17 and Supplementary Table 6), as predicted KO-derived intermediates to kauralexin aldehydes (KA3/KB3) and diacids (KA2/KB2). To examine the endogenous role of ZmKO2 in kauralexin biosynthesis, a maize *Mu* transposon insertion (mu1068966) impacting the first intron was examined (Supplementary Fig. 16). In comparison to wild-type W22 parents, both fungal-elicited *ZmKO2* transcript levels and, specifically, C-19-oxygenated kauralexins were significantly reduced in the *Zmko2* mutant (Supplementary Fig. 16). Collectively, our results support KO gene duplication and functionalization of ZmKO2 for defence.

A steroid 5α reductase enables partitioning of kauralexin and gibberellin biosynthesis. Of all the kauralexins, the A-series kauralexins are typically the dominant accumulated end products in most maize inbred lines (Supplementary Fig. 18). This pattern and the enzyme co-expression results are consistent with the hypothesis

Fig. 3 | Kaurene oxidase-like 2 (ZmKO2) catalyses the synthesis of C-19 oxygenated kauralexins. a, Manhattan plot of the Goodman diversity panel association analysis (258 inbred lines) using ratio of fungal-elicited C-19-oxygenated kauralexins to total kauralexins as mapping trait. Negative $\log_{10}$-transformed *P* values from the compressed mixed linear model are plotted on the *y* axis. The dashed line denotes the 5% Bonferroni-corrected threshold for 246,477 SNP markers with the most statistically significant SNP located at position 79,119,114 (B73 RefGen_v2) on chromosome 9. **b**, Local Manhattan plot surrounding the peak on chromosome 9. **c**,**d**, qrtPCR fold changes of *ZmKO1* (**c**) and *ZmKO2* (**d**) transcript levels present in untreated control stems (Con) or those damaged and treated with either 100 µl of water alone (Dam) or spore suspensions (1×10^7 ml^{-1}) of *F.v.*, *A.p.*, *C.h.* or *R.m.* Error bars in the bar charts indicate mean ± s.e.m. ($n=4$ biologically independent replicates indicated by green circle plot symbols). Within plots, different letters (a–c) represent significant differences (one-way ANOVA followed by Tukey's test corrections for multiple comparisons; $P<0.05$). **e**, GC–MS selected ion chromatograms of hexane extracts derived from *Agrobacterium*-mediated transient *N. benthamiana* co-expression assays of ZmAN2, ZmKSL2 and ZmKO1 or ZmKO2 yield *ent*-isokaurenoic acid. Four independent experiments were performed and showed similar results. **f**, GC–MS selected ion chromatograms traces of identical DiTPS/KO combinations, with the exception of ZmKSL3 substituted for ZmKSL2 yield *ent*-kaurenoic acid. Four independent experiments were performed and showed similar results. **g**, Schematic representation of the reactions catalysed by ZmKO1/ZmKO2.

that *ent*-kaurene can substantially contribute to kauralexin biosynthesis[22,42]. As a logical extension, large-scale pathogen-elicited production of direct GA biosynthetic precursors, such as *ent*-kaurene, occurs while the GA-related transcripts display only occasional and modest decreases (Fig. 3c, Supplementary Fig. 2 and Supplementary Table 2). An over-abundance of *ent*-kaurene could contribute to a physiological 'Achilles' heel' by promoting dysregulated GA precursor levels and pathway production[43]. Opposing results using enzyme



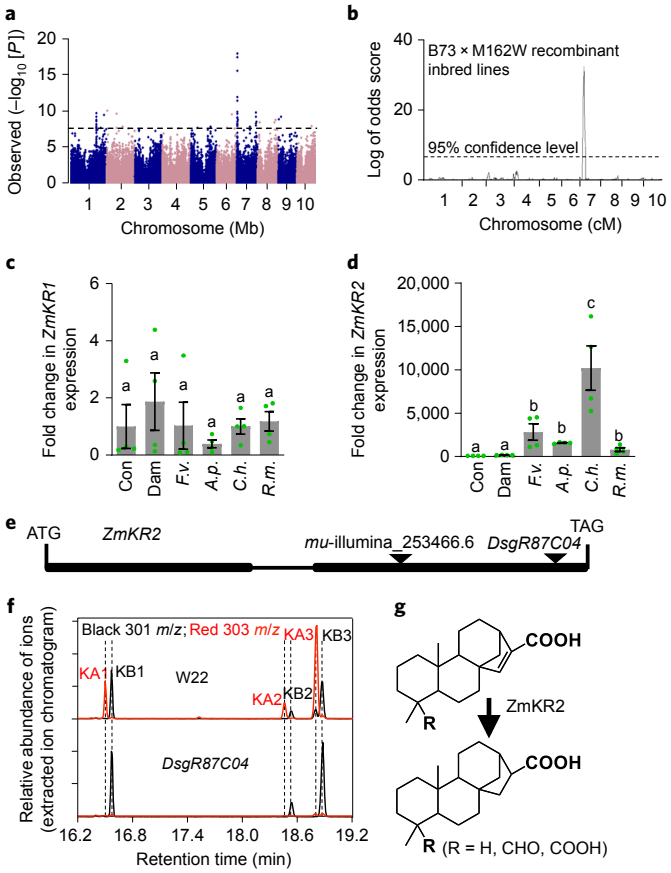


Fig. 4 | Kauralexin reductase 2 (ZmKR2) is a steroid 5α-reductase family enzyme required for the indirect production of *ent*-kaurene associated defences. a, Manhattan plot of the Goodman diversity panel association analysis (258 inbred lines) using the ratio of fungal-elicited total A- to B-series kauralexins as a mapping trait. Negative $\log_{10}$-transformed *P* values from the compressed mixed linear model are plotted on the *y* axis. The dashed line denotes the 5% Bonferroni-corrected threshold for 246,477 SNP markers, with the most statistically significant SNP located at position 7,279,886 (B73 RefGen_v2) on chromosome 7. **b**, Linkage analysis of the ratio of the fungal-elicited total A-series to B-series kauralexins in a B73 × M162W recombinant inbred line mapping population. Logarithm of the odds score profile with the permutation threshold is indicated by the dashed horizontal line. **c**, qrtPCR fold changes of *ZmKR1* (Zm00001d018846). **d**, *ZmKR2* (Zm00001d018847) transcript levels present in untreated control stems (Con) or those damaged and treated with either 100 µl of water alone (Dam) or spore suspensions (1×10^7 ml^{-1}) of *F.v.*, *A.p.*, *C.h.* or *R.m.*. Error bars in **c** and **d** indicate mean ± s.e.m. ($n = 4$ biologically independent replicates indicated by green circle plot symbols). Within plots, different letters (a–c) represent significant differences (one-way ANOVA followed by Tukey's test corrections for multiple comparisons; $P < 0.05$). **e**, Schematic of transposon insertions (*dsg* and *Mu* elements) in the second exon of the *ZmKR2* gene. **f**, GC–MS extracted ion chromatogram (*m/z* 301, B-series kauralexins, *m/z* 303, A-series kauralexins) of extracts from fungal elicitor-challenged stems of W22 and the *DsgR87C04* mutant after 72 h. Four biological repeats were performed and showed similar results. **g**, Schematic representation of ZmKR2 catalysed reactions.

co-expression assays demonstrate that *ent*-kaurene is a minor (<5%) ZmKSL2 byproduct and *Zmksl2* mutants display dramatic reductions in pathogen-elicited kauralexins (Fig. 1f,g). Based on the absence of A-series kauralexins in select inbreds (Supplementary Fig. 18), we hypothesized that *ent*-kaurene-derived kauralexins are actually products of a subsequent pathway step.

In a forward genetics approach, we performed a replicated GWAS of fungal-elicited tissues using the ratio of A- to B-series kauralexins as a trait to identify highly significant SNPs on chromosome 7 (Fig. 4a and Supplementary Fig. 18). Similar analyses in a B73 × M162W biparental mapping population further confirmed the presence of a single significant quantitative trait loci (QTL) on chromosome 7 (Fig. 4b) independent of total kauralexin abundance (Supplementary Fig. 18). Within the shared genetic mapping interval (Supplementary Table 7), examination of the maize genome revealed two candidate kauralexin reductase (*KR*) genes, GRMZM2G394968 (*ZmKR1*) and GRMZM2G073929 (*ZmKR2*), 58.3 kb apart and annotated as 3-oxo-5-α-steroid 4-dehydrogenase (that is, steroid 5α-reductase: SRD5α) family proteins, which share 74% sequence identity at the amino acid level. In plants and mammals, the 5α-reductase family functions in steroid hormone metabolism and contains three subfamilies, namely SRD5α type 1/2, SRD5α type 3 and Glycoprotein Synaptic 2, all of which are predicted products of gene duplication events in early eukaryotic ancestors[44]. In plants the SRD5α, termed de-etiolated 2 (DET2), catalyses the reduction of campesterol to campestanol, thereby enabling brassinosteroid biosynthesis[45]. Phylogenetic analyses placed ZmDET2 in the SRD5α type 1/2 subfamily, while the human testosterone 5α-reductase (SRD5A3) and ZmKR1/2 are members of the SRD5α type 3 subfamily (Supplementary Fig. 19)[46].

Unlike *ZmKR1*, *ZmKR2* transcripts exhibit inducible accumulation in response to fungal pathogens and display strong co-expression with kauralexin biosynthetic genes (Fig. 4c,d Supplementary Table 2 and Supplementary Fig. 20). To examine the function of *ZmKR2* in planta, we isolated a *Dsg* transposable element insertion mutant impacting second exon (Fig. 4e)[47]. Elicitor-induced accumulation of both *ZmKR2* transcripts and A-series kauralexins were significantly decreased in *Zmkr2* (Fig. 4f,g and Supplementary Fig. 20).

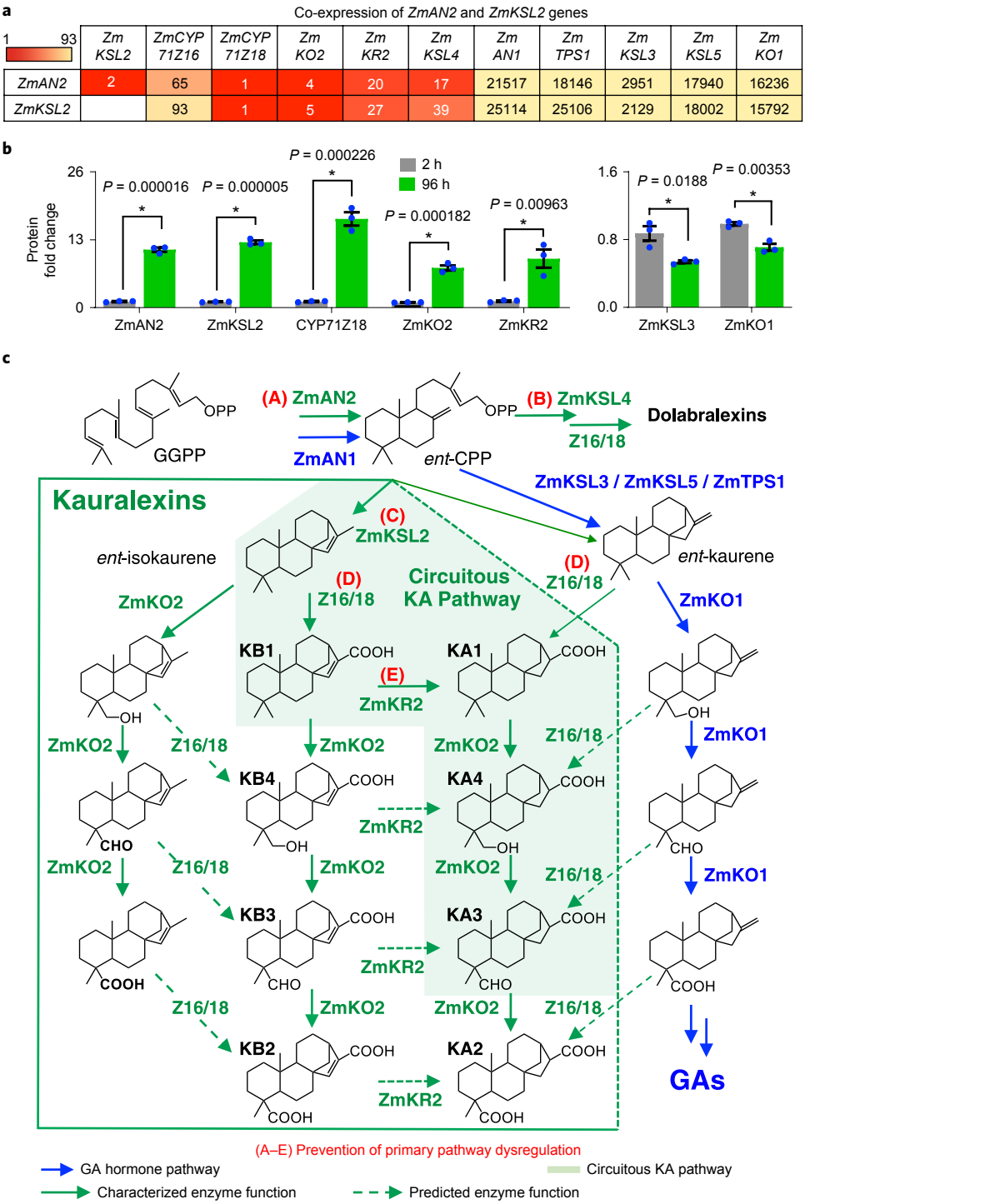


Fig. 5 | Coregulation, positional isomer specificity, enzyme promiscuity and a 5α-steroid reductase family enzyme partition growth and defence-related maize diterpenoid pathways. a, Heat map illustrating the MR correlation of co-expression of *ZmAN2* and *ZmKSL2* genes with diterpenoid genes involved in defence and the GA pathway derived from 2,094 RNA-seq samples. **b**, LC–MS quantification of fungal elicitor-induced fold changes in kauralexin and GA pathway proteins using the W22 inbred. Error bars in charts indicate mean ± s.e.m. ($n = 3$ biologically independent replicates). The asterisk denotes significant differences using Student *t*-tests (two-tailed distribution, unpaired), with $P < 0.05$. **c**, Pathway integration and partitioning of parallel diterpenoid defence (green arrows) and hormone (blue arrows) pathways in maize. Solid line arrows represent enzyme catalysis supported by genetics and/or heterologous co-expression studies. Dashed line arrows indicate predicted pathways enabled by enzyme promiscuity that require in vitro substrate feeding studies to empirically demonstrate. KA3 is the predominant accumulated end product in most maize inbreds (Supplementary Fig. 18) and ZmKR2 enables a circuitous KA pathway. Capital letters (red) denote mechanisms to reduce parallel pathway dysregulation: A, inducible co-expression of the entire defence pathway; B, multiple defence pathways to consume shared GA precursors; C, production of *ent*-isokaurene minimizes production of GA pathway precursors; D, promiscuous P450s remove traces of *ent*-kaurene produced by the defence enzyme ZmKSL2; E, ZmKR2 indirectly yields *ent*-kaurene associated defences while avoiding the direct production of *ent*-kaurene.

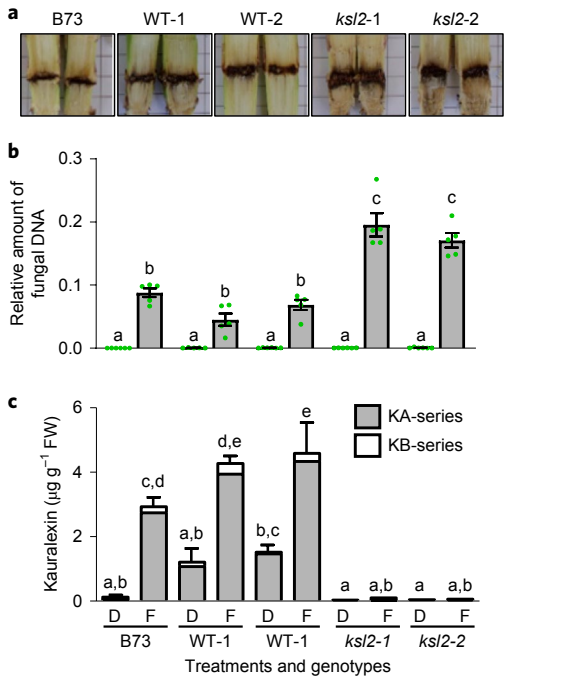


Fig. 6 | Kauralexins are required to suppress *F. graminearum* stalk rot. **a**, Representative disease levels in stems of B73, two wild-type (WT) siblings and two *Zmksl2* mutant lines inoculated with *F. graminearum* (10 μl of 1.5×10^5 conidia ml^{-1}) after 10 d. Ten biological repeats were performed and showed similar results. **b**, Ratio of *F. graminearum* DNA to maize DNA levels. **c**, Total A-series (grey) and B-series (white) kauralexins present in corresponding maize stems damaged and treated with either H_2O (D) or *F. graminearum* (F) from the *F. graminearum* resistance assay. Error bars in charts **b** and **c** indicate mean ± s.e.m. ($n = 5$ biologically independent replicates indicated by green circle plot symbols). Within plots, different letters (a–e) represent significant differences (one-way ANOVA followed by Tukey's test corrections for multiple comparisons; $P < 0.05$).

A second insertion mutant from the Illumina *Mu* population (253466.6) produced identical biochemical results (Supplementary Fig. 20). To demonstrate pathway activity, combinatorial transient expression of all five enzymes, namely ZmAN2, ZmKSL2, ZmCYP71Z18, ZmKO2 and ZmKR2, in *N. benthamiana* yielded the combined formation of KB1-3 and KA1-3 (Supplementary Fig. 21). Thus, ZmKR2 enables maize plants to predominantly accumulate A-series kauralexins (Supplementary Fig. 18) using a pathway that is independent from the growth hormone precursor *ent*-kaurene.

Separation of the maize kauralexin pathway. Combined MR co-expression analyses of all characterized genes mediating *ent*-copalyl-derived maize diterpenoids including early predicted GA metabolic genes support the kauralexin pathway and provide powerful predictive relationships (Fig. 5a). To examine endogenous relationships at the level of mRNA translation products we performed mass spectrometry-based proteomics on *Fusarium*-elicited maize stem tissue. Statistically significant seven-fold, or greater, increases in ZmAN2, ZmKSL2, ZmCYP71Z18, ZmKO2 and ZmKR2 occurred between early (2 h) and late (96 h) sampling times (Fig. 5b). In contrast, levels of established or predicted GA pathway proteins were either undetectable or moderately reduced (Fig. 5c and Supplementary Table 8). Targeted metabolomics, transcriptomics, heterologous enzyme co-expression, genetics and proteomics collectively support the fungal-induced role of ZmAN2, ZmKSL2, ZmCYP71Z16/18, ZmKO2 and ZmKR2 in kauralexin biosynthesis (Fig. 5c).

Maize fungal resistance is the major predicted biological function of kauralexins and this hypothesis is partially supported by increased fungal susceptibility in *Zman2* mutants[21]. However, due to the complexity of branched ZmKSL2 and ZmKSL4 pathways, both of which require ZmAN2 products, the specific protective role of kauralexins in planta has not been previously examined[21,23]. Using *Zmksl2* mutants, we examined alterations in the *Fusarium graminearum* pathosystem where elicited *ZmAN2* transcript accumulation was first observed[29]. Ten days after stalk inoculation with *F. graminearum*, *Zmksl2* mutant plants displayed visible and quantitative increases in disease susceptibility as estimated by ratios of *F. graminearum* to maize DNA (Fig. 6a,b). Unlike the uniform elicitation of sesquiterpenoid zealexins (Supplementary Fig. 22), *Zmksl2* mutants consistently displayed a lack of inducible kauralexin accumulation, which illustrates a specific endogenous role in pathogen protection (Fig. 6).

Discussion

Stress protection in plants is facilitated by dynamic networks of specialized metabolites that serve diverse roles in species interactions and ecological adaptation. Given the global footprint of Poaceous crops on arable lands, the precise chemical languages employed by key species are increasingly sought after for control and trait improvement[19,48,49]. Decades of extensive studies in rice have revealed an array of specialized diterpenoids with potent antimicrobial and allelopathic functions that mechanistically contribute to stable resistance to major fungal diseases, such as rice blast (*Magnaporthe oryzae*)[19,50,51]. In maize, related yet distinct diterpenoids have been implicated in diverse protective roles providing fungal, insect and drought resistance[21,23–25]. The present effort highlights the power of combining multi-omics, GWAS and enzyme co-expression to define specialized metabolic pathways in the genus *Zea* (Fig. 5c and Supplementary Fig. 11).

Approximately 12 million years ago, two maize progenitors created an allopolyploidization event, resulting in the genome-wide duplication of gene copies which was eventually followed by the loss of many homoeologous gene pairs during a return to the diploid state[52,53]. Tandem duplicate genes display common associations with transposable elements, continue to arise through non-homologous recombination and account for least 10% of all annotated maize genes[40]. Widespread gene duplications in maize provide a rich foundation to facilitate functional radiation of primary metabolic enzymes to expanded specialized metabolism[6,53]. As a common evolutionary mechanism in plant diterpenoid metabolism, GA biosynthetic genes are repeatedly recruited for gene duplication and diversification, thereby enabling vast chemical diversity of terpenoid defences[17]. In addition to kauralexin pathway genes, related metabolic expansion is exemplified by the multi-gene diTPS and P450 families forming modular diterpenoid-metabolic pathways in rice, wheat (*Triticum aestivum*), switchgrass (*Panicum virgatum*) and others[8,19,23,54]. Beyond basic paradigms of gene family expansion and divergence, precisely how a series of closely related and catalytically conserved phytohormone-derived pathways are functionally partitioned to avoid negatively impacting primary metabolic processes is thus far unpredictable. In maize, three kauralexin biosynthetic enzymes, namely ZmAN2, ZmKSL2 and ZmKO2, display phylogenetic relationships consistent with duplication and functional divergence from core diTPS and P450 genes involved in phytohormone pathways (Supplementary Figs. 3, 13). ZmAN2 retains the same enzymatic function as ZmAN1, but underwent specialization towards regulation with patterns of strict stress-inducible gene expression (Supplementary Fig. 2)[29]. Similarly, the magnitude of stress-elicited transcript accumulation discriminates ZmKSL2 from the GA pathway enzyme ZmKSL3 (Fig. 1c, Supplementary Fig. 2 and Supplementary Table 2). In addition, ZmKSL2 displays catalytic specificity distinct from the *ent*-KS ZmKSL3 by predominantly

forming *ent*-isokaurene with only minimal *ent*-kaurene synthase activity (Fig. 1d–f). Formation of specialized diterpenoids via an *ent*-isokaurene pathway node is also a major pathway branch in rice catalysed by the pairwise activity of *ent*-CPS (OsCPS1/2) and the *ent*-isokaurene synthase (OsKSL6)[36]. Less than 50% protein-level identity between OsKSL6 and ZmKSL2 suggests that the formation of *ent*-isokaurene as a branch point towards specialized metabolites evolved independently in rice and maize rather than being derived from a shared *ent*-isokaurene synthase progenitor. In contrast to the diTPS-catalysed formation of olefin precursors, KO of the CYP701 family seem to have been recruited for gene duplication and functionalization to a lesser extent. For example, KO in *Physcomitrella patens* (CYP701B1) and *Arabiopsis thaliana* (CYP701A3) remain as single copy genes required for GA biosynthesis[55,56]; however, despite serving targeted roles the enzymes also display remarkable substrate promiscuity[28]. Similar to the expanded family of rice CYP701A KO with distinct functions in GA and specialized metabolism, we demonstrate that maize ZmKO1 and ZmKO2 share a nearly exact chromosome location, gene structure and enzyme activity. However, the comparatively recent tandem gene duplications (Supplementary Table 2) can be discriminated by differential expression, thus offering alternate regulation of KO pathway steps (Fig. 3c–f). In addition to three enzymes recruited from the GA pathway, a small family of steroid 5α-reductases conserved in both plants and animals for hormone biosynthesis contributed to the origin and function of ZmKR2 in kauralexin biosynthesis[44–46]. Verified by multiple defined mutants, ZmKR2 catalysis in planta enables the circuitous production of *ent*-kaurene-associated kauralexins.

Although the duplication and functionalization of genes involved in primary pathways underlies plant specialized metabolism[6–9], there must be layered fail-safe mechanisms to minimize the interference of functionally distinct parallel pathways. For example, in *Fusarium*-elicited maize tissues, kauralexin concentrations (Supplementary Fig. 18) can average 100,000-fold higher than total bioactive GAs quantified in maize-*Fusarium* assays[57]. The kauralexin pathway reveals multiple interactive mechanisms (A–E) to reduce the constitutive presence of hormone precursors (Fig. 5c). (A) All kauralexin pathway genes display low expression levels in healthy tissues and are highly co-expressed during pathogen challenge (Figs. 1c, 2b, 3d and 4d and Supplementary Table 2). (B) Additional class I diTPS, such as ZmKSL4, provide an additional pathway branch that uses ZmAN2-derived *ent*-copalyl diphosphate (*ent*-CPP) precursors for pathogen-elicited defensive production (Fig. 5a,c)[23]. (C) Kauralexins are formed predominantly from the ZmKSL2 product *ent*-isokaurene, a positional isomer not used for GA biosynthesis[58], thereby avoiding the need for direct *ent*-kaurene production. (D) Independent to the GA pathway, highly promiscuous activities of ZmCYP71Z16/18 enable the committed production of kauralexins from both *ent*-isokaurene and *ent*-kaurene. Thus, in response to pathogen challenge, ZmCYP71Z16/18 can facilitate conversion of *ent*-kaurene derived from ZmKSL2 and other *ent*-kaurene synthases into antibiotic defences.

In a large-scale effort to understand mechanisms of pathogen resistance, statistical modelling previously ranked the *ent*-kaurene-associated metabolite KA3 as being highly associated with fungal protection in diverse commercial hybrids[21]. A-series kauralexins, and specifically KA3, accumulate as predominant end products in most inbreds and underlie *Fusarium* resistance (Fig. 6c and Supplementary Fig. 18). In the final mechanism, (E), ZmKR2 enables the production of A-series kauralexins from *ent*-isokaurene derived precursors to yield *ent*-kaurene-associated A-series kauralexins indirectly. The term *ent*-kaurene-associated kauralexins refers to the combined promiscuous activities of ZmCYP71Z16/18 and ZmKR2 that typically mask the specific molecular-level source of A-series kauralexins. Collectively, significant activation of the kauralexin pathway at the transcriptome and proteome levels is paired with comparatively modest signatures of GA pathway suppression (Figs. 3c and 5b, Supplementary Fig. 2 and Supplementary Tables 2 and 8) and supports additive layers (A–E), thereby aiding in the prevention of GA pathway dysregulation (Fig. 5c). More generalized mechanisms could also play a role in separating defence and hormone pathways. For example, although the 2-*C*-methyl-D-erythritol 4-phosphate (MEP) pathway is generally accepted as rate-limiting in diterpenoid biosynthesis, the position of diTPSs at metabolic branching points suggests that kinetic differences in diTPS activity may also contribute to pathway partitioning[59]. Furthermore, substrate allocation to distinct pathways might be guided by specific diTPS and P450 interactions, either by means of functional metabolons, as demonstrated in sorghum for dhurrin biosynthesis, or via close physical proximity of plastidial diTPS and endoplasmic reticulum-localized P450s through membrane hemifusion, as suggested for GA biosynthesis[60,61].

The duplication of hormone pathway genes to fuel plant defence metabolite production is common in Poaceous crops and probably plants in general. Using maize as a model and kauralexin biosynthesis as a comprehensive example, we detail how multiple hormone pathways and mechanisms can combine to enable the partitioning of primary and specialized metabolism. Many features of the kauralexin pathway enable the biosynthesis of closely related functional antibiotics while minimizing the unregulated accumulation of GA metabolic precursors that could otherwise promote dysregulated phytohormone signalling commonly targeted by diverse pathogenic microbes[62,63].

Methods

Plant and fungal materials. Seeds of the Goodman diversity panel for GWAS mapping and the Nested Association Mapping (NAM) recombinant inbred lines (RIL) (B73×M162W subpopulation) for linkage analysis were provided by Georg Jander (Boyce Thompson Institute) and Peter Balint-Kurti (US Department of Agriculture—Agricultural Research Service), respectively (Supplementary Table 9). Seeds of NAM parental lines and uniform Mu mutant UFMu-08760 (containing mu1068966) were obtained from the Maize Genetic COOP Stock Center. Seeds of tdsgR87C04 and mu-illumina_253466.6 were provided by Hugo K. Dooner (Waksman Institute of Microbiology) and Alice Barkan (Institute of Molecular Biology)[47,64]. *Zea perennis* (Ames 21874), *Zea diploperennis* (PI 462368), *Zea luxurians* (PI 422162), *Zea m. parviglumis* (PI 384069) and *Zea m. mexicana* (Ames 21851) were acquired from Mark Millard (North Central Regional Plant Introduction Station). Unless otherwise mentioned, maize inbreds (B73, Mo17, W22) used for replicated elicitation experiments were germinated in MetroMix 200 (Sun Gro Horticulture Distribution) supplemented with 14-14-14 Osmocote (Scotts Miracle-Gro) and grown in a greenhouse as previously described[65]. Fungal stock cultures of *R. microsporus* (Northern Regional Research Laboratory, stock no. 54029), *F. verticillioides* (Northern Regional Research Laboratory, stock no. 7415), *F. graminearum* (Northern Regional Research Laboratory, stock no. 31084) *A. parasiticus* (nor-1) and *C. heterostrophus* were grown on V8 agar for 12 d before the quantification and use of spores[66]. Heat-killed *F. venenatum* (strain PTA-2684) hyphae was commercially obtained (Monde Nissin Corporation) and used as a non-infectious elicitor lacking known mycotoxins.

RNA-seq analyses of fungal-elicited defence genes. To examine the induced defence transcriptome activated by a necrotrophic fungal pathogen, 18-day-old soil-grown maize (*Z. mays* var. Golden Queen) seedlings were inoculated with *C. heterostrophus*. Inoculations involved application of six 10 µl H_2O (plus 0.1% Tween 20) droplets (control) or those containing a suspension of SLB spores (1×10^6 ml^{-1}) onto the central adaxial portion of newly emerged leaf 4. The droplets were spaced 15 cm from the leaf tip, with three inoculation sites on each side of the midrib approximately 2 cm apart. After treatment all plants were placed into a 100% humidity chamber, followed by snap freezing of control and inoculated leaves in N_2 after 24 h. For RNA-seq analysis, libraries were constructed using an Illumina TruSeq Stranded RNA LT Kit following manufacturer protocols. Quantity of total RNA was adjusted to 1.3 µg. Deep sequencing was performed using an Illumina NextSeq500 Instrument at the University of Georgia Genomics Facility. For RNA-seq data analysis, raw fastq reads were trimmed for adaptors and preprocessed to remove low quality reads using Trimmomatic v.0.32 (ref. [67]). Qualified reads were then aligned to the *Z. mays* AGPv4.34 reference genome using Bowtie2 v.2.2.3.0 (ref. [68]) and TopHat v.2.0.13 (ref. [69]). Gene expression values were computed using Cufflinks v.2.2.1 (ref. [70]).

MR analyses of coregulated transcripts. RNA-seq datasets from the Goodman diversity panel and the Nested Association Mapping qTeller data

(http://www.qteller.com/NAM4) were derived from the National Center for Biotechnology Information Sequence Read Archive project IDs SRP115041 and SRP011480, respectively[71]. The combined RNA-seq dataset (B73 RefGen_V4) was composed of 300 inbred lines represented by 2,094 samples derived from developmentally diverse tissues. From this dataset, calculations of MR were used as a measure of co-expression by calculating the geometric mean of the product of two-directional ranks derived from Pearson correlation coefficients across all gene pairs[30,72].

Maize stem inoculation with live fungi and heat-killed fungal elicitors. Using a surgical scalpel, 35-day-old plants were slit in the centre, spanning both sides of the stem, to create a 8–10 cm long parallel longitudinal incision. The incision targeted and spanned the upper nodes, internodes and the most basal portion of unexpanded leaves. All fungal spore inoculations (1×10^7 ml^{-1}) of *R. microsporus*, *F. verticillioides*, *A. parasiticus* (nor-1) and *C. heterostrophus* were performed in 100 µl of H_2O. For experiments using heat-killed fungal elicitors, approximately 500 µl of commercial *F. venenatum* hyphae was introduced into each slit stem, which was then sealed with clear plastic packing tape to minimize desiccation of the treated tissues. For the *Fusarium* stalk rot resistance assay, a 1-mm-diameter hole was created through the second above-ground node in the stalk of 35-old-plants. The hole was inoculated with either 10 µl of either H_2O alone or 10 µl of a *F. graminearum* spore suspension (1.5×10^5 ml^{-1}). After 10 d, stems were split open, photographed and collected using a pool of two individual plants for each of the five final collected replicates. Within each experiment, treated maize stem tissues were collected in liquid nitrogen at specific time points as indicated.

Nucleic acid isolation and qrtPCR. Total RNA was isolated with TRIzol (Invitrogen) according to the manufacturer's protocol. First-strand complementary DNA was synthesized with the RETROscript reverse transcriptase kit (Ambion) using random decamer primers. qrtPCR was performed using Power SYBR Green Master mix (Applied Biosystems) and 250 nM primers on a Bio-Rad CFX96TM Real-Time PCR Detection System. Mean cycle threshold values were normalized to *ZmEF-1α* (GenBank accession no. AF136829)[66]. Fold-change calculations were performed using the equation $2^{-\Delta\Delta Ct}$. All qrtPCR primers used in the study are listed in Supplementary Table 10. Primers for confirmation and tracking of maize mutants derived from CRISPR–Cas9, *Dsg* and *Mu*-based disruption are listed (Supplementary Table 11). For determination of the fungal biomass, total genomic DNA was extracted plant stem tissues and subjected to qrtPCR using the *F. graminearum*-specific primers for a deoxynivalenol mycotoxin biosynthetic gene (*FgTri6*) (Supplementary Table 10)[73]. Plant DNA quantification was performed using the maize ribosomal protein L17 gene (*ZmRLP17b*, Zm00001d049815) and specific primers ZmRLP17b_F (5′-CAAAGTCTCGCCACTCCA-3′) and ZmRLP17b_R (5′-CGTCCGTGAGCACGGTA-3′). Relative amounts of fungal DNA were calculated by the $2^{-\Delta\Delta Ct}$ method, normalized to a conserved genomic sequence of *ZmRLP17b* DNA.

Genetic mapping of fungal-elicited defences. The Goodman diversity panel and B73 × M162W NAM subpopulation was planted in summer 2016 at the Biology Field Station, UCSD. To avoid plant response variation due to the differential expression and action of fungal effectors on diverse inbred lines, all mapping experiments used heat-killed *F. venenatum* hyphae and slit-stem treatments. Elicited maize stems were collected after 5 d, frozen in liquid nitrogen, ground to a fine powder and stored at −80 °C for further metabolite analyses.

Metabolite-based GWAS was performed as previously described[65]. GWAS was initially conducted using the Unified Mixed Linear Model in TASSEL 5.0, with final analyses conducted with the R package GAPIT[74–77]. SNPs with less than 20% missing genotype data and minor allele frequencies >5% from both an Illumina 50 K array and a genotyping-by-sequencing strategy were employed in the association analysis[78–80]. To improve association analysis, the kinship matrix (K) was used jointly with population structure (Q)[81]. Manhattan plots were constructed in the R package qqman [82]. Genetic marker data for the recombinant lines (NAM B73 × M162W subpopulation) were downloaded from www.panzea.org and used for composite interval mapping implemented in Windows QTL Cartographer v.2.5 (https://brcwebportal.cos.ncsu.edu/qtlcart/WQTLCart.htm) using previously described settings[65].

Identification and quantification of metabolites. All plant tissue samples were frozen in liquid nitrogen, ground to a fine powder and stored at −80 °C for further analyses. For sample preparation by vapour phase extraction, ~50 mg aliquots were weighed, solvent extracted in a bead homogenizer and derivatized using trimethylsilyldiazomethane as described previously[24,83]. 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 × 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. GC–MS-based quantification of kauralexins and related compounds was based on the use of heneicosanoic acid as an internal standard[24]. 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} helium flow, using a HP5-MS column (30 m, 250 µm internal diameter, 0.25 µm film) with a sample volume of 1 µl under 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 for 3 min. MS data from 90 to 600 mass-to-charge ratio (*m*/*z*) were collected after a 9 min solvent delay. Product identification was conducted using authentic standards and by comparing reference mass spectra with Wiley, National Institute of Standards and Technology and the Adams libraries.

Analysis of the second replication of the Goodman diversity panel examining *Fusarium*-elicited kauralexin levels with liquid chromatography–mass spectrometry (LC–MS) follows from previous descriptions[65]. Briefly, tissue 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 H_2O. Each combined sample consisted of a comiscible acidified solvent mixture of primarily 1-propanol/acetonitrile/ethyl acetate/H_2O in the approximate proportion of 11/39/28/22, which was then centrifuged at 15,000 rpm for 20 min. Approximately 150 µl of the particulate-free supernatant was carefully removed and filtered for LC–MS automated sample analyses using 5-µl injections. The LC consisted of an Agilent 1260 Infinitely Series HP Degasser (G4225A), 1260 binary pump (G1312B) and 1260 autosampler (G1329B). The binary gradient mobile phase consisted of 0.1% (v/v) formic acid in H_2O (solvent A) and 0.1% (v/v) formic acid in methanol (solvent B). Analytical samples were chromatographically separated on a Zorbax Eclipse Plus C18 Rapid Resolution HD column (Agilent; 1.8 mm, 2.1350 mm) using a 0.35 ml min^{-1} flow rate. The mobile phase gradient was: 0 to 2 min, 5% B constant ratio; 3 min, 24% B; 18 min, 98% B; 25 min, 98% B; and 26 min, 5% B for column re-equilibration before the next injection. Eluted analytes underwent electrospray ionization 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}, 379 kPa, 225 °C) and sheath gas (350 °C, 12 l min^{-1}). The transfer inlet capillary was 3,500 V and both MS1 and MS2 heaters were at 100 °C. Negative ionization [M–H]$^-$ mode scans (0.1-atomic mass unit steps, 2.25 cycles s21) from *m*/*z* 100 to 1,000 were acquired. In order of relative retention times and [M–H]$^-$ parent ions, all eight kauralexins were KA1 (18.82 min, *m*/*z* 303), KB1 (18.61 min, 301 *m*/*z* 301), KA2 (15.74 min, *m*/*z* 333), KB2 (14.62 min, *m*/*z* 331), KA3 (16.10 min, *m*/*z* 317), KB3 (15.45 min, *m*/*z* 315), KA4 (15.42 min, *m*/*z* 319) and KB4 (14.14 min, *m*/*z* 317).

cDNA RACE library construction and cloning of full-length cDNA. Total RNA was isolated from 35-day-old B73 meristem tissues elicited with heat-killed *F. venenatum* hyphae collected after 48 h treatment, as described previously. Total RNA, approximately 2 µg, was subjected to TURBO DNA free treatment (Ambion) and 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. Full-length open reading frames, or those truncated to lack predicted N-terminal plastid transit peptides, were amplified using gene-specific oligonucleotides (Supplementary Table 12). For *Agrobacterium*-mediated transient expression in *N. benthamiana*, full-length open reading frames, including *ZmAN2, ZmKSL2, ZmCYP71Z18, ZmCYP71Z16* and *ZmKR2*, were cloned into the expression vector pLIFE33.

Transient co-expression assays in *N. benthamiana*. For transient expression in *N. benthamiana*, full-length open reading frames in the pLIFE33 vector were transformed into *Agrobacterium tumefaciens* strain GV3101. Resulting cells were grown at 28 °C for 24 h in Luria broth media supplemented with 50 mg l^{-1} of kanamycin, 30 mg l^{-1} of gentamicin and 50 mg l^{-1} of rifampicin. Cells were collected and resuspended to a final optical density$_{600}$ of 0.8 in 10 mM MES buffer with 10 mM $MgCl_2$. Equal volumes of different cultures, as well as the silencing suppressor strain P19, were combined and infiltrated into the leaves of 6-week-old *N. benthamiana* plants using a needleless syringe[84]. Five days post transfection, inoculated leaves were collected for hexane extraction of diterpenoids and GC–MS analyses.

Co-expression of diTPSs and P450s in *E. coli*. Combinatorial expression of diTPSs and P450s was performed using a previously described *E. coli* system engineered for enhanced diterpenoid production[32,85]. For functional analysis, diTPS genes were N-terminally truncated to remove predicted plastid transit peptides. ZmKSL2Δ34 (Zm00001d041082) and ZmKSL3Δ53 (Zm00001d002349) were synthesized and subcloned into the pET28b(+) expression vector (EMD Millipore). For additional co-expression of P450s, synthetic N-terminally modified and codon-optimized genes of ZmCYP71Z16 (Zm00001d014136)[23], ZmCYP71Z18 (Zm00001d014134)[23], ZmKO1 (CYP701A43, Zm00001d046344) and ZmKO2 (CYP701A26, Zm00001d046342) were synthesized and subcloned into the multiple cloning site 2 of the pETDuet-1 expression vector (EMD Millipore) carrying the ZmCPR2 gene (Zm00001d026483) in multiple cloning site 1. Class II diTPS were expressed as previously described with the GGPP synthase from *Abies grandis*[32]. The pACYC-Duet:AgGGPS-ZmAN2 (Zm00001d029648), pACYC-Duet:AgGGPPS-ZmCPS3 (Zm00001d024512) or pACYC-Duet:AgGGPPS-

ZmCPS4 (Zm00001d048874) constructs were expressed alone or in combination with pET28b(+):ZmKSL2 and pETDuet:ZmCPR2:ZmCYP71Z16 or pETDuet:ZmCPR2:ZmCYP71Z18[23,31]. The desired construct combination was then co-transformed into *E. coli* strain BL21DE3-C41 cells (Lucigen) together with pCDFDuet:IRS expression vector for co-expression assays[32,85]. Cultures were grown in 50 ml Terrific Broth medium to an optical density$_{600}$ of ~0.6 at 37 °C and cooled to 16 °C, at which time protein expression was induced by adding 1 mM isopropyl-thio-galactoside (IPTG) followed by incubation for 72 h with supplement 25 mM sodium pyruvate. For P450 co-expression, cultures were further supplemented with 4 mg l^{-1} riboflavin and 75 mg l^{-1} δ-aminolevulinic acid. Enzyme products were extracted with 50 ml of 1/1 ethyl acetate:hexane (v/v), concentrated under N_2 stream and resuspended in 1 ml hexane for GC–MS analysis. For in vitro activity analysis of diTPSs, constructs in the pET28b(+) vector were expressed in *E. coli* BL21DE3-C41 cells and Ni^{2+}-nitrilotriacetic acid affinity-purified as described elsewhere[86]. Single-vial diTPS enzyme assays were performed using 50 µg of recombinant protein and 15 µM of (*E*,*E*,*E*)-GGPP (Sigma–Aldrich) as substrate, with gentle shaking for 1 h at 30 °C, followed by product extraction with hexane for GC–MS analysis.

Creation of *Zmksl2* using CRISPR–Cas9. *ZmKSL2* guide RNA (gRNA) target site selection used the B73 reference genome sequence criteria as described[87]. 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[88]. 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 Table 11) from each plant were sequenced. Two independent homozygous T0 plants with 1 bp insertion at the same position were identified, referred to as ZmKSL2–7 and ZmKSL2–20. The homozygous plants were outcrossed with B73 and the resulting F1 plants were self-pollinated to generate F2 progenies. Forty-eight F2 plants were genotyped, and a homozygous plant without the CRISPR transgene was selected and backcrossed to B73. Finally, two homozygous plants and two wild-type siblings were selected for bioassays by genotyping from self-pollinated plants after B73 backcrossing four successive times.

Identification of *Zmkr2* and *Zmko2* mutants. The *Dsg* insertion (tdsgR87C04) in ZmKR2 (Zm00001d018847, B73RefGen_V4) was verified by designing PCR primer pairs, with one gene-specific pairs (Supplementary Table 11) from ZmKR2 and one primer from the insertion, GFP gene (GFP_AC-DS: TTCGCTCATGTGTTGAGCAT)[47]. qrtPCR analysis confirmed that the *Dsg* insertion severely affects transcription of ZmKR2 in response to fungal elicitation (Supplementary Fig. 20). The Mu insertion line mu-illumina_253466.6 was confirmed in ZmKR2 by PCR genotyping with one pair of gene-specific primers from ZmKR2 (Supplementary Table 11) and a degenerate MuTIR primer (GCCTC(T/C)ATTTCGTCGAATCC(C/G))[64]. Mu elements in mu-Illumina mutants have terminal inverted repeats, so the primer for MuTIR works in both directions. The mu1068966 insertion in the first intron of ZmKO2 (Zm00001d046342, B73RefGen_V4) was verified by PCR genotyping with a pair of gene-specific primers (Supplementary Table 11) and a primer MuTIR6 from mu insertion. qrtPCR analysis further confirmed that the mu1068966 insertion significantly compromised transcription of *ZmKO2* in response to SLB elicitation (Supplementary Fig. 17).

Analyses of maize diterpenoid pathway proteins. W22 maize plants were grown individually in 1G pots for 35 d. All plants were stem elicited with heat-killed *F. venenatum* hyphae under staged timings such that all time points (0, 2, 4, 72, 96 and 120 h) could be collected within the same hour and precise plant age. Stem tissues from four plants were collected and pooled to generate a single homogenous sample per time point, ground in liquid nitrogen and stored at −80 °C. Tissue powders were suspended in extraction buffer (8 M urea/100 mM Tris/5 mM Tris(2-carboxyethyl)phosphine (TCEP)/phosphatase inhibitors, pH 7). Proteins were precipitated by adding four volumes of cold acetone and incubated at 4 °C for 2 h. Samples were centrifuged at 4,000*g* for 5 min at 4 °C. Supernatant was removed and discarded. Proteins were resuspended in urea extraction buffer and precipitated by cold acetone. Protein pellets were washed by cold methanol with 0.2 mM sodium orthovanadate to further remove non-protein contaminants and resuspended in the original extraction buffer. Proteins were then digested with Lys-C (Wako Chemicals, 125-05061) at 37 °C for 15 min, diluted eight-fold with 1 M urea containing 100 mM Tris and secondarily digested with trypsin (Roche, 03 708 969 001) for 4 h. Digested peptides were purified on Waters Sep-Pak C18 cartridges and eluted with 60% acetonitrile. TMT-10 labelling was performed in 50% acetonitrile/150 mM Tris, pH 7 and checked by LC–MS/MS to confirm >99% efficiency. Labelled peptides from each time point sample were pooled together for 2D-nanoLC–MS/MS analysis. An Agilent 1100 high-performance liquid chromatography (HPLC) system was used to deliver a flow rate of 600 nl min^{-1} to a custom 3-phase capillary chromatography column through a splitter. Column phases consisted of a 30-cm long reverse phase (RP1: 5 µm Zorbax SB-C18, Agilent), 8-cm long strong cation exchange (SCX: 3 µm PolySulfoethyl, PolyLC) and 40-cm long reverse phase 2 (RP2: 3.5 µm BEH C18, Waters) coupled with an electrospray tip of fused silica tubing pulled to a sharp point (inner diameter <1 µm). Peptide mixtures were loaded onto RP1 and the three column sections were joined and mounted on a custom electrospray adaptor for online nested elutions. Peptides were eluted from the RP1 section to the SCX section using a 0–80% acetonitrile gradient for 60 min, and then fractionated by the SCX column section using a series of 20 step salt gradients of ammonium acetate over 20 min, followed by high-resolution reverse phase separation on the RP2 section of the column using an acetonitrile gradient of 0–80% for 150 min. Spectra were acquired on a Q-exactive-HF mass spectrometer (Thermo Electron Corporation) operated in positive ion mode with a source temperature of 275 °C and spray voltage of 3 kV. Xcalibur 4.0 (Thermo Scientific) was used for automated data-dependent acquisition and employed the top 20 ions with an isolation window of 1.0 Da and normalized collision energy of 30. The resolving power was set at 60,000 full width at half maximum (FWHM) for MS and 30,000 FWHM for MS–MS scans, respectively. Dynamic exclusion was used to improve the duty cycle. The raw data was extracted and searched using Spectrum Mill v.B.06 (Agilent). MS–MS spectra with a sequence tag length of 1 or less are considered to be poor spectra and were discarded. The remaining MS–MS spectra were searched against the maize W22 gene set. Search parameters were set to Spectrum Mill's default settings, with the enzyme parameter limited to full tryptic peptides with a maximum mis-cleavage of 1. A 1/1 concatenated forward–reverse database was constructed to calculate the false discovery rate. Cut-off scores were dynamically assigned to each dataset to obtain the false discovery rates of 0.1% for peptides and 1% for proteins. Proteins that share common peptides were grouped using principles of parsimony to address protein database redundancy. Total TMT-10 reporter intensities were used for relative protein quantitation. Peptides shared among different protein groups were removed before quantitation. Isotope impurities of TMT-10 reagents were corrected using correction factors provided by the manufacturer (Thermo Fisher Scientific). Median normalization was performed to normalize the protein TMT-10 reporter intensities in which the log ratios between different TMT-10 tags were adjusted globally such that the median log ratio was 0.

Purification, analyses and NMR of KA4 and KB4. For the purification of kauralexin A4 and B4, 1 kg of 80-day-old field-grown B97 husk tissue was ground to a powder in liquid nitrogen, extracted with 1 l of methanol, filtered and dried using a rotary evaporator. The resulting oily residue was then separated by preparative flash chromatography (CombiFlashRf; Teledyne ISCO) on a 15.5-g Silica (RediSepRf High Performance GOLD) column. The mobile phase consisted of solvent A (100% hexane) and solvent B (100% ethyl acetate), with a continuous gradient of 0–100% B from 1 to 50 min using a flow rate of 19 ml min^{-1}. Fractions were collected and analysed every 5 min by LC–MS. In the 25–30 min fraction, an enrichment of putative candidates for KA4 and KB4 was observed. This fraction was purified further by HPLC using repeated 1 mg injections on a Zorbax RX-C18 (length, 250 mm; particle size, 3 µm; internal diameter, 4.6 mm; Agilent) column and a mobile phase consisting of solvent A (1:4 acetonitrile (ACN)/H_2O) and solvent B (100% ACN) with a continuous gradient of A–B from 0 to 27 min using a flow rate of 1 ml min^{-1}. The recollected fractions spanning 15–16 min and 17–18 min contained KB4 and KA4 at approximately 85% purity and were used to generate samples for NMR. Purified KA4 and KB4 were dissolved in chloroform-d (Cambridge Isotope Laboratories), and NMR spectra were acquired on a Bruker 600-MHz spectrometer equipped with a 1.7-mm cryo probe triple resonance interface. Chemical shifts were calculated by reference to known chloroform-d (^{13}C 77.2 ppm, ^{1}H 7.26 ppm) signals. All spectra were acquired using standard experiments on a Bruker Avance III console and TopSpin 2.1.6 software, including 1D 1 H and 2D heteronuclear single quantum coherence spectroscopy, correlated spectroscopy and heteronuclear multiple bond correlations (600 MHz).

Phylogenetic analysis of protein sequences and gene sequences used for heterologous expression. Protein sequence alignments derived from UniProtKB and Genbank IDs (Supplementary Table 13) were performed using ClustalW as implemented in the BioEdit software package (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). The maximum-likelihood phylogenetic trees were built using MEGA7 (http://www.megasoftware.net/megabeta.php) with bootstrap values based on 1,000 iterations. All gene sequences used in this study for enzyme characterization are detailed (Supplementary Table 14).

Statistical analyses. Statistical analyses were conducted using GraphPad Prism 8.0 (GraphPad Software, Inc.) and JMP Pro 13.0 (SAS Institute Inc.). One-way analyses of variance (ANOVAs) were performed to evaluate statistical differences. Tukey tests were used to correct for multiple comparisons between control and treatment groups. Student's unpaired *t*-tests were used for pairwise comparisons. P values < 0.05 were considered significant.

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

Data availability

Raw read sequences have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE120135. All other data that support the findings of this are available from the corresponding author upon request.

Received: 4 March 2019; Accepted: 26 July 2019;
Published online: 16 September 2019

References

1. Dixon, R. A. Natural products and plant disease resistance. *Nature* **411**, 843–847 (2001).
2. Gershenzon, J. & Dudareva, N. The function of terpene natural products in the natural world. *Nat. Chem. Biol.* **3**, 408–414 (2007).
3. Pichersky, E. & Lewinsohn, E. Convergent evolution in plant specialized metabolism. *Ann. Rev. Plant Biol.* **62**, 549–566 (2011).
4. Turlings, T. C. J. & Erb, M. Tritrophic interactions mediated by herbivore-induced plant volatiles: mechanisms, ecological relevance, and application potential. *Annu. Rev. Entomol.* **63**, 433–452 (2018).
5. Lopez-Nieves, S. et al. Relaxation of tyrosine pathway regulation underlies the evolution of betalain pigmentation in Caryophyllales. *New Phytol.* **217**, 896–908 (2018).
6. Moghe, G. D. & Last, R. L. Something old, something new: conserved enzymes and the evolution of novelty in plant specialized metabolism. *Plant Physiol.* **169**, 1512–1523 (2015).
7. Chae, L., Kim, T., Nilo-Poyanco, R. & Rhee, S. Y. Genomic signatures of specialized metabolism in plants. *Science* **344**, 510–513 (2014).
8. Chen, F., Tholl, D., Bohlmann, J. & Pichersky, E. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J.* **66**, 212–229 (2011).
9. Peters, R. J. Two rings in them all: the labdane-related diterpenoids. *Nat. Prod. Rep.* **27**, 1521–1530 (2010).
10. Ma, K. W. & Ma, W. Phytohormone pathways as targets of pathogens to facilitate infection. *Plant Mol. Biol.* **91**, 713–725 (2016).
11. Robert-Seilaniantz, A., Navarro, L., Bari, R. & Jones, J. D. Pathological hormone imbalances. *Curr. Opin. Plant Biol.* **10**, 372–379 (2007).
12. Robert-Seilaniantz, A., Grant, M. & Jones, J. D. G. Hormone crosstalk in plant disease and defense: more than just jasmonate–salicylate antagonism. *Annu. Rev. Phytopathol.* **49**, 317–343 (2011).
13. Hedden, P. & Thomas, S. G. Gibberellin biosynthesis and its regulation. *Biochem. J.* **444**, 11–25 (2012).
14. Mitchell, J. W., Skaggs, D. P. & Anderson, W. P. Plant growth-stimulating hormones in immature bean seeds. *Science* **114**, 159–161 (1951).
15. Phinney, B. O., West, C. A., Ritzel, M. & Neely, P. M. Evidence for 'gibberellin-like' substances from flowering plants. *Proc. Natl Acad. Sci. USA* **43**, 398–404 (1957).
16. Yabuta, T. Biochemistry of the 'bakanae' fungus of rice. *Agric. Hort. (Tokyo)* **10**, 17–22 (1935).
17. Zi, J., Mafu, S. & Peters, R. J. To gibberellins and beyond! Surveying the evolution of (di)terpenoid metabolism. *Annu. Rev. Plant Biol.* **65**, 259–286 (2014).
18. Gao, Y., Honzatko, R. B. & Peters, R. J. Terpenoid synthase structures: a so far incomplete view of complex catalysis. *Nat. Prod. Rep.* **29**, 1153–1175 (2012).
19. Schmelz, E. A. et al. Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins. *Plant J.* **79**, 659–678 (2014).
20. Lu, X. et al. Inferring roles in defense from metabolic allocation of rice diterpenoids. *Plant Cell* **30**, 1119–1131 (2018).
21. Christensen, S. A. et al. Commercial hybrids and mutant genotypes reveal complex protective roles for inducible terpenoid defenses in maize. *J. Exp. Bot.* **69**, 1693–1705 (2018).
22. Fu, J. et al. A tandem array of ent-kaurene synthases in maize with roles in gibberellin and more specialized metabolism. *Plant Physiol.* **170**, 742–751 (2016).
23. Mafu, S. et al. Discovery, biosynthesis and stress-related accumulation of dolabradiene-derived defenses in maize. *Plant Physiol.* **176**, 2677–2690 (2018).
24. Schmelz, E. A. et al. Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. *Proc. Natl Acad. Sci. USA* **108**, 5455–5460 (2011).
25. Vaughan, M. M. et al. Accumulation of terpenoid phytoalexins in maize roots is associated with drought tolerance. *Plant Cell Environ.* **38**, 2195–2207 (2015).
26. Bensen, R. J. et al. Cloning and characterization of the maize An1 gene. *Plant Cell* **7**, 75–84 (1995).
27. Mao, H., Shen, Q. & Wang, Q. CYP701A26 is characterized as an ent-kaurene oxidase with putative involvement in maize gibberellin biosynthesis. *Biotechnol. Lett.* **39**, 1709–1716 (2017).
28. Mafu, S. et al. Probing the promiscuity of ent-kaurene oxidases via combinatorial biosynthesis. *Proc. Natl Acad. Sci. USA* **113**, 2526–2531 (2016).
29. Harris, L. J. et al. The maize An2 gene is induced by *Fusarium* attack and encodes an ent-copalyl diphosphate synthase. *Plant Mol. Biol.* **59**, 881–894 (2005).
30. Wisecaver, J. H. et al. A global coexpression network approach for connecting genes to specialized metabolic pathways in plants. *Plant Cell*, **29**, 944–959 (2017).
31. Murphy, K. M., Ma, L. T., Ding, Y. Z., Schmelz, E. A. & Zerbe, P. Functional characterization of two class II diterpene synthases indicates additional specialized diterpenoid pathways in maize (*Zea mays*). *Front. Plant Sci.* **9**, 1542 (2018).
32. 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).
33. Morrone, D. et al. Evident and latent plasticity across the rice diterpene synthase family with potential implications for the evolution of diterpenoid metabolism in the cereals. *Biochem. J.* **435**, 589–595 (2011).
34. Zerbe, P. & Bohlmann, J. Plant diterpene synthases: exploring modularity and metabolic diversity for bioengineering. *Trends Biotechnol.* **33**, 419–428 (2015).
35. Bak, S. et al. Cytochromes P450. *Arabidopsis Book* **9**, e0144 (2011).
36. Wu, Y. S., Hillwig, M. L., Wang, Q. & Peters, R. J. Parsing a multifunctional biosynthetic gene cluster from rice: biochemical characterization of CYP71Z6 & 7. *FEBS Lett.* **585**, 3446–3451 (2011).
37. Wang, X. Y. et al. Genome alignment spanning major poaceae lineages reveals heterogeneous evolutionary rates and alters inferred dates for key evolutionary events. *Mol. Plant* **8**, 885–898 (2015).
38. Flint-Garcia, S. A. et al. Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant J.* **44**, 1054–1064 (2005).
39. Hedden, P. & Sponsel, V. A century of gibberellin research. *J. Plant Growth Regul.* **34**, 740–760 (2015).
40. Kono, T. J. Y., Brohammer, A. B., McGaugh, S. E. & Hirsch, C. N. Tandem duplicate genes in maize are abundant and date to two distinct periods of time. *G3* **8**, 3049–3058 (2018).
41. Helliwell, C. A., Poole, A., James Peacock, W. & Dennis, E. S. *Arabidopsis* ent-kaurene oxidase catalyzes three steps of gibberellin biosynthesis. *Plant Physiol.* **119**, 507–510 (1999).
42. Meyer, J., Berger, D. K., Christensen, S. A. & Murray, S. L. RNA-Seq analysis of resistant and susceptible sub-tropical maize lines reveals a role for kauralexins in resistance to grey leaf spot disease, caused by *Cercospora zeina*. *BMC Plant Biol.* **17**, 197 (2017).
43. Yang, D. L. et al. Altered disease development in the eui mutants and Eui overexpressors indicates that gibberellins negatively regulate rice basal disease resistance. *Mol. Plant* **1**, 528–537 (2008).
44. Langlois, V. S., Zhang, D., Cooke, G. M. & Trudeau, V. L. Evolution of steroid-5alpha-reductases and comparison of their function with 5beta-reductase. *Gen. Comp. Endocrinol.* **166**, 489–497 (2010).
45. Li, J. M., Biswas, M. G., Chao, A., Russell, D. W. & Chory, J. Conservation of function between mammalian and plant steroid 5 alpha-reductases. *Proc. Natl Acad. Sci. USA* **94**, 3554–3559 (1997).
46. Uemura, M. et al. Novel 5 alpha-steroid reductase (SRD5A3, type-3) is overexpressed in hormone-refractory prostate cancer. *Cancer Sci.* **99**, 81–86 (2008).
47. 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. USA* **114**, 5165–5170 (2017).
48. McMullen, M. D., Frey, M. & Degenhardt, J. in *Handbook of Maize: Its Biology* (eds Bennetzen, J. L. & Hake, S. C.) 271–289 (Springer, 2009).
49. Wouters, F. C., Blanchette, B., Gershenzon, J. & Vassão, D. G. Plant defense and herbivore counter-defense: benzoxazinoids and insect herbivores. *Phytochem. Rev.* **15**, 1127–1151 (2016).
50. Bagnaresi, P. et al. Comparative transcriptome profiling of the early response to *Magnaporthe oryzae* in durable resistant vs susceptible rice (*Oryza sativa* L.) genotypes. *PLoS ONE* **7**, e51609 (2012).
51. Toyomasu, T. et al. Reverse-genetic approach to verify physiological roles of rice phytoalexins: characterization of a knockdown mutant of OsCPS4 phytoalexin biosynthetic gene in rice. *Physiol. Plant.* **150**, 55–62 (2014).
52. Brohammer, A. B., Kono, T. J. Y., Springer, N. M., McGaugh, S. E. & Hirsch, C. N. The limited role of differential fractionation in genome content variation and function in maize (*Zea mays* L.) inbred lines. *Plant J.* **93**, 131–141 (2018).
53. Schnable, J. C., Springer, N. M. & Freeling, M. Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proc. Natl Acad. Sci. USA* **108**, 4069–4074 (2011).
54. Pelot, K. A. et al. Functional diversity of diterpene synthases in the biofuel crop switchgrass. *Plant Physiol.* **178**, 54–71 (2018).
55. Miyazaki, S., Katsumata, T., Natsume, M. & Kawaide, H. The CYP701B1 of *Physcomitrella patens* is an ent-kaurene oxidase that resists inhibition by uniconazole-P. *FEBS Lett.* **585**, 1879–1883 (2011).

56. Morrone, D., Chen, X. M., Coates, R. M. & Peters, R. J. Characterization of the kaurene oxidase CYP701A3, a multifunctional cytochrome P450 from gibberellin biosynthesis. *Biochem. J.* **431**, 337–344 (2010).
57. Vrabka, J. et al. Production and role of hormones during interaction of *Fusarium* species with maize (*Zea mays* L.) seedlings. *Front. Plant Sci.* **9**, 1936–1936 (2019).
58. Hedden, P. & Phinney, B. O. Comparison of ent-kaurene and ent-isokaurene synthesis in cell-free systems from etiolated shoots of normal and dwarf-5 maize seedlings. *Phytochemistry* **18**, 1475–1479 (1979).
59. Bohlmann, J., Meyer-Gauen, G. & Croteau, R. Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proc. Natl Acad. Sci. USA* **95**, 4126–4133 (1998).
60. Laursen, T. et al. Characterization of a dynamic metabolon producing the defense compound dhurrin in sorghum. *Science* **354**, 890–893 (2016).
61. Mehrshahi, P., Johnny, C. & DellaPenna, D. Redefining the metabolic continuity of chloroplasts and ER. *Trends Plant Sci.* **19**, 501–507 (2014).
62. Wiemann, P. et al. Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites. *PLoS Pathog.* **9**, e1003475 (2013).
63. Lu, X., Hershey, D. M., Wang, L., Bogdanove, A. J. & Peters, R. J. An ent-kaurene-derived diterpenoid virulence factor from *Xanthomonas oryzae* pv. *oryzicola*. *New Phytol.* **206**, 295–302 (2015).
64. Williams-Carrier, R. et al. Use of Illumina sequencing to identify transposon insertions underlying mutant phenotypes in high-copy Mutator lines of maize. *Plant J.* **63**, 167–177 (2010).
65. Ding, Y. Z. et al. Selinene volatiles are essential precursors for maize defense promoting fungal pathogen resistance. *Plant Physiol.* **175**, 1455–1468 (2017).
66. Huffaker, A. et al. Novel acidic sesquiterpenoids constitute a dominant class of pathogen-induced phytoalexins in maize. *Plant Physiol.* **156**, 2082–2097 (2011).
67. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
68. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
69. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
70. Trapnell, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
71. Kremling, K. A. G. et al. Dysregulation of expression correlates with rare-allele burden and fitness loss in maize. *Nature* **555**, 520 (2018).
72. Obayashi, T. & Kinoshita, K. Rank of correlation coefficient as a comparable measure for biological significance of gene coexpression. *DNA Res.* **16**, 249–260 (2009).
73. 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).
74. Zhang, Z. W. et al. Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* **42**, 355–U118 (2010).
75. 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).
76. Lipka, A. E. et al. GAPIT: genome association and prediction integrated tool. *Bioinformatics* **28**, 2397–2399 (2012).
77. Bradbury, P. J. et al. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635 (2007).
78. Cook, J. P. et al. Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant Physiol.* **158**, 824–834 (2012).
79. Elshire, R. J. et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **6**, e19379 (2011).
80. 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).
81. VanRaden, P. M. Efficient methods to compute genomic predictions. *J. Dairy Sci.* **91**, 4414–4423 (2008).
82. Turner, S. D. qqman: an R package for visualizing GWAS results using Q–Q and Manhattan plots. Preprint at https://doi.org/10.1101/005165 (2014).
83. 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).
84. 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).
85. Kitaoka, N., Lu, X., Yang, B. & Peters, R. J. The application of synthetic biology to elucidation of plant mono-, sesqui-, and diterpenoid metabolism. *Mol. Plant* **8**, 6–16 (2015).
86. Zerbe, P. et al. Gene discovery of modular diterpene metabolism in nonmodel systems. *Plant Physiol.* **162**, 1073–1091 (2013).
87. Brazelton, V. A. et al. A quick guide to CRISPR sgRNA design tools. *GM Crops Food* **6**, 266–276 (2015).
88. Char, S. N. et al. An agrobacterium-delivered CRISPR/Cas9 system for high-frequency targeted mutagenesis in maize. *Plant Biotechnol. J.* **15**, 257–268 (2017).

Acknowledgements

We thank A. Steinbrenner, P. Weckwerth, K. Dressano, J. Chan, K. O'Leary, M. Broemmer, H. Riggleman, S. Reyes, S. Delgado and J. Tran for help in planting, treatments and sampling (UCSD). Thanks also to L. Smith (UCSD) for shared UCSD Biology Field Station management. This work was supported by a grant from the National Science Foundation Plant-Biotic Interactions Program (grant no. 1758976 to E.S. and P.Z.), by a Department of Energy Joint Genome Institute Community Science Program grant (no. CSP2568 to P.Z., E.S. and A.H.) and by a fellowship provided by the National Science Foundation Graduate Research Fellowship Program (to K.M.M.).

Author contributions

Y.D., K.M.M., A.H., P.Z. and E.A.S. designed the experiments and analysed the data. Y.D., E.P., S.A.C., L.J., R.J.S., J.B., P.Z., K.A.K. and E.S.B. designed, performed and analysed the transcriptome data. Y.D., S.M., K.M.M., P.Z. and E.A.S. performed MS experiments and MS-related data analysis. S.M., Y.D., K.M.M., Q.W., E.S. and M.W. performed and analysed the enzyme co-expression data. B.Y., S.N.C. and Y.D. designed gRNA constructs and generated the Zmksl2 maize mutants. A.S., G.C.-F. and C.C.H. performed and analysed the NMR data. Y.D., P.Z. and E.A.S. wrote the manuscript with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41477-019-0509-6.

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

Correspondence and requests for materials should be addressed to E.A.S.

Peer Review Information: *Nature Plants* thanks Hugues Renault and other, anonymous, reviewers for their contribution to the peer review of this work.

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

nature research

Corresponding author(s): Eric Schmelz

Last updated by author(s): 7/18/2019

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 Authors & Referees 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 using TopSpin 2.1.6 software. qRT-PCR data were collected using Bio-RAD CFX Manager 3.1 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 using the Compressed MLM. Linkage analysis was done with Windows QTL Cartographer V2.5. Trimmomatic v0.32 was used for processing raw RNA-seq fastq reads, and Bowtie2 v2.2.3.0 and TopHat v2.0.13. were employed to align qualified reads to the reference genome. Cufflinks v2.2.1 was used to compute gene expression values.

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/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

Sequence data that support the findings of this study (Supplementary Table 1-2) have been deposited in NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) with the accession code "GSE120135". All plant accessions used, gene identifiers, unique protein sequences, heterologously expressed

genes and web links to publicly available datasets are listed in detail in the Methods and Supplementary Tables. All unique materials are readily available from the authors on request.

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 sample size calculation was performed. All experiments were performed in three or more independent biological replicates to ensure reproducibility and statistical relevance. Uniform vigorous plants produce exceedingly low levels of kauralexins and fungal elicitation elicits highly consistent qualitative and quantitative differences in production; thus, a modest number of biological replicates (n=3-5) is sufficient to detect statistical differences. GWAS studies were repeated twice and intentionally used different plant growth conditions to ensure statistical effects observed were robust.
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 or more 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	Blinding was not attempted in this study due to the large number of chemicals, solutions, and vectors that had to be prepared prior to treatments. Initial RNA-seq data analyses were done blind.

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
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

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

nature plants

SUPPLEMENTARY INFORMATION

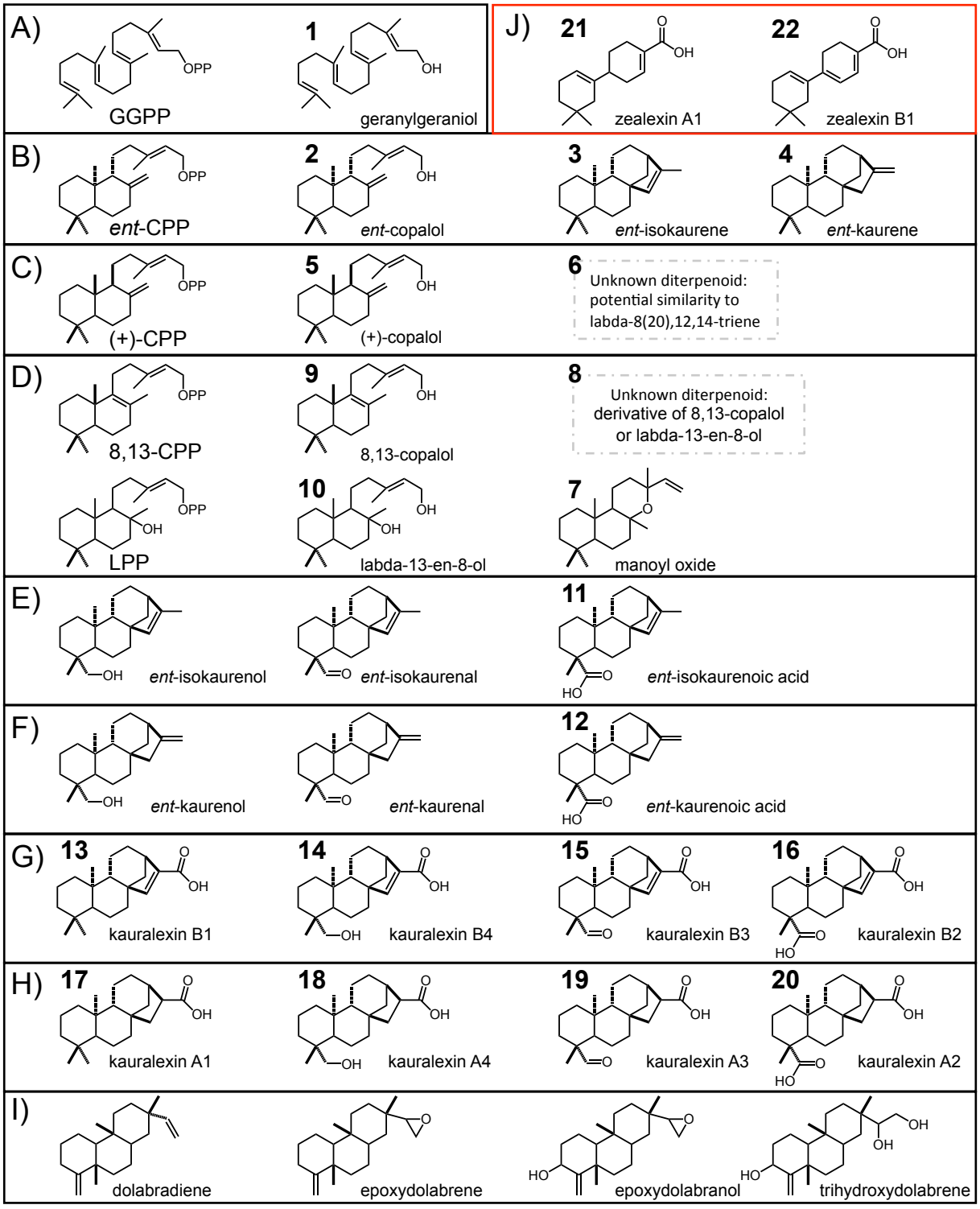
https://doi.org/10.1038/s41477-019-0509-6

In the format provided by the authors and unedited.

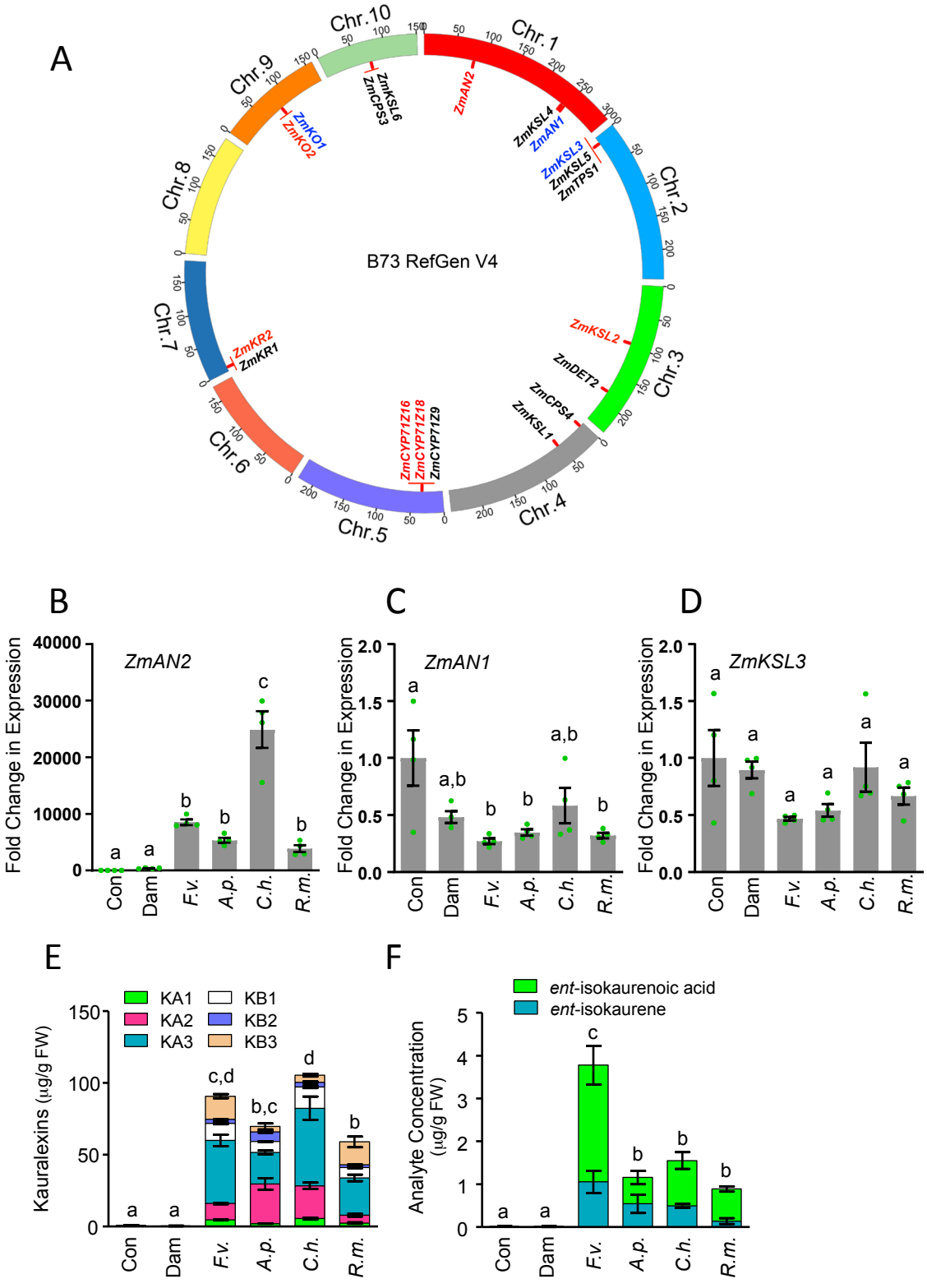
Multiple genes recruited from hormone pathways partition maize diterpenoid defences

Yezhang Ding[1], Katherine M. Murphy[2], Elly Poretsky[1], Sibongile Mafu[2], Bing Yang[3], Si Nian Char[3], Shawn A. Christensen[4], Evan Saldivar[1], Mengxi Wu[1], Qiang Wang[5], Lexiang Ji[6], Robert J. Schmitz[7], Karl A. Kremling[8], Edward S. Buckler[8,9], Zhouxin Shen[1], Steven P. Briggs[1], Jörg Bohlmann[10], Andrew Sher[1], Gabriel Castro-Falcon[11], Chambers C. Hughes[11], Alisa Huffaker[1], Philipp Zerbe[2] and Eric A. Schmelz[1]*

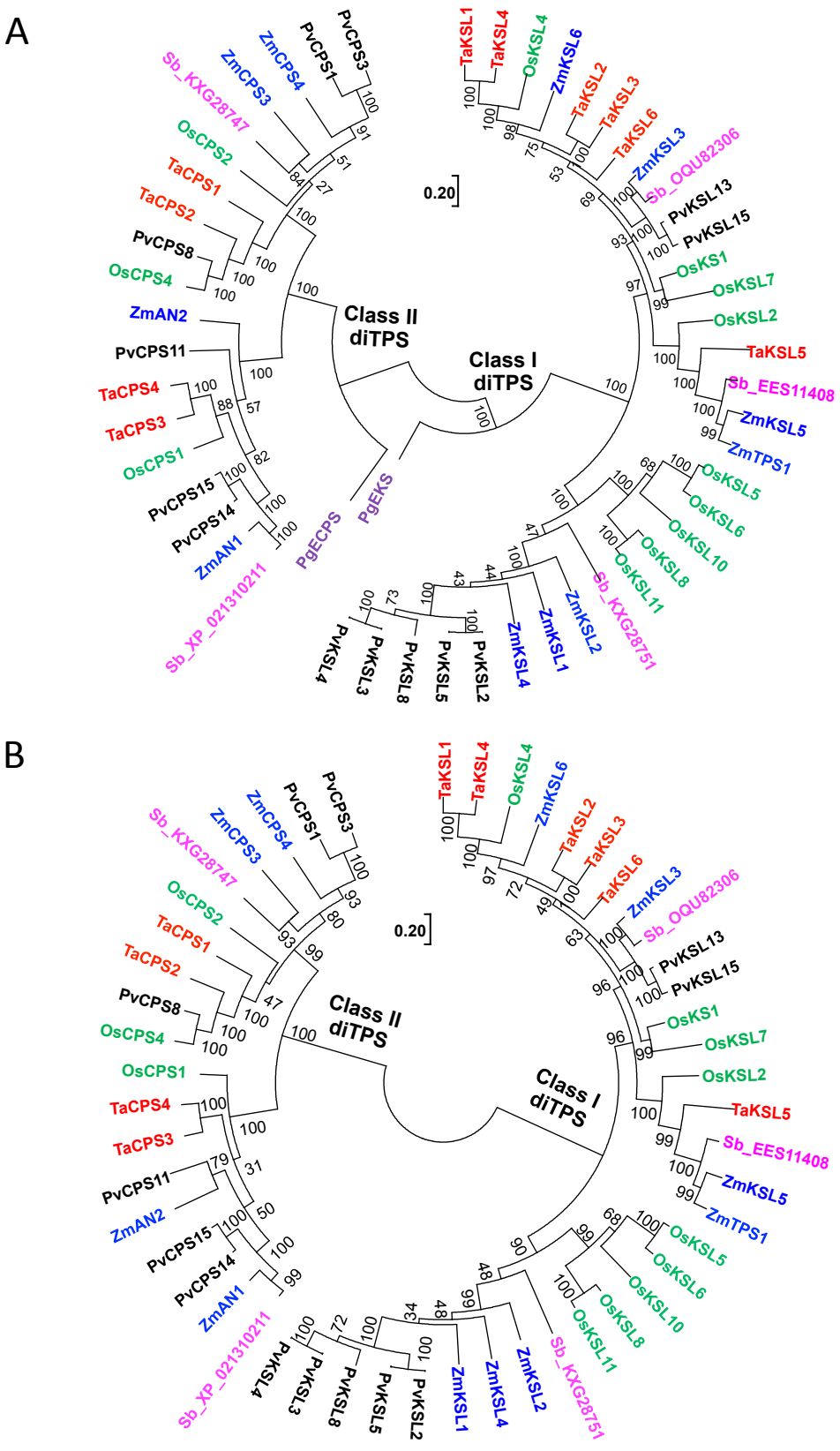
[1]Section of Cell and Developmental Biology, University of California San Diego, La Jolla, CA, USA. [2]Department of Plant Biology, University of California Davis, Davis, CA, USA. [3]Department of Genetics, Development and Cell Biology, Iowa State University, Ames, IA, USA. [4]Chemistry Research Unit, Center for Medical, Agricultural, and Veterinary Entomology, US Department of Agriculture—Agricultural Research Service, Gainesville, FL, USA. [5]Institute of Ecological Agriculture, Sichuan Agricultural University, Chengdu, China. [6]Institute of Bioinformatics, University of Georgia, Athens, GA, USA. [7]Department of Genetics, University of Georgia, Athens, GA, USA. [8]Department of Plant Breeding and Genetics, Cornell University, Ithaca, NY, USA. [9]Robert W. Holley Center for Agriculture and Health, US Department of Agriculture—Agricultural Research Service, Ithaca, NY, USA. [10]Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, Canada. [11]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA. *e-mail: eschmelz@ucsd.edu



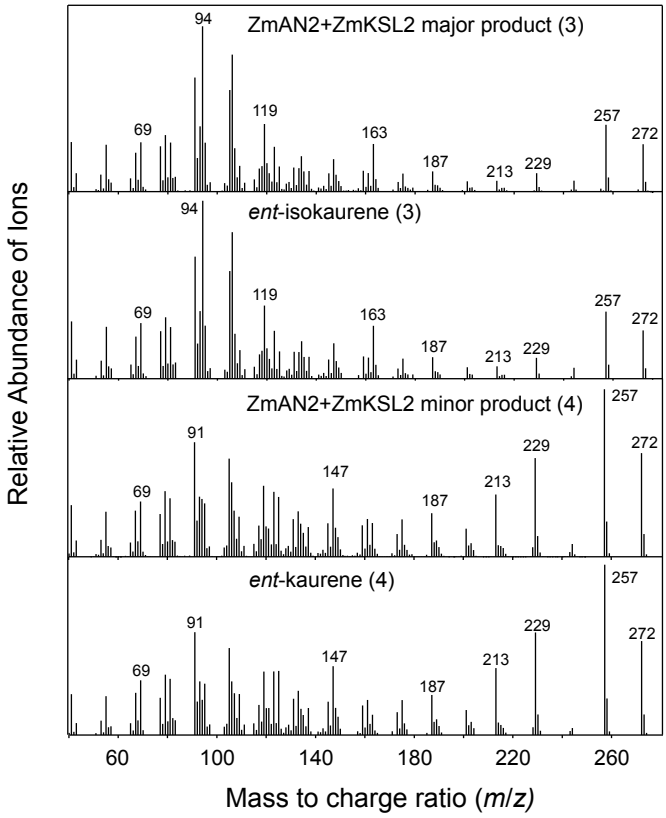
Supplementary Fig. 1. Maize isoprenoid precursors and products in the current study relevant to elucidation of the kauralexin pathway. (A) Precursor required for all type II DiTPS, geranylgeranyl-diphosphate (GGPP) and the dephosphorylated product geranylgeraniol (1). **(B)** Products of the *ent*-CPS ZmAN1 and ZmAN2, include; *ent*-copalyl diphosphate (*ent*-CPP) and the dephosphorylated product *ent*-copalol (2). Products of *ent*-CPS paired with ZmKSL2 are *ent*-isokaurene (3) and *ent*-kaurene (4). **(C)** Products of the (+)-CPS ZmCPS3 include (+)-copalyl-diphosphate [(+)-CPP], the dephosphorylated product (+)-copalol (5) and pairings with ZmKSL2 yield an unknown diterpenoid possibly similar to labda-8(20),12,14-triene (6). **(D)** Products of ZmCPS4, a labda-8,13-dien-15-yl diphosphate (8,13-CPP) synthase include 8,13-CPP, the dephosphorylated product 8,13-copalol (9), an unknown diterpenoid derivative of 8,13-copalol or labda-13-en-8-ol (8), labda-13-en-8-ol diphosphate (LPP), the dephosphorylated product labda-13-en-8-ol (10) and manoyl oxide (7). **(E)** Products of ZmAN2+ZmKSL2 plus kaurene oxidases ZmKO1 or ZmKO2, *ent*-isokaurenol, *ent*-isokaurenal, *ent*-isokaurenoic acid (11). **(F)** Products of ZmAN2+ZmKSL3 plus kaurene oxidases ZmKO1 or ZmKO2, *ent*-kaurenol, *ent*-kaurenal, *ent*-kaurenoic acid (12). **(G)** B-series kauralexins (KB) include KB1 (13), KB4 (14), KB3 (15), KB2 (16). **(H)** A-series kauralexins (KA) include KA1 (17), KA4 (18), KA3 (19), KA2 (20). **(I)** Known dolabralexin products derived from ZmAN2 + ZmKSL4 include dolabradiene, epoxydolabrene (15,16-epoxydolabrene), epoxydolabranol (3b-hydroxy-15,16-epoxydolabrene), and trihydroxydolabrene (3b,15,16-trihydroxydolabrene). **(J)** Representative acidic maize sesquitepenoids include zealexin A1 and zealexin B1. **Note:** Numbered compounds (1-22) in approximate order of occurrence represent those directly measured in the present study, unnumbered compounds represent intermediates or enzyme products not directly detected or measured in samples.



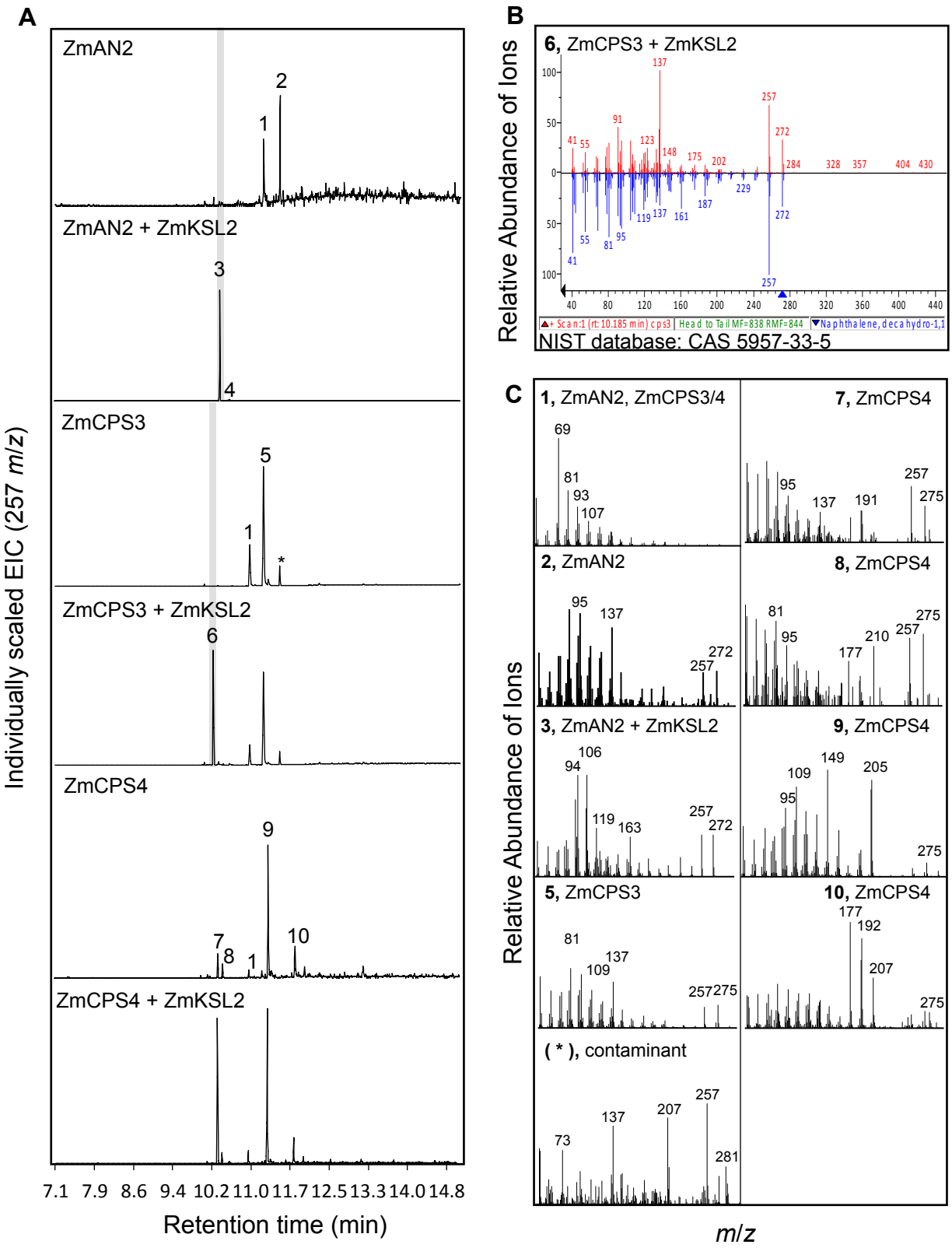
Supplementary Fig. 2. The kauralexin pathway spans 5 chromosomes, is fungal inducible and results in the accumulation of precursors and products. (A) Chromosomal location of maize gibberellin (blue), kauralexin (red), and other (black) pathway genes examined in the current study. Chromosome sizes are scaled in megabases (Mbs) according to B73 RefGen_V4, for additional details refer to Supplementary Table 2. Average (n = 4, ±SEM) abundance of (B) *ZmAN2*, (C) *ZmAN1* and (D) *ZmKSL3* transcripts in stems by elicited by diverse fungi and resulting (E) kauralexins and (F) pathway precursors. Stems of 30-day old maize (var. Mo17) were damaged and treated with either 100 μL of either water alone (Dam) or spore suspensions (1 ×10^7 μL^{-1}) of *F. verticillioides* (*F.v.*), *A. parasiticus* nor-1 (*A.p.*), *C. heterostrophus* (*C.h.*), or *R. microsporus* (*R.m.*). Tissues were harvested at 2- and 4-d for qRT-PCR transcripts and metabolite levels, respectively. Intact stems were used for controls (Con). Gene expression was normalized to the internal reference *ZmEF-1α*. Error bars in the bar charts (B, C, D, E, and F) indicate mean ± SEM (n = 4 biologically independent replicates). Within plots, different letters (a–d) represent significant differences (one-way ANOVA followed by Tukey's test corrections for multiple comparisons, $P < 0.05$).



Supplementary Fig. 3. Maximum likelihood phylogenetic tree of DiTPSs from maize (*Zea mays,* blue), sorghum (*Sorghum bicolor,* pink), rice (*Oryza sativa,* green), wheat (*Triticum aestivum,* red), and switchgrass (*Panicum virgatum,* black). The tree is rooted (A) with spruce diterpene synthases (purple) or unrooted (B). Bootstrap values calculated from 1,000 iterations are indicated at the nodes. Protein accession numbers are listed in Supplementary Table 13.

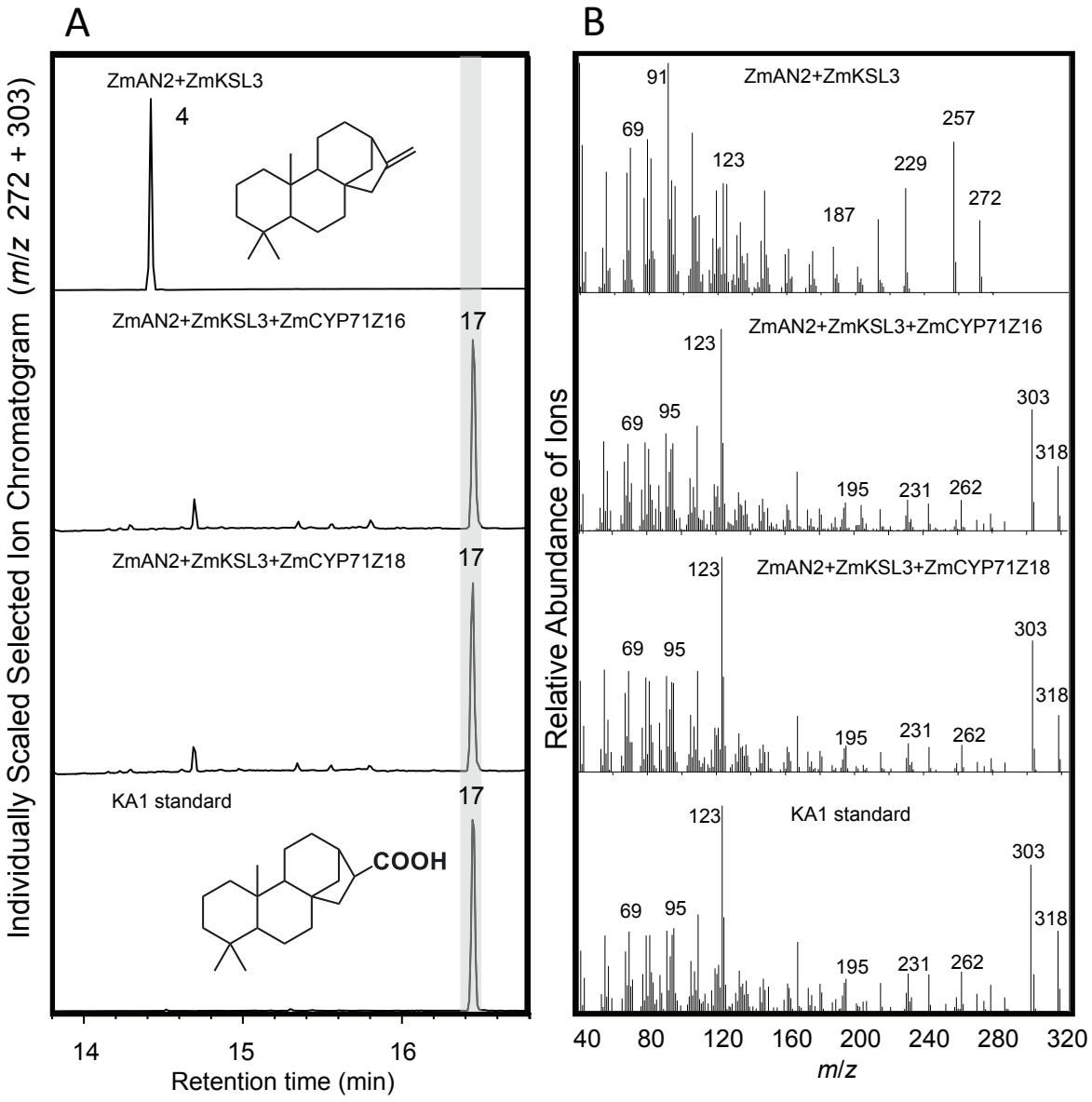


Supplementary Fig. 4. *In planta* functional characterization of maize class I enzyme ZmKSL2 as an *ent*-isokaurene synthase. Corresponding EI-mass spectra (from Fig. 1d) of the major (≈ 95%) and minor (≈5%) products present in *Agrobacterium*-mediated transient *N*. *benthamiana* co-expression assays using the class I diTPS *ent*-copalyl diphosphate synthase ZmAN2 paired with the class II diTPS ZmKSL2. Included are corresponding mass spectra of authentic *ent*-isokaurene and *ent*-kaurene standards. EI-spectra labeling is referenced to Supplementary Fig. 1 (3, *ent*-isokaurene; 4, *ent*-kaurene). Four independent experiments were performed and showed similar results.

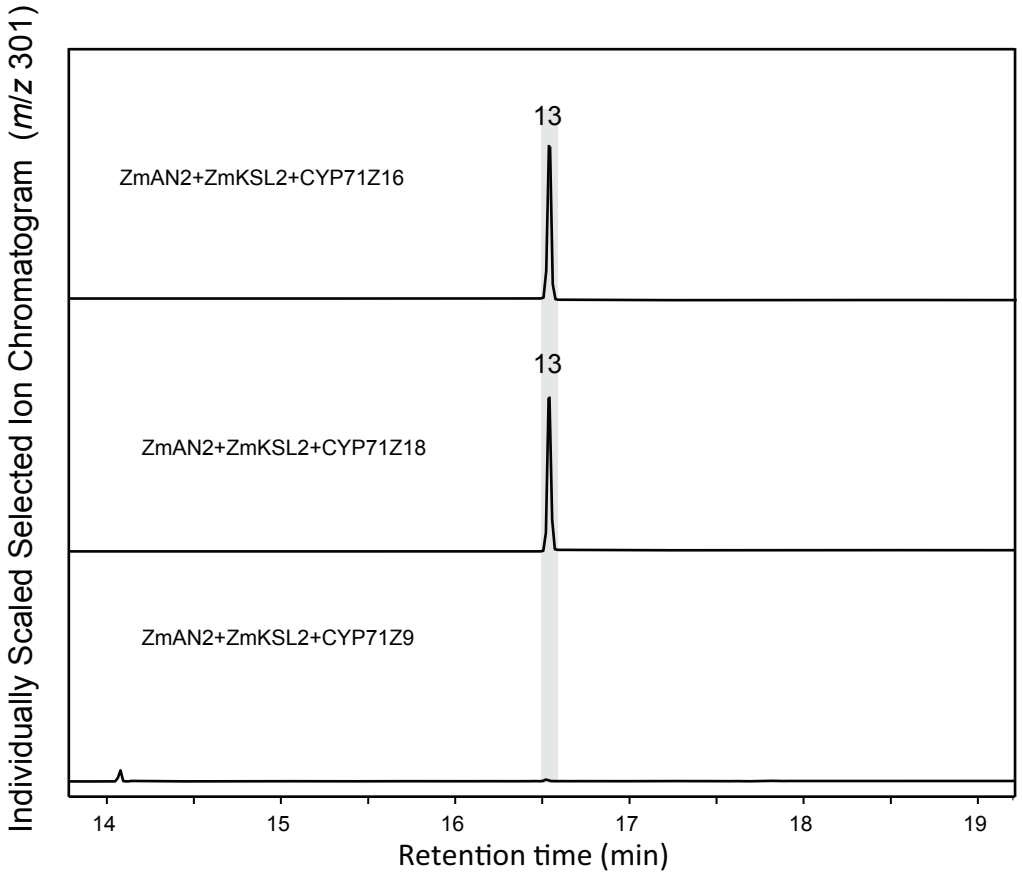


Supplementary Fig. 5. Substrate specificity of ZmKSL2 using paired maize diterpene precursors. (A) Extracted ion chromatograms (EIC) of products resulting from in *E. coli* co-expression of ZmKSL2 with characterized maize class II diTPSs, namely *ent*-CPS (ZmAN2), (+)-CPS (ZmCPS3) and a labda-8,13-dien-15-yl diphosphate (8,13-CPP) synthase (ZmCPS4). Compound **1** is geranylgeraniol (GG), compound **2** is *ent*-copalol, compound **3** is *ent*-isokaurene, compound **4** is *ent*-kaurene (trace abundance), compound **5** is (+)-copalol, asterisk (*) denotes a plasticizer contaminant, compound **6** is an unknown diterpene displaying partial spectral resemblance to **(B)** labda-8(20),12,14-triene (i.e. biformene, CAS 5957-33-5) based on the National Institute of Standards and Technology Database**,** compound **7** is manoyl oxide, compound **8** is an unidentified labda-13-en-8-ol or 8,13-CPP derivative, compound **9** is 8,13-copalol, compound **10** is labda-13-en-8-ol. (**C**) EI-mass spectra of all compounds in panel **A**. Four independent experiments were performed and showed similar results.

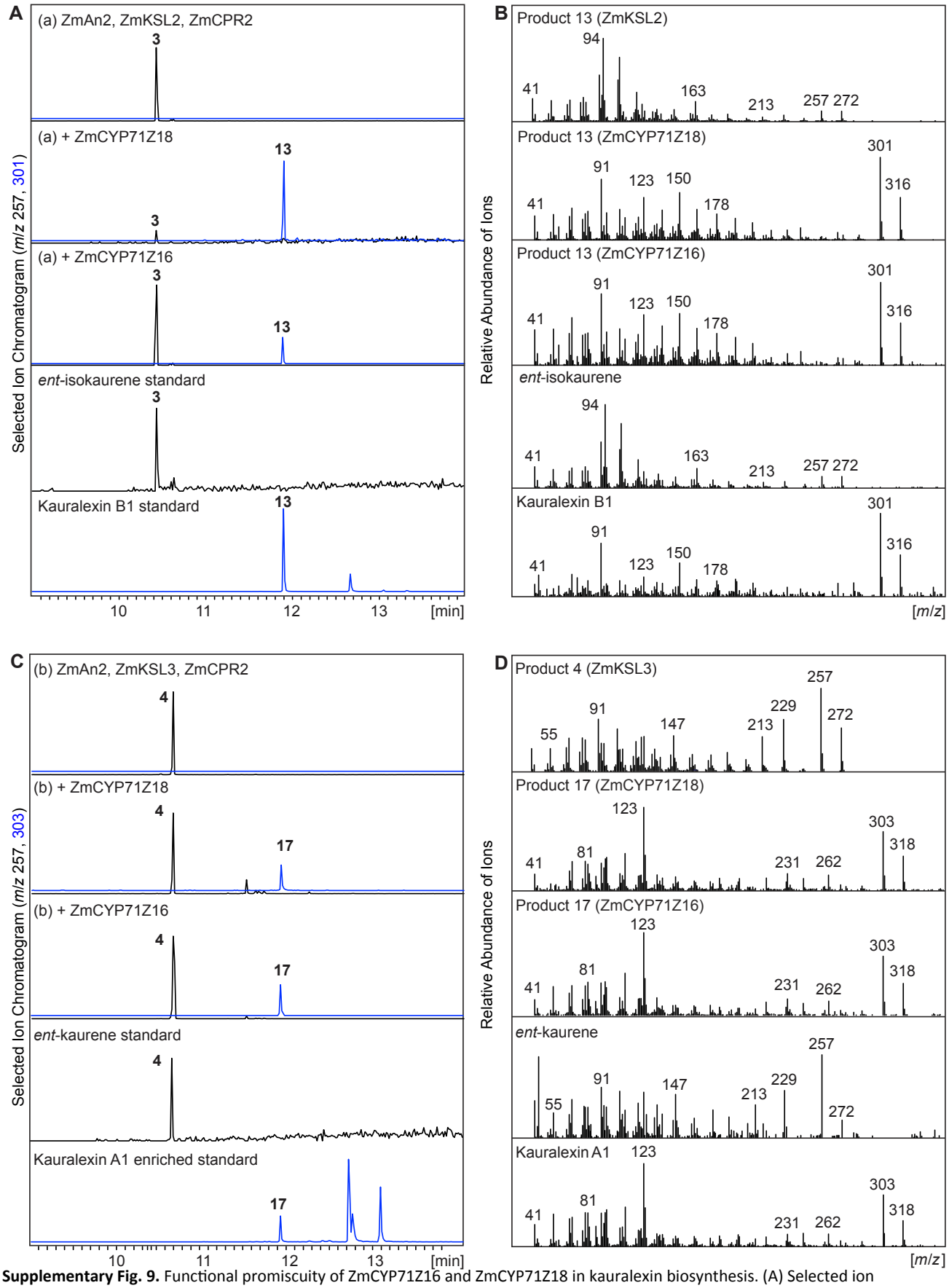
Supplementary Fig. 6. Restriction digest-based discrimination of 2 closely related *ZmCYP71* transcripts reveals the predominant presence of Zm*CYP71Z18*. (A), DNA sequence alignment of Zm*CYP71Z18* and Zm*CYP71Z16* with primer annealing sites for both genes (yellow-highlighted). Both genes contain two exons and one intron (labelled above). The Hind III recognition site (AAGCTT) in Zm*CYP71Z18* is highlighted in green with an arrow denoting the cut site. The Hind III cut site is missing in the Zm*CYP71Z16* amplicon. **(B)**, Photo of gel electrophoresis separation of undigested as well as Hind III-digested RT-PCR product. *ZmCYP71Z18* cDNA amplicon is clipped to yield fragments 217 bp and 159 bp in size. **(C)**, Photo of gel electrophoresis separation of undigested, double digested with Hind III and EcoRV (recognition site, GATATC, highlighted in light-blue in *ZmCYP71Z16*), Hind III-digested, and EcoRV-digested PCR products of B73 genomic DNA. *ZmCYP71Z18* PCR product is clipped to yield fragments 217 bp and 159 bp in size, and *ZmCYP71Z16* PCR product is clipped to yield fragments 213 bp and 163 bp in size. Two independent experiments were performed and showed similar results.



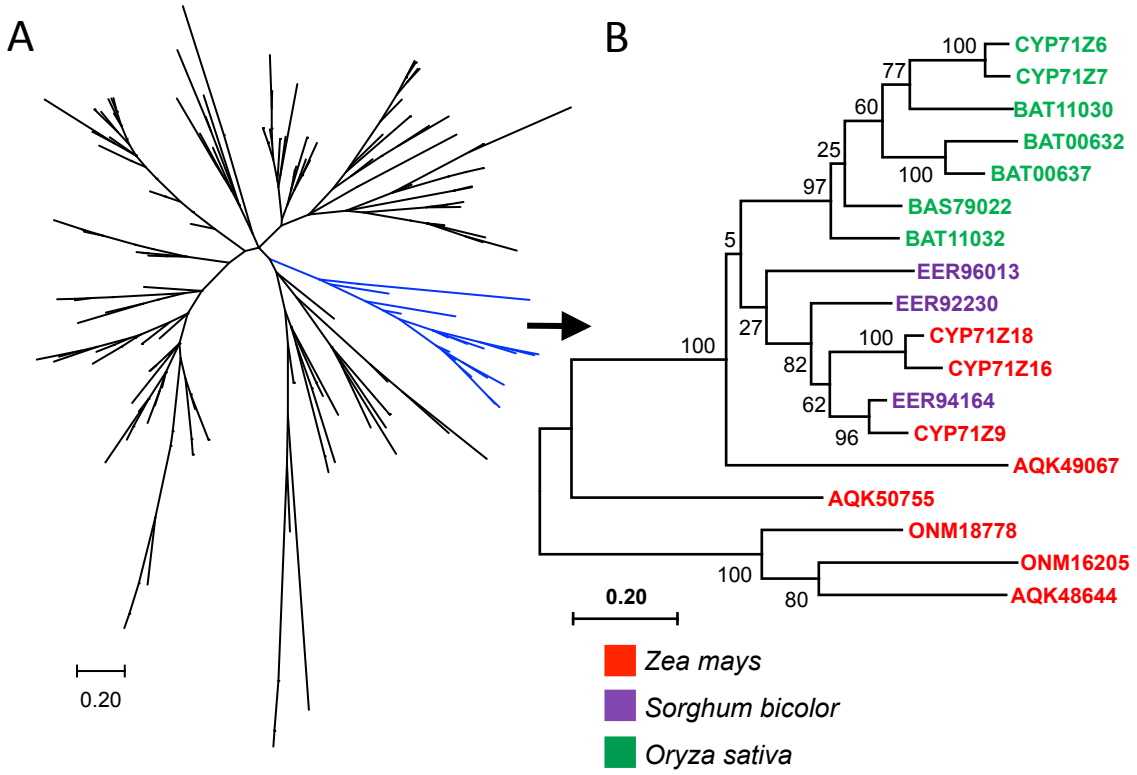
Supplementary Fig. 7. Both ZmCYP71Z16 and ZmCYP71Z18 catalyze the oxidation of *ent*-kaurene to kauralexin A1. (A) Representative GC-MS selected ion chromatograms (SIC) and **(B)** EI-mass spectra of hexane extracts derived from *Agrobacterium*-mediated transient *N. benthamiana* co-expression assays of ZmAN2, ZmKSL3, and ZmCYP71Z16 or ZmCYP71Z18 as compared to authentic kauralexin A1 as a standard. Peaks labeling is referenced to Supplementary Fig. 1 (4, *ent*-kaurene; 17, kauralexin A1). Four independent experiments were performed and showed similar results.



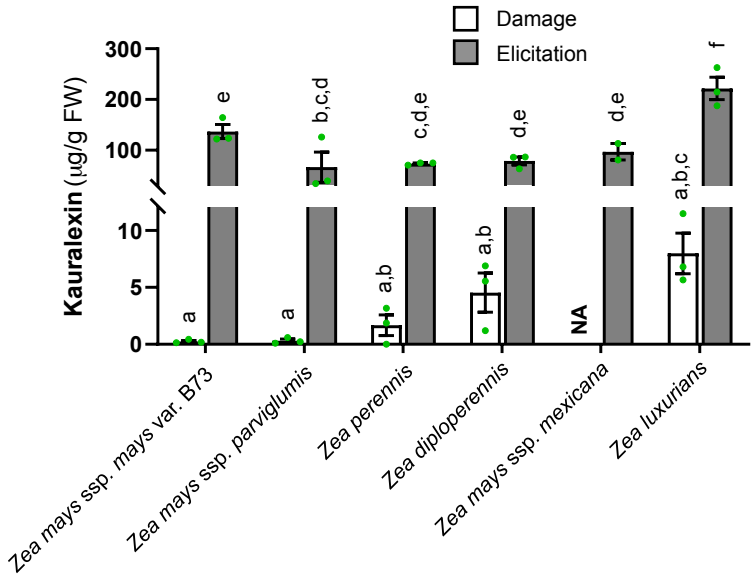
Supplementary Fig. 8. Unlike ZmCYP71Z16 and ZmCYP71Z18, ZmCYP71Z9 does not result in predominant oxidation of *ent*-isokaurene to kauralexin B1. Representative GC-MS selected ion chromatograms (SIC) hexane extracts derived from *Agrobacterium*-mediated transient *N*. *benthamiana* co-expression assays of ZmAN2 and ZmKSL2 with either ZmCYP71Z16, ZmCYP71Z18, or ZmCYP71Z9. Peaks labeling is referenced to Supplementary Fig. 1 (13, kauralexin B1). Four independent experiments were performed and showed similar results.



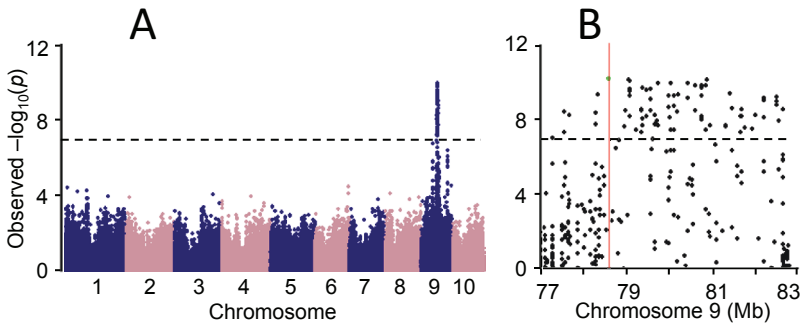
Supplementary Fig. 9. Functional promiscuity of ZmCYP71Z16 and ZmCYP71Z18 in kauralexin biosynthesis. (A) Selected ion chromatograms (SIC; *m/z* 257 for diterpene olefin scaffold, *m/z* 301 & 303 for oxygenated diterpenoids) and (B) EI-spectra of products resulting from in *E. coli* co-expression of ZmCYP71Z16 and ZmCYP71Z18 with ZmAN2, the cytochrome P450 reductase ZmCPR2 and the *ent*-isokaurene synthase ZmKSL2 or (C) identical *E. coli* co-expression combinations that substitute ZmKSL2 for the ent-kaurene synthase ZmKSL3 and (D) corresponding EI-mass spectra of enzyme products and authentic standards. Peaks labeling is referenced to Supplementary Fig. 1 (3, *ent*-isokaurene; 13, kauralexin B1, 4, *ent*-kaurene; 17, kauralexin A1). Four independent experiments were performed and showed similar results.



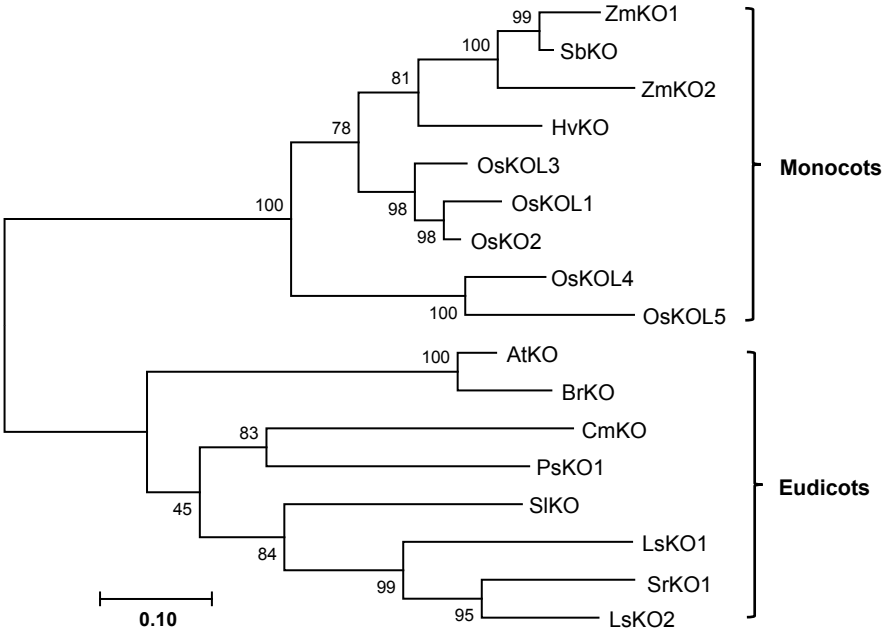
Supplementary Fig. 10. ZmCYP71Z16 and ZmCYP71Z18 belong to a narrow subclade of Poaceous Cyp71 family of P450s. Maximum-likelihood phylogenetic analysis of the CYP71 family from maize, sorghum and rice. **(A)** Members of the CYP71 family from three species of *Zea mays, Sorghum bicolor, and Oryza sativa* with protein sequence identity >40% are included. ZmCYP71Z18 was used a seed query (blue phylogenetic branch). B, Members of the CYP71Z18 subclade from three species of *Z. mays, S. bicolor, and O. sativa.* Bootstrap values calculated from 1000 iterations are indicated at the nodes. All Protein accession numbers for the full tree are listed in Supplementary Table 13.



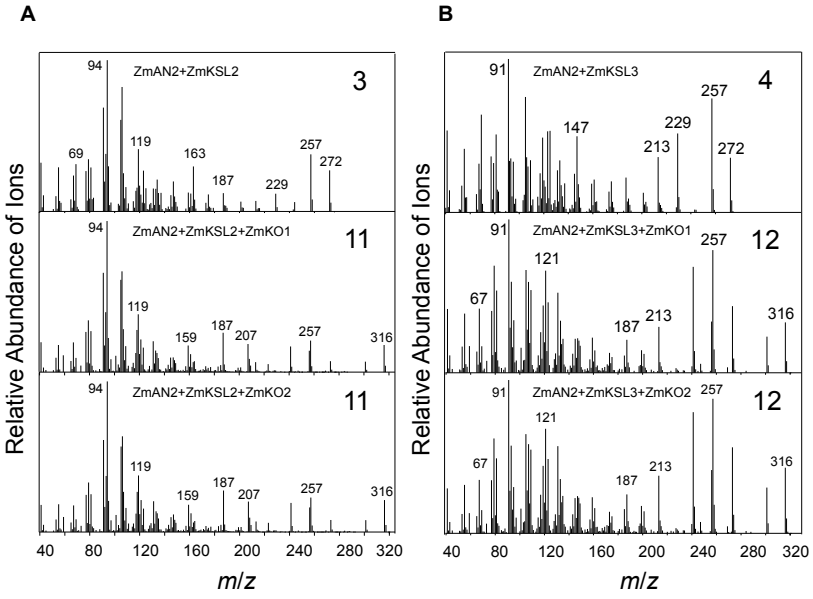
Supplementary Fig. 11. Fungal-elicited kauralexin accumulation occurs in all *Zea* species examined. Average (n = 3, ± SEM) kauralexin levels in stems of 4-week old *Zea mays* ssp. *mays* var. B73, *Z. mays* spp. *parviglumis* (AMES21889), *Z. perennis* (AMES21874), *Z. diploperennis* (PI462368), *Z. mays spp. mexicana* (AMES21851), *and Z. luxurians* (PI422162) treated with a heat-killed *F. venenatum* hyphae. All stem tissues were harvested 5 days after treatment and analyzed by GC-MS. Error bars in the bar charts indicate mean ± SEM (n = 3 biologically independent replicates). Within plots, different letters (a–f) represent significant differences (one-way ANOVA followed by Tukey's test corrections for multiple comparisons, $P < 0.05$). NA represents missing samples/data due to low seed germination rates.



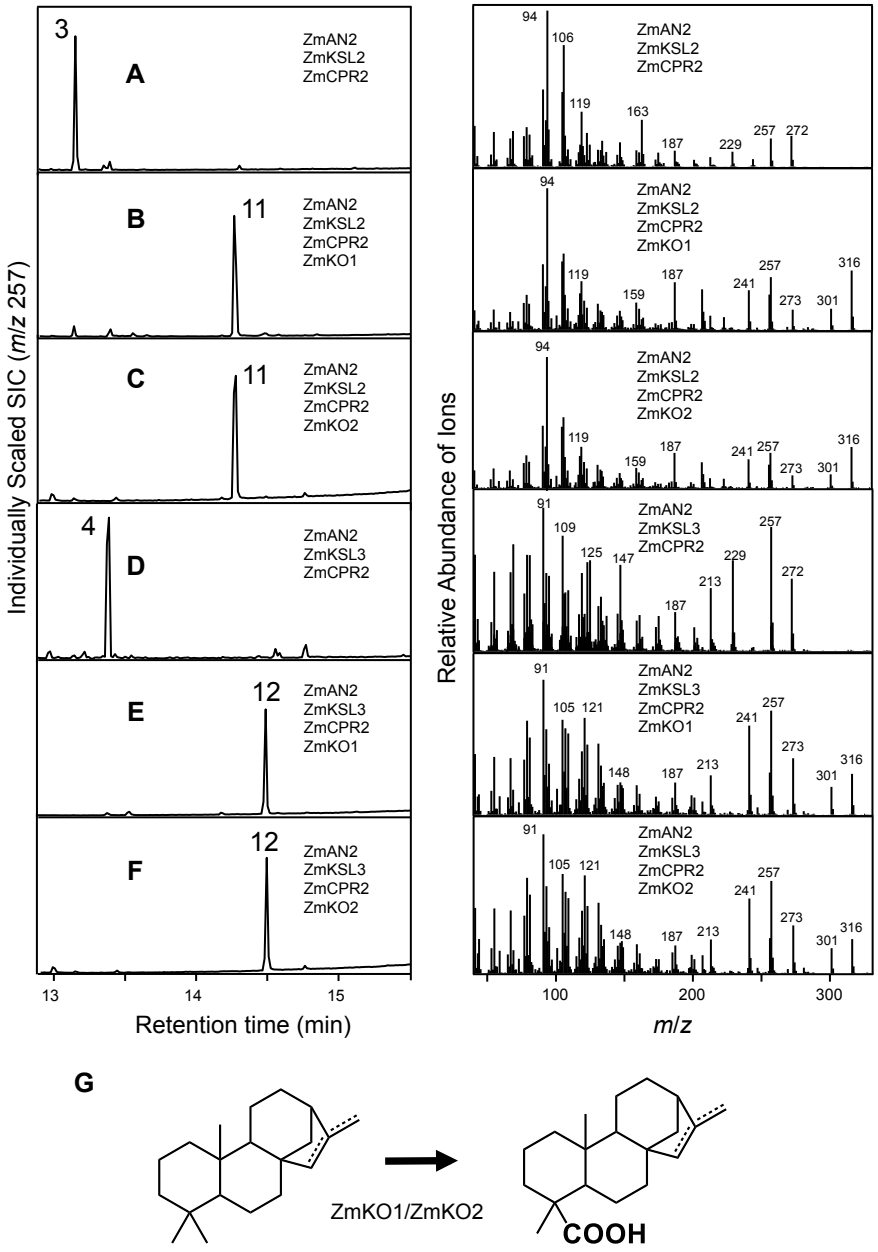
Supplementary Fig. 12. Replicated GWAS supports the role for *ZmKO2* in the production of diterpenoid defenses. **(A)** Second and independently replicated Manhattan plot of the Goodman diversity panel association analysis (258 inbred lines) using the ratio of C-19 oxygenated kauralexins to total kauralexins as a mapping trait. Stem tissue samples of greenhouse grown 35 day old plants were taken 3 days after elicitation with heat-killed *F. venenatum* hyphae. Negative log10-transformed P values from the compressed mixed linear model are plotted on the y axis. Dashed line denotes the 5% Bonferroni-corrected threshold for 246,477 SNP markers with the most statistically significant SNP located at position 79,119,114 (B73 RefGen_v2) on Chromosome 9. **(B)** Local Manhattan plot surrounding the peak on chromosome 9.



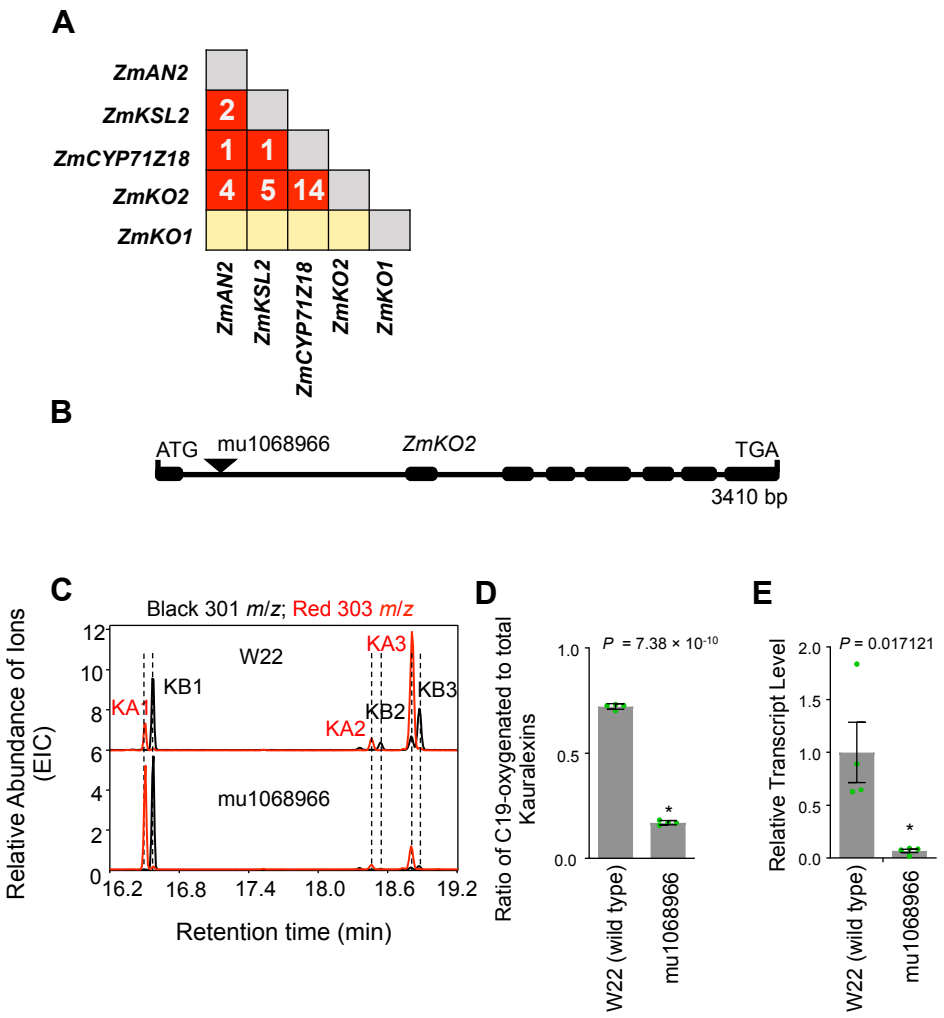
Supplementary Fig. 13. Unlike monotypic kaurene oxidases (KO) present in many plants, maize contains both ZmKO1 and ZmKO2 consistent with a gene duplication event. Maximum likelihood phylogenetic tree of kaurene oxidase (KO) family from four monocot species (*Oryza sativa*, *Sorghum bicolor, Zea mays, and Hordeum vulgare*), and seven dicot species (*Solanum lycopersicum*, *Brassica rapa, Arabidopsis thaliana, Stevia rebaudiana, Cucurbita maxima, Pisum sativum,* and *Lactuca sativa*). Bootstrap values calculated from 1000 iterations are indicated at the nodes. Protein accession numbers are listed in Supplementary Table 13.



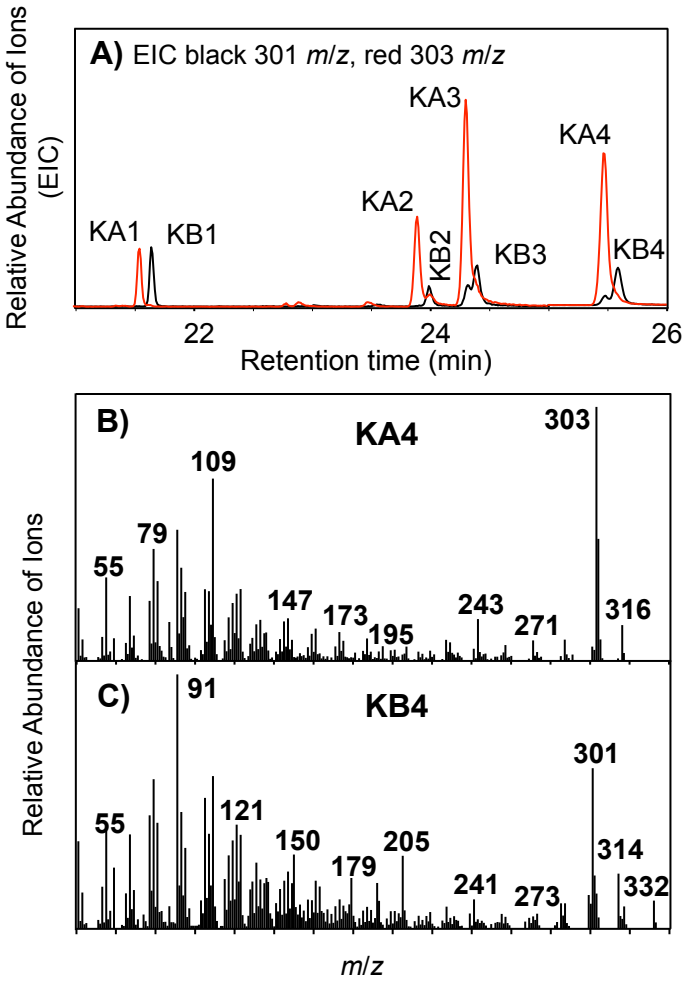
Supplementary Fig. 14. EI spectra confirm the catalytic promiscuity of ZmKO1 and ZmKO2 that results in the C19 oxidation of both *ent*-kaurene and *ent*-isokaurene. Corresponding EI mass spectra of predominant analytes present in **Fig. 3e**-**f** corresponding to *Agrobacterium*-mediated transient *N. benthamiana* co-expression assays of **(A)** ZmAN2 and ZmKSL2 paired with either ZmKO1 or ZmKO2 (**Fig. 3e**) and **(B)** ZmAN2 and ZmKSL3 paired with either ZmKO1 or ZmKO2 (**Fig. 3f**). EI-spectra labeling is referenced to Supplementary Fig. 1 (3, *ent*-isokaurene; 11, *ent*-isokaurenoic acid, 4, *ent*-kaurene; 12, *ent*-kaurenoic acid). Four independent experiments were performed and showed similar results.



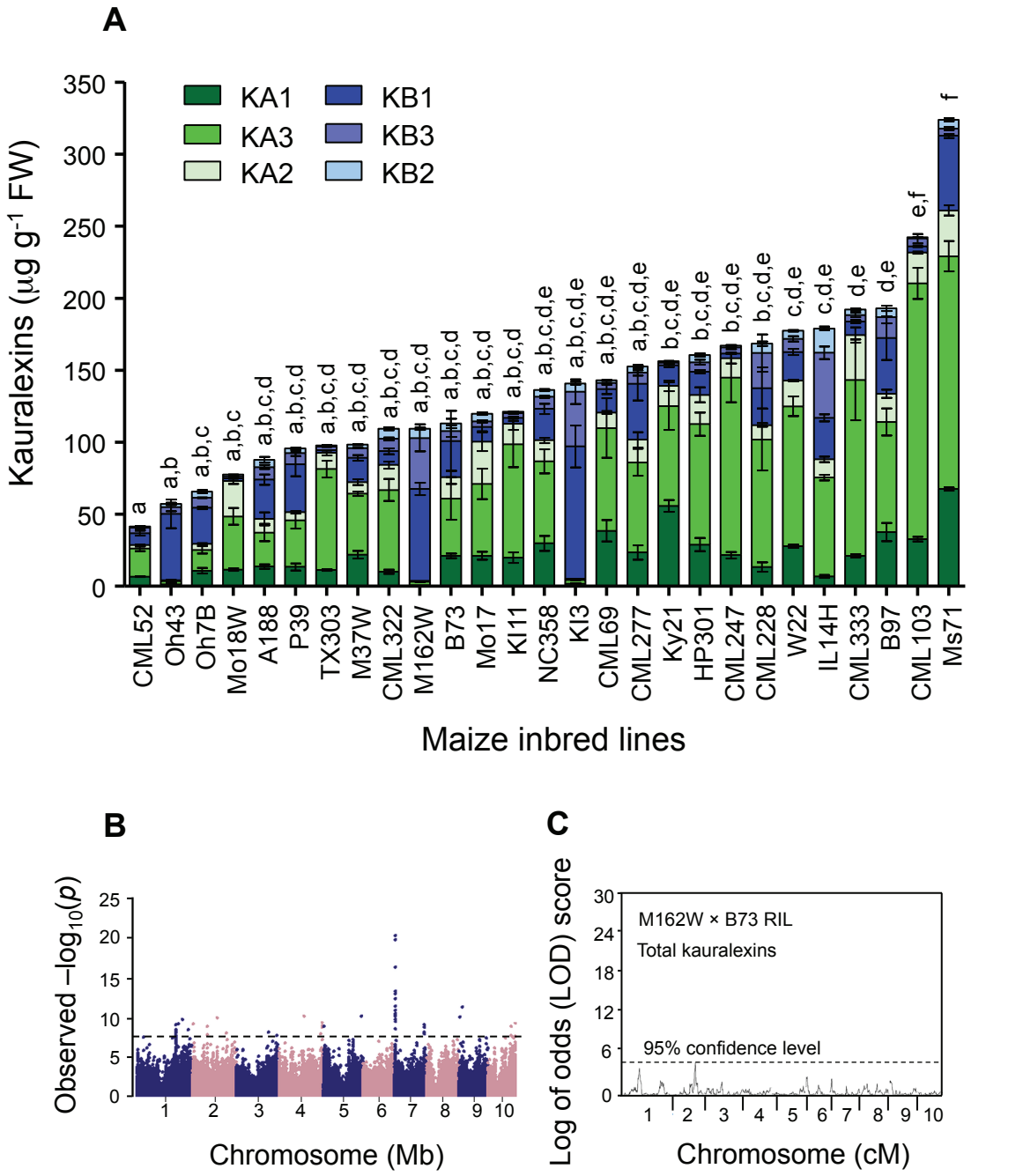
Supplementary Fig. 15. Enzyme co-expression in *E. coli* confirms ZmKO1 and ZmKO2 as kaurene oxidases acting on both *ent*-kaurene and *ent*-isokaurene. (A-C) Gas chromatography-mass spectrometry (GC-MS) total ion chromatograms (TIC) of extracts from *E. coli* engineered for production of *ent*-isokaurene co-expressing proteins in core set 1 (A) ZmCPR2/ZmAN2/ZmKSL2; (B) core set 1 + ZmKO1; and (C) core set 1 + ZmKO2. (D-F) GC-MS TIC of extracts from *E. coli* engineered for production of *ent*-kaurene co-expressing proteins in core set 2 (D) ZmCPR2/ZmAN2/ZmKSL3; (E) core set 2 + ZmKO1; and (F) core set 2 + ZmKO2. TIC product peaks are labelled as follows: 1, *ent*-isokaurene; 2, *ent*-kaurene; 3, *ent*-isokaurenoic acid; 4, *ent*-kauranoic acid. EI-mass spectra of corresponding dominant products in panels A - F are displayed in the adjacent right side panels. Peak labeling is referenced to Supplementary Fig. 1 (3, *ent*-isokaurene; 11, *ent*-isokaurenoic acid, 4, *ent*-kaurene; 12, *ent*-kaurenoic acid). Four independent experiments were performed and showed similar results.



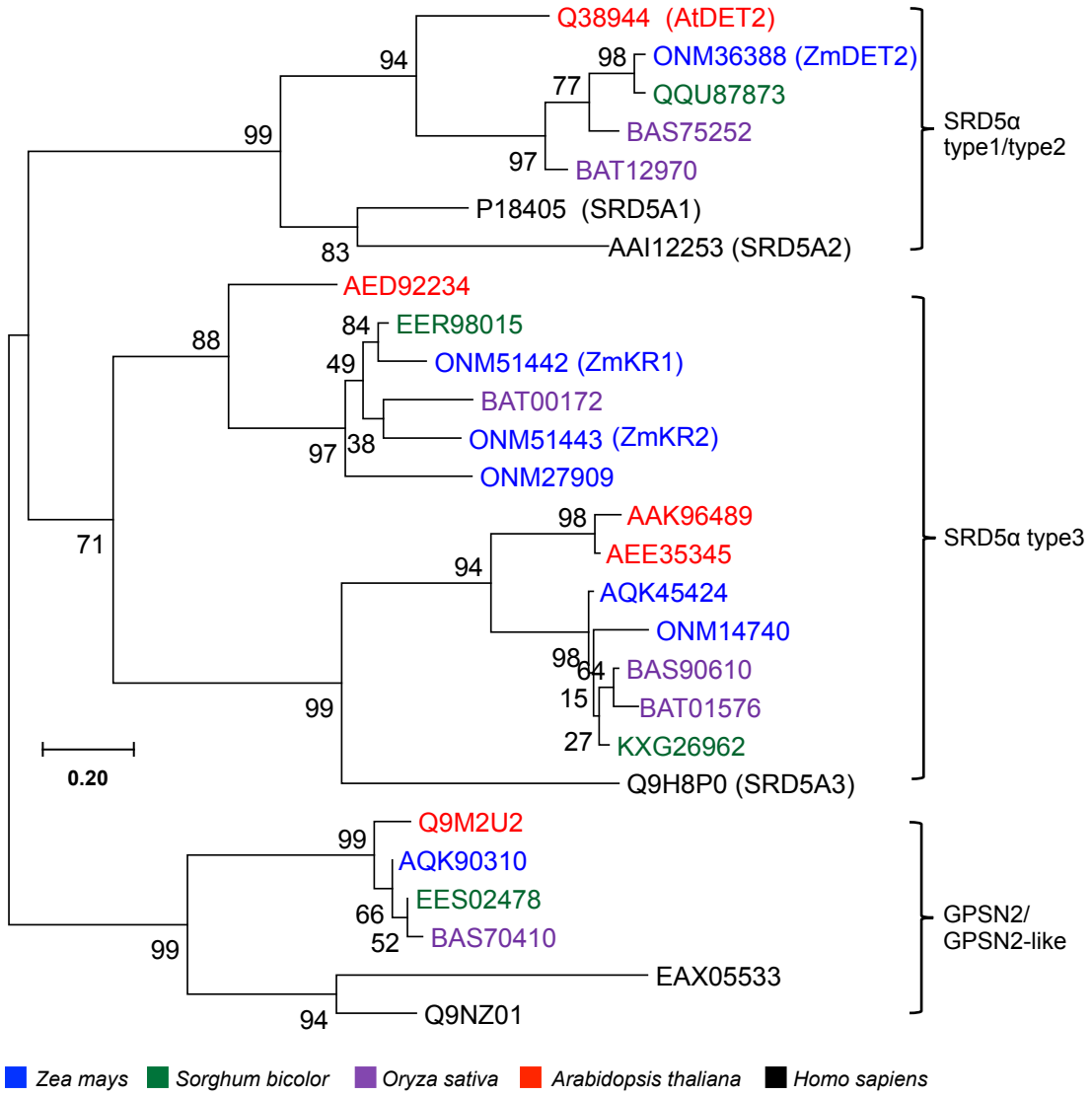
Supplementary Fig. 16. *ZmKO2* transcripts are highly co-expressed with kauralexin pathway genes and an intronic transposon insertion promotes reduced levels of C19-oxidized kauralexins. (A) Heatmap depicting the correlation of coexpression of ZmKO1 and ZmKO2 with kauralexin biosynthetic genes in an RNA-seq dataset of 2094 samples. Numbers in squares represent mutual rank (MR) scores while empty squares represent MR scores > 250. **(B)** Schematic of the analyzed Uniform Mu transposon insertion (mu1068966) present in the first intron of *ZmKO2* from a collection derived from the inbred line W22. **(C)** Representative GC-MS extracted ion chromatogram (EIC) of extracts from *C. heterostrophus* (*C.h.*)-challenged W22 and mu1068966 mutant leaves, Leaves of 60-d-old plants were wounded and treated with either 100 mL of water alone or spore suspensions (1×10^7 ml^{-1}) of *C.h.* for 72 h. Four biological repeats were performed and showed similar results. **(D)** Ratio of the total C-19 oxidized kauralexins to the total kauralexins in *C.h.*-challenged W22 and Uniform Mu (mu1068966) leaves at 72 h. **(E)** qRT-PCR transcript levels of *ZmKO2* in *C.h.*-challenged leaves of W22 and mu1068966 at 24 h with normalization relative to the *ZmEF-1α* reference control gene. Error bars in the bar charts (D and E) indicate mean ± SEM (n = 4 biologically independent replicates). All asterisk (*) denote significant differences ($P < 0.05$) using Student t-tests (two-tailed distribution, unpaired.)



Supplementary Fig. 17. Kauralexin A4 and B4 are co-occurring members of four A-series and B-series kauralexin pairs assigned by relative GC-MS retention times. (A) Relative DB35-GC retention times and EI-mass spectra of the kauralexin alcohols KA4 (*ent*-kaur-19-ol-17-oic acid) and KB4 (*ent*-kaur-15-en-19-ol-17-oic acid) as methyl ester derivatives are co-present in fungal elicited maize tissue and detectable using a cool-on column injection mode. **(B)** The methyl ester of KA4 lacks an identifiable parent ion (*m*/*z* 334) consistent with a predominant loss of H_2O yielding *m*/*z* 316 as the highest MW ion detectable. **(C)** The methyl ester of KB4 contains the predicted parent ion *m*/*z* 332. In each case, A-series kauralexins share a predominant fragment ion *m*/*z* 303 and likewise B-series kauralexins share the fragment ion *m*/*z* 301. Two independent experiments were performed and showed similar results. **Note:** KA4 and KB4 can readily escape detection due to 2 reasons as follows: 1) typical GC/MS sample introduction methods using split-splitless injection modes for KA4 and KB4 do not yield appreciable detection, and 2) Cool-on column injection mode enables KA4/KB4 detection but only on new pristine GC columns that display dramatic losses in peak resolution with each subsequent injection. Thus despite being present, using existing routine and high-throughput GC/MS methods employed, KA4 and KB4 are neither detected nor includable in analyses. In addition to EI-spectra (B-C), structures of KA4 and KB4 were confirmed by NMR (**Supplementary Table 6).**



Supplementary Fig. 18. *ent*-kaurene associated A-series kauralexins predominate in most maize inbreds and consistently map to a chromosome 7 interval containing a steroid 5α-reductase. (A) kauralexin A-series (3 shades of green: KA1, KA3, KA3) and B-series (3 shades of blue: KB1, KB3, KB2) concentrations in the stem tissues of the 27 diverse inbred lines elicited with heat-killed *Fusarium* hyphae for 3d. Error bars in the bar chart indicate mean ± SEM (n = 3 biologically independent replicates). Within plots, different letters (a–f) represent significant differences (one-way ANOVA followed by Tukey's test corrections for multiple comparisons, $P < 0.05$). **(B)** Manhattan plot summary of the second replicated GWAS mapping trial (230 inbred lines) using the ratio of A-series kauralexins to B-series kauralexins as a trait in the greenhouse grown Goodman diversity panel following 3 days of fungal elicitation. Negative log10-transformed P values from the compressed mixed linear model are plotted on the y axis. Dashed line denotes the 5% Bonferroni-corrected threshold for 246,477 SNP markers with the most statistically significant SNP located at position 7,272,291 (B73 RefGen_v2) on Chromosome 7. **(C)** Unlike kauralexin A-series metabolites (**Fig. 4B**), linkage analysis of total kauralexins in the B73 × M162W RIL population following 3 days of fungal elicitation is not influenced by chromosome 7. The logarithm of the odds (LOD) score profiles with the permutation thresholds indicated by the horizontal lines are shown.

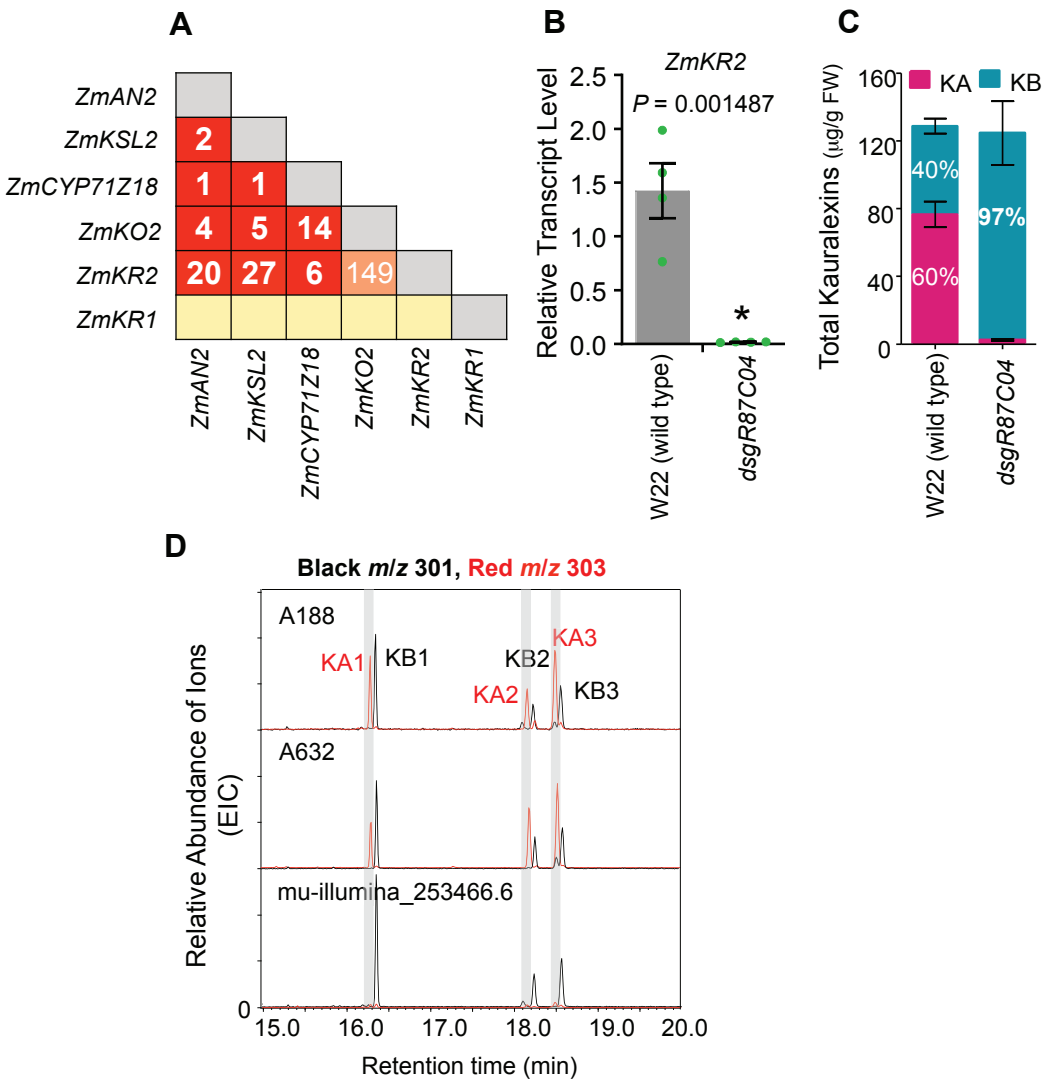


Supplementary Fig. 19. Kauralexin reductase 2 (ZmKR2) is part of a small and high conserved family of plant and animal steroid 5α-reductases responsible for steroid hormone biosynthesis. Maximum likelihood phylogenetic tree of SRD5α family from human (*Homo sapiens*), arabidopsis (*Arabidopsis thaliana*), rice (*Oryza sativa*), sorghum (*Sorghum bicolor*) and maize (*Zea mays*). ZmKR2 specifically resides in SRD5α type 3 family which includes the human protein SRD5A3 (uniprot Q9H8P0) catalyzing the conversion of testosterone to the highly active agonist 5α-dihydrotestosterone (Uemura *et al.*, 2008). Bootstrap values calculated from 1000 iterations are indicated at the nodes. The protein accession numbers are listed in the Supplementary Table 13. Note, AED92234 (GI ID:15237245, At5g16010) was previously assigned to SRD5α type3 subfamily (Langlois *et al*. 2010).

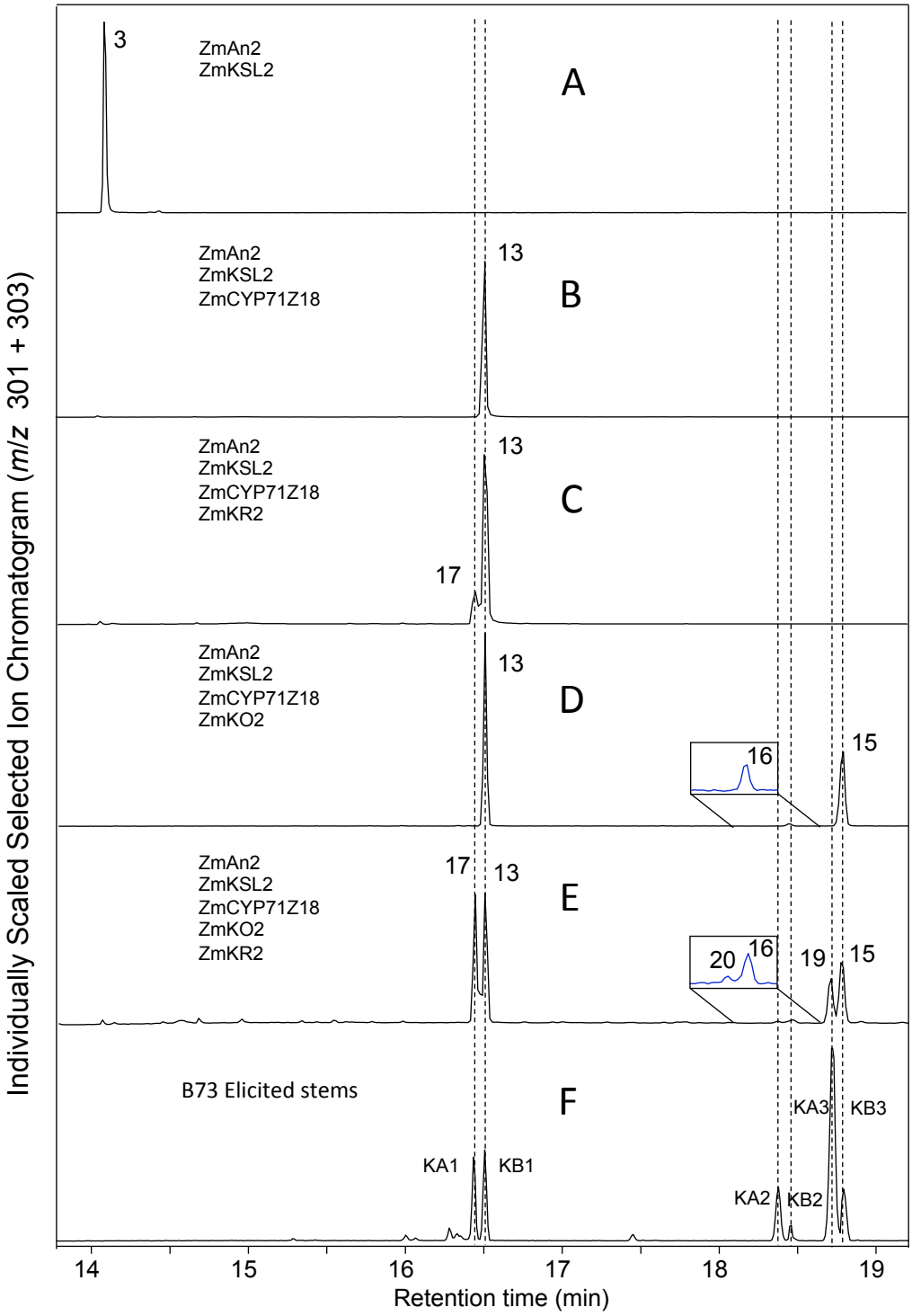
Uemura M, Tamura K, Chung S, Honma S, Okuyama A, Nakamura Y, Nakagawa H (2008) Novel 5 alpha-steroid reductase (SRD5A3, type-3) is overexpressed in hormone-refractory prostate cancer. *Cancer Science* 99:81-86

Langlois VS, Zhang D, Cooke GM, Trudeau VL (2010) Evolution of steroid-5alpha reductases and comparison of their function with 5beta-reductase. *General and Comparative Endocrinology*, 166: 489-497.

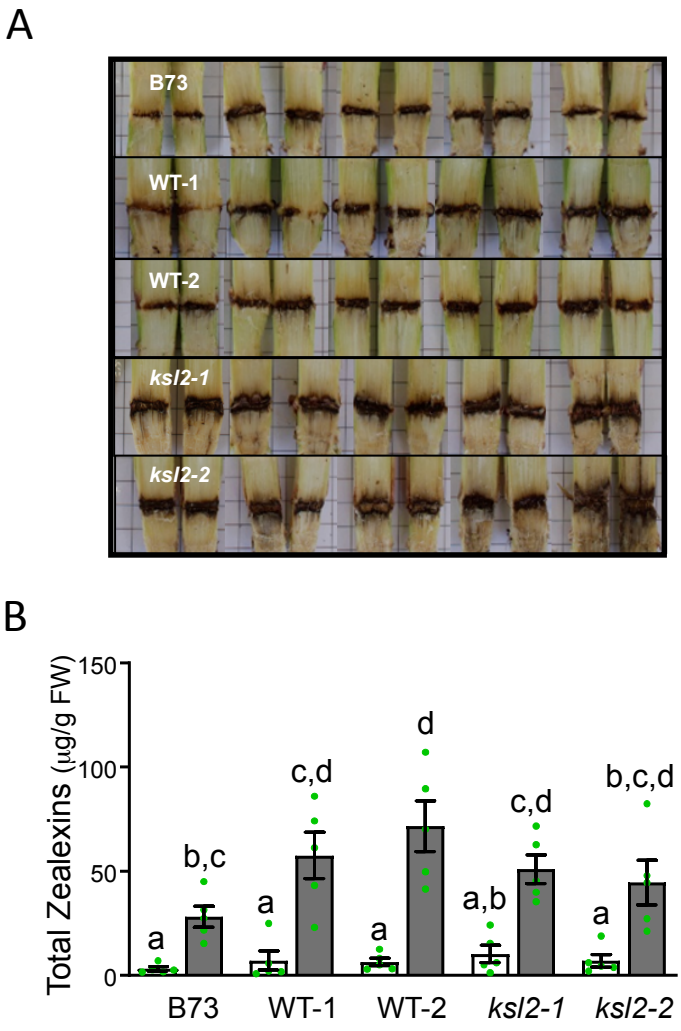
Hartwig T, Chuck GS, Fujioka S, Klempien A, Weizbauer R, Potluri, DP, Choe S, Johal GS, Schulz B. (2011). Brassinosteroid control of sex determination in maize. *Proc Natl Acad Sci U S A*. 2011;108(49):19814–19819.



Supplementary Fig. 20. The steroid 5α-reductase family gene GRMZM2G073929 is co-expressed with kauralexin pathway genes and encodes a functional kauralexin reductase (ZmKR2) based on multiple defined *Zmkr2* mutants. **(A)** Heatmap depicting the mutual rank (MR) correlation of expression of two candidate maize kauralexin reductase genes (ZmKR1/2) with established kauralexin pathway genes using a combined RNA-seq dataset containing 2094 samples. Numbers in squares represent mutual rank (MR) scores while empty squares represent MR scores > 250. **(B)** Abundance of *ZmKR2* transcripts in stems of W22 and the *dsgR87C04 Zmkr2* mutant elicited by heat-killed *Fusarium* hyphae for 48h. Transcript abundance was determined by qRT-PCR relative to the *EF-1α* reference gene. **(C)** Concentration of A- and B-series kauralexins in the stem tissues of W22 and the *dsgR87C04* mutant after 72 h of elicitation with heat-killed *F. venenatum* hyphae. Error bars in the bar charts (B and C) indicate mean ± SEM (n = 4 biologically independent replicates). Asterisk (*) denote significant differences using Student *t*-tests (two-tailed distribution, unpaired) with $P < 0.05$. **(D)** Representative GC-MS chromatograms of A188, A632, and the mu-illumina_253466.6 (*Zmkr2)* mutant after 72 h of elicitation with heat-killed *F. venenatum* hyphae. Four biological repeats were performed and showed similar results.



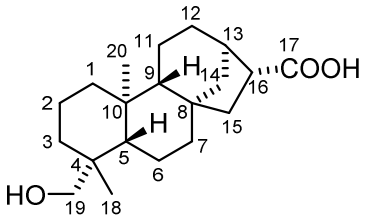
Supplementary Fig. 21. Heterologous co-expression of ZmKR2 *in planta* completes a functional characterization of the core maize kauralexin biosynthetic pathway. (A-E) Gas chromatography-mass spectrometry (GC-MS) select ion chromatograms (SIC) of hexane extracts of *N. benthamiana* leaves transiently expressing combinations of diterpene synthases **(A)** ZmAN2 and ZmKSL2 with **(B)** ZmCYP71Z18, **(C)** ZmCYP71Z18 and ZmKR1, **(D)** ZmCYP71Z18 and ZmKO2, and **(E)** ZmCYP71Z18, ZmKO2 and ZmKR2. **(F)** GC-MS SIC of extracts from B73 stems elicited with heat-killed *F. venenatum* hyphae were used for authentic standards of six kauralexins readily captured by GC/MS analyses. Peak labeling is referenced to Supplementary Fig. 1 (3, *ent*-isokaurene; 13, kauralexin B1; 15, kauralexin B3; 16, kauralexin B2; 17, kauralexin A1; 19, kauralexin A3; 20, kauralexin A2). Four independent experiments were performed and showed similar results.



Supplementary Fig. 22. ***Zmksl2*** **mutants display increased** ***Fusarium graminearum*** **stalk rot yet unaltered levels of inducible acidic sequiterpenoid defenses, namely zealexins. (A)** Disease symptoms in stems of B73, two WT siblings and two *ksl2* mutant lines in response to *Fusarium graminearum.* Photos were taken at 10 days after inoculation with 10 µL of 1.5×10^5 conidia mL^{-1} *F. graminearum*. Two independent experiments were performed and showed similar results. **(B)** Average (n = 5; ± SEM) estimate of corresponding total zealexins (combination of zealexin A1 and B1) present in *F. graminearum* (*F.g*) challenged maize stems. Error bars in the bar chart indicate mean ± SEM (n = 5 biologically independent replicates). Within plots, different letters (a–d) represent significant differences (one-way ANOVA followed by Tukey's test corrections for multiple comparisons, $P < 0.05$).

Supplementary Table 6. NMR analyses confirm the co-occurrence of kauralexin C19 alcohols, namely KA4 (*ent*-kaur-19-ol-17-oic acid) and KB4 (ent-kaur-15-en-19-ol-17-oic acid).

Supplementary Table 6a. Combined NMR spectral data for KA4 in $CDCl_3$ at 600 MHz.

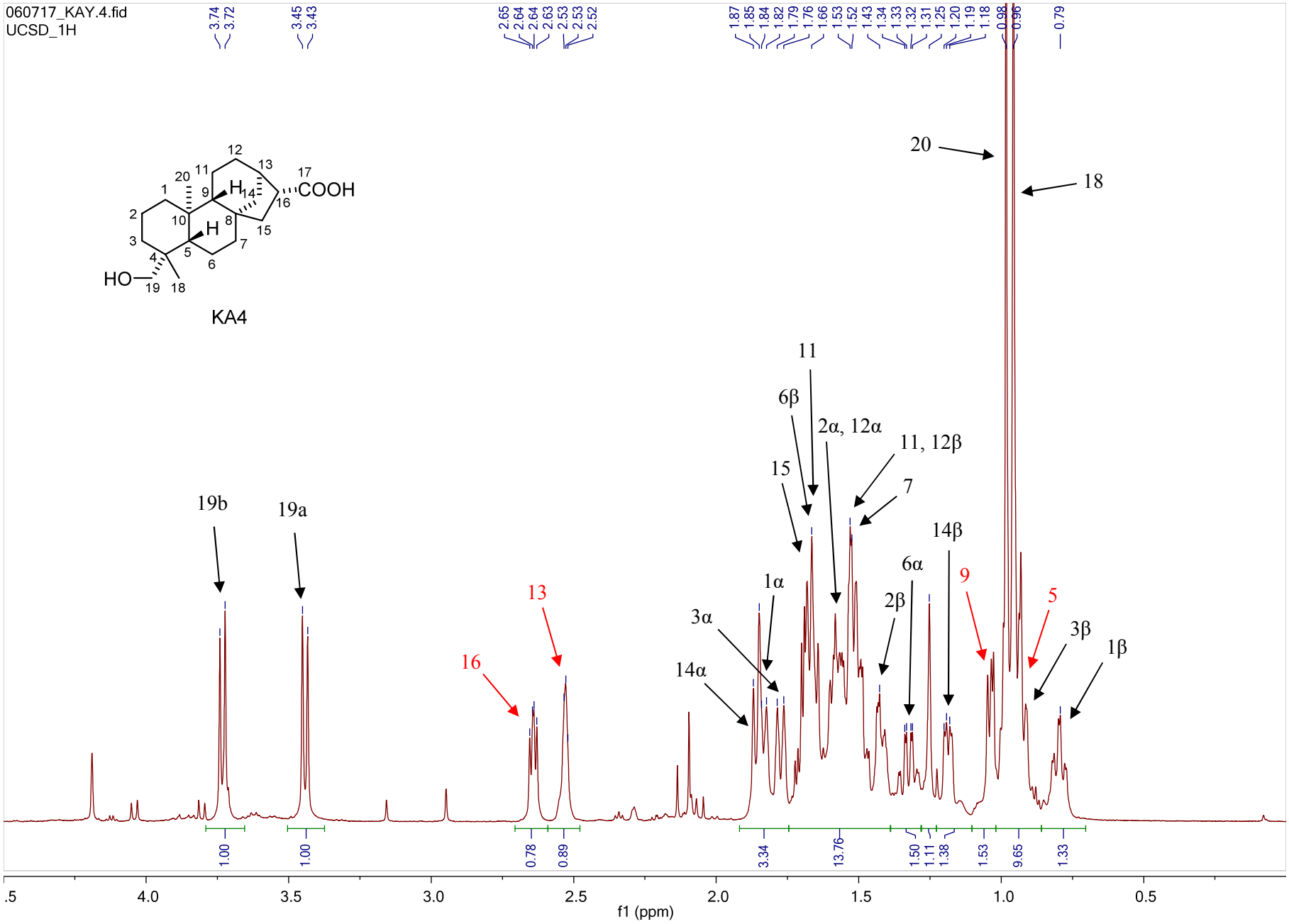


KA4

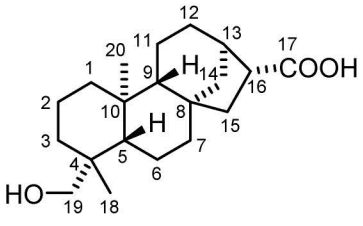
Carbon #[a]	δ_C	δ_H (mult., (*J* (Hz))	COSY	HMBC	NOESY
1	40.6, CH_2	1.83 (d, 13.0) α 0.79 (dd, 13.0, 3.9) β	1β 1α, 2α	2, 9, 10 2, 9, 10	20 5, 9
2	18.3, CH_2	1.57 (overlap) α 1.42 (m) β	1β	1, 3, 10 1, 3, 10	19b, 20
3	35.7, CH_2	1.78 (d, 13.6) α 0.93 (overlap) β	3β 3α	 1, 2, 4, 5, 19	19b
4	38.7, C	----	----	----	
5	56.9, CH	0.94 (overlap)	6α	6, 20	9
6	21.1, CH_2	1.32 (dq, 12.7, 3.3) α 1.67 (overlap) β	5, 6β 6α	5, 7, 8, 10	19a 18, 19a
7	41.4, CH_2	1.51 (overlap)		5	
8	45.3, C	----	----	----	
9	56.3, CH	1.04 (d, 7.6)	11a	1, 8, 10, 12, 14, 20	1β, 5, 7, 15
10	39.3, C	----	----	----	
11	18.6, CH_2	1.52 (overlap) a 1.66 (overlap) b	9	10, 12, 13 10, 12, 13	
12	31.3, CH_2	1.59 (overlap) α 1.52 (overlap) β	 13		 16
13	41.5, CH	2.53 (br)	12β, 14β	8, 11, 12, 15, 17	12α, 12β, 14α, 14β
14	38.1, CH_2	1.86 (d, 11.7) α 1.19 (dd, 11.7, 4.5) β	14β 13, 14α	8, 12, 13, 9 8, 12, 15	13, 20 13
15	44.7, CH_2	1.68 (overlap)	16		
16	45.4, CH	2.64 (dd, 8.7, 6.3)	15	12, 13, 14, 15, 17	12β
17	181.8, C	----	----	----	
18	27.2, CH_3	0.96 (s)		3, 4, 5, 7, 19	3β, 6β
19	65.7, CH_2	3.44 (d, 10.9) a 3.73 (d, 10.9) b	19b 19a	3, 4, 5, 18 3, 4, 5, 18	6α, 6β, 20 2α, 3α, 20
20	18.1, CH_3	0.98 (s)		1, 5, 9, 10	1α, 2α, 19a, 19b

[a] Carbon chemical shifts were based on HSQC and HMBC data

^{1}H NMR (600 MHz, $CDCl_3$) of KA4

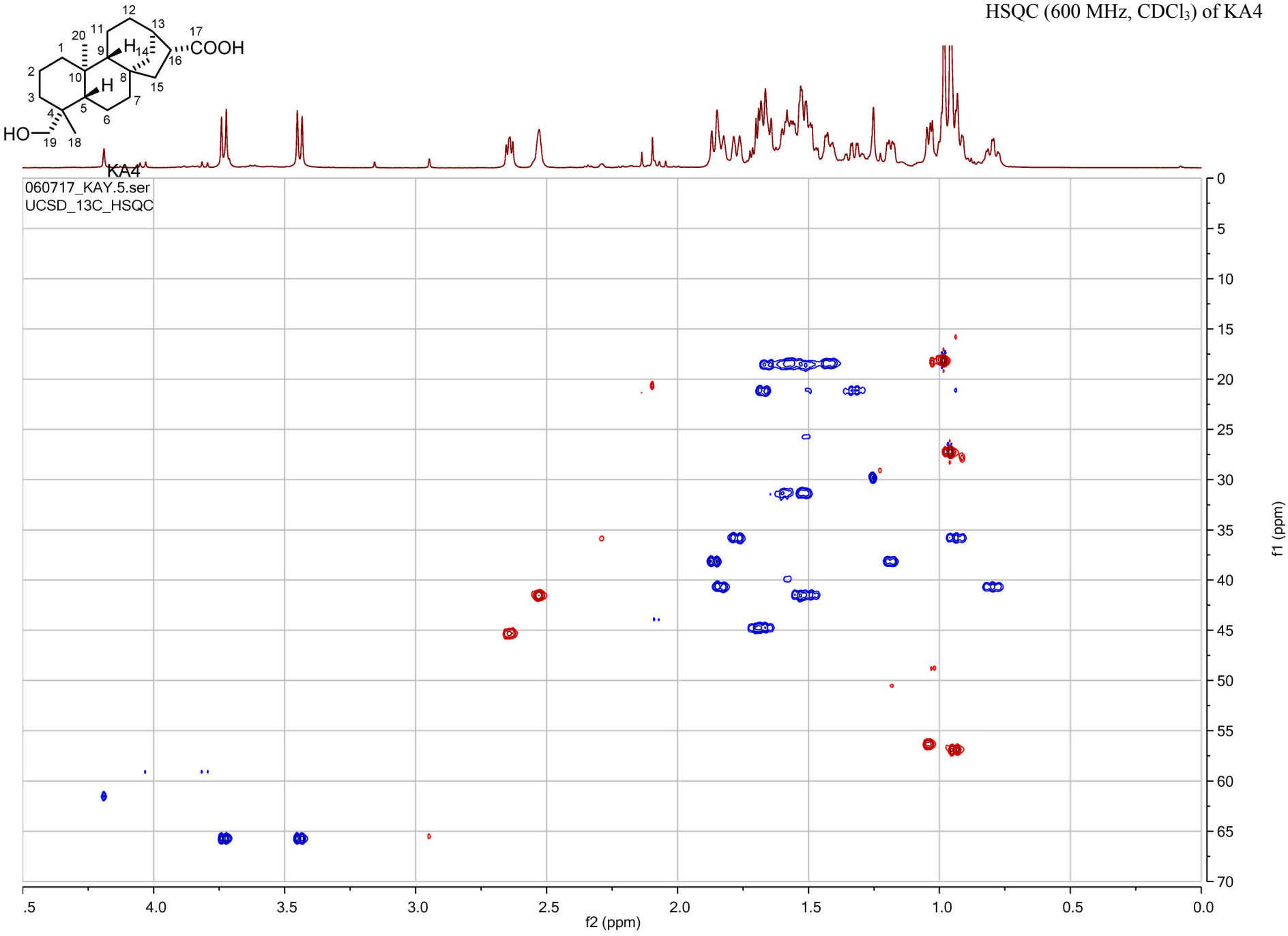


COSY (600 MHz, $CDCl_3$) of KA4

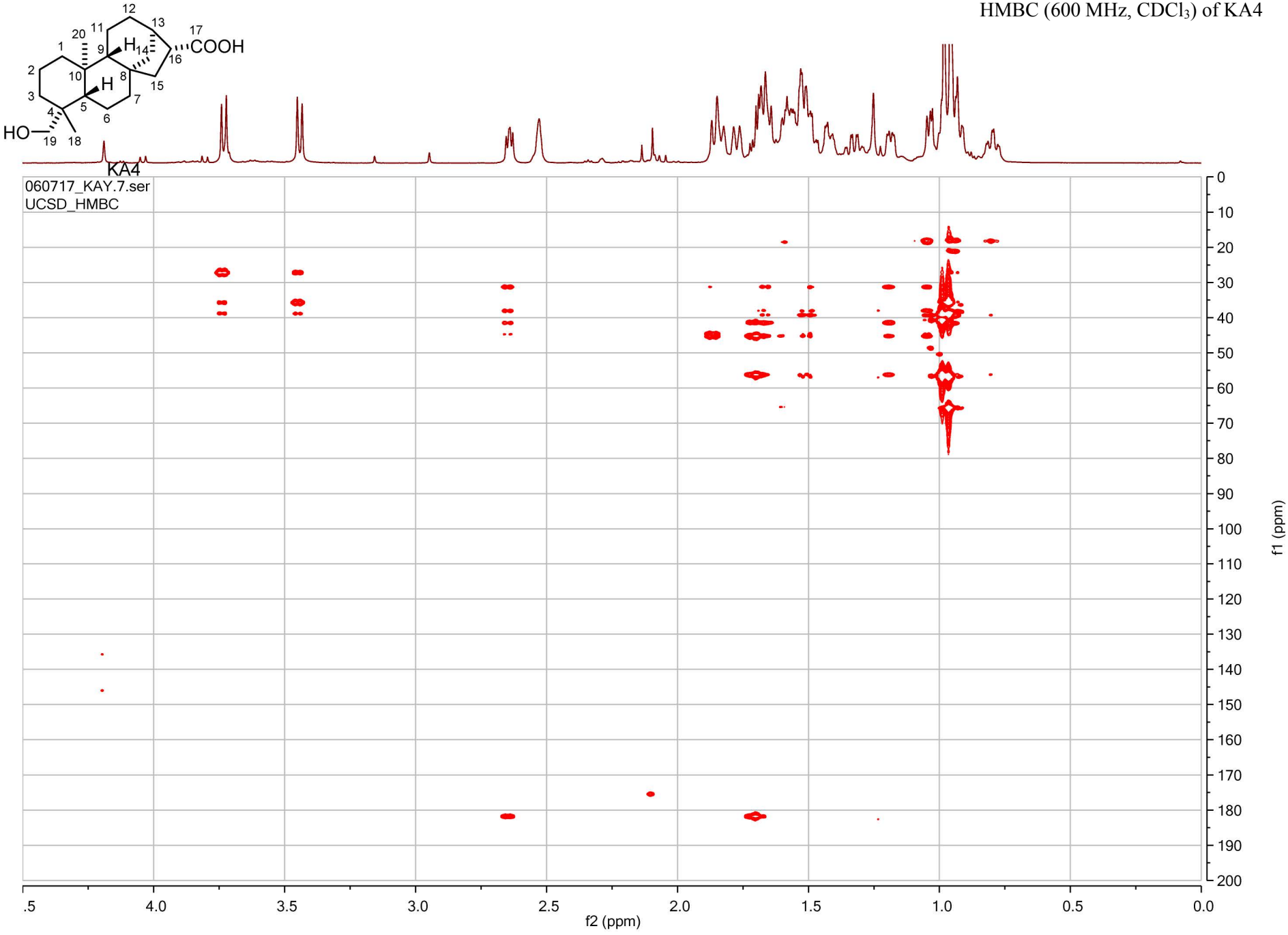


KA4

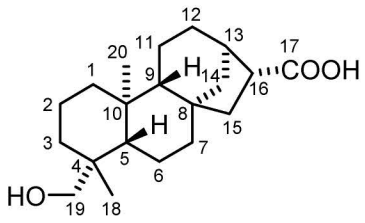
HSQC (600 MHz, $CDCl_3$) of KA4



HMBC (600 MHz, $CDCl_3$) of KA4

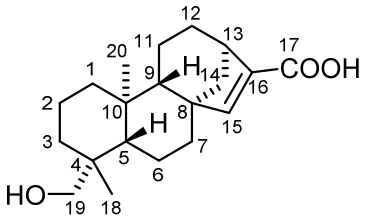


NOESY (600 MHz, $CDCl_3$) of KA4



KA4

Supplementary Table 6b: NMR spectral data for KB4 in $CDCl_3$ at 600 MHz.

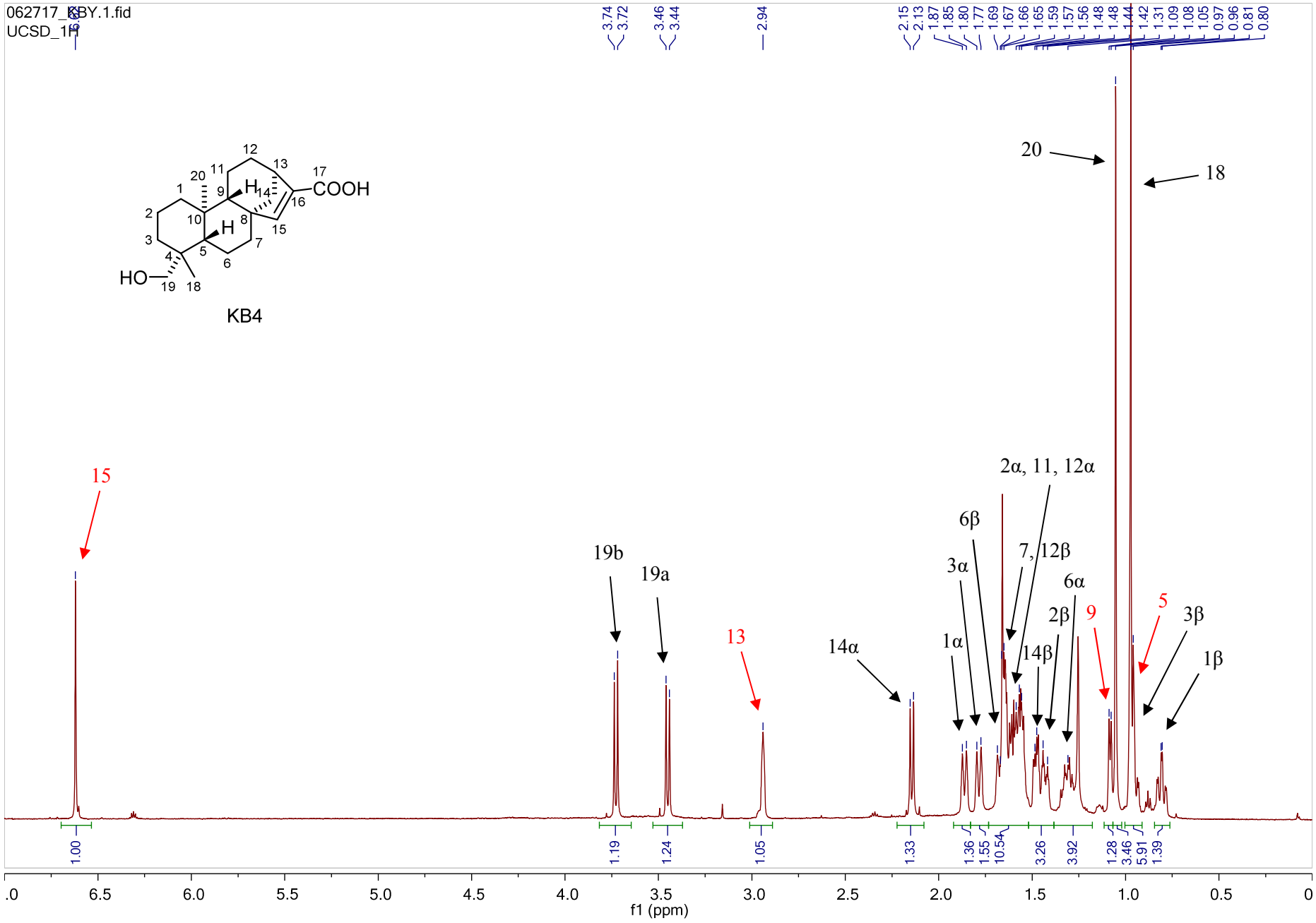


KB4

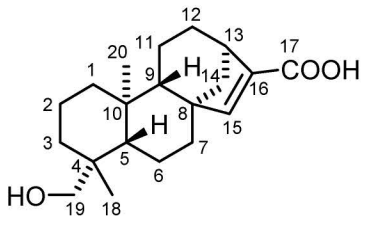
Carbon #[a]	δ_C	δ_H (mult., (*J* (Hz))	COSY	HMBC	NOESY
1	43.0, CH_2	1.86 (d, 13.0) α 0.81 (dd, 13.0, 5.2) β	1β 1α, 2α	2, 3, 4, 5 2, 3, 4, 5, 9	2α, 2β, 20 9
2	18.2, CH_2	1.59 (overlap) α 1.43 (m) β	1β, 3β		1α, 19b, 20 1α, 1β
3	35.6, CH_2	1.78 (d, 13.7) α 0.96 (overlap) β	3β 2α, 3α	4, 5, 6	2β, 19b
4	38.8, C	----	----	----	
5	56.4, CH	0.97 (overlap)	6α		
6	19.0, CH_2	1.32 (overlap) α 1.67 (overlap) β	5, 6β 6α	5, 7, 8, 9, 10	14α, 19a, 20 18, 19a
7	38.6, CH_2	1.66 (overlap)			15
8	50.6, C	----	----	----	
9	46.7, CH	1.08 (d, 6.7)	11	1, 5, 8, 10, 11, 12, 14	1β, 5, 11, 15
10	39.7, C	----	----	----	
11	18.3, CH_2	1.57 (overlap)	9		
12	25.3, CH_2	1.56 (overlap) α 1.65 (overlap) β	 13		13
13	40.4, CH	2.94 (br)	12β, 14β	8, 11, 12, 14, 15, 16	12α, 12β, 14α, 14β
14	43.0, CH_2	2.14 (d, 10.4) α 1.48 (dd, 10.4, 5.2) β	14β 13, 14α	8, 12, 13, 15, 16 7, 8, 12, 13, 15, 16	6α, 13, 20 13
15	155.8, CH	6.62 (s)		7, 8, 12, 13, 14, 16, 17	7, 9
16	137.1, C	----	----	----	
17	168.7, C	----	----	----	
18	27.0, CH_3	0.97 (s)		3, 4, 5	3α
19	65.4, CH_2	3.45 (d, 10.8) a 3.73 (d, 10.8) b	19b 19a	3, 4, 5, 18 3, 4, 5, 18	6α, 6β, 18, 20 2α, 3α, 18, 20
20	18.2, CH_3	1.05 (s)		1, 5, 9, 10	1α, 6α, 11, 14α, 19a, 19b

[a] Carbon chemical shifts were based on HSQC and HMBC data

^{1}H NMR (600 MHz, $CDCl_3$) of KB4

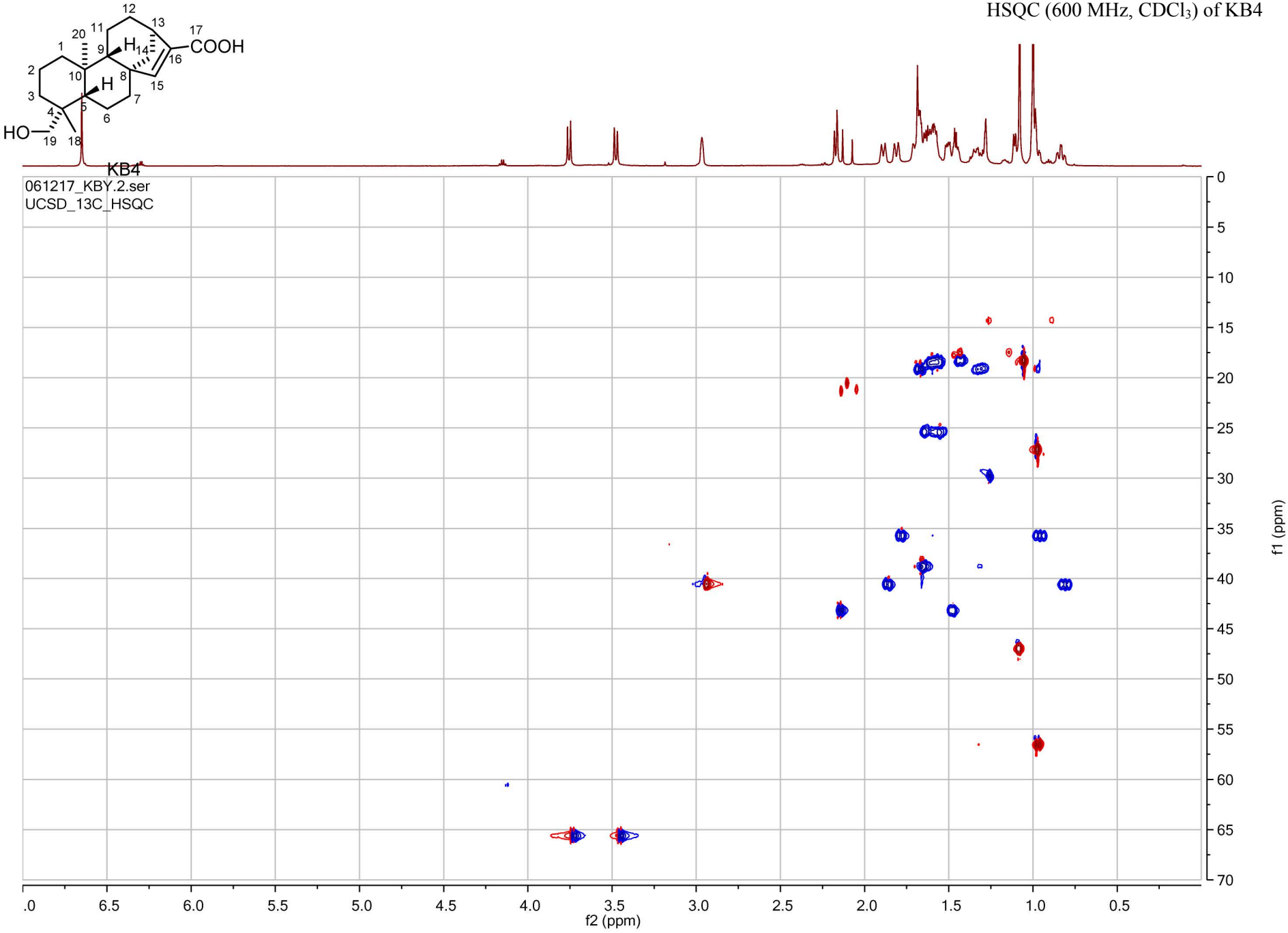


COSY (600 MHz, $CDCl_3$) of KB4

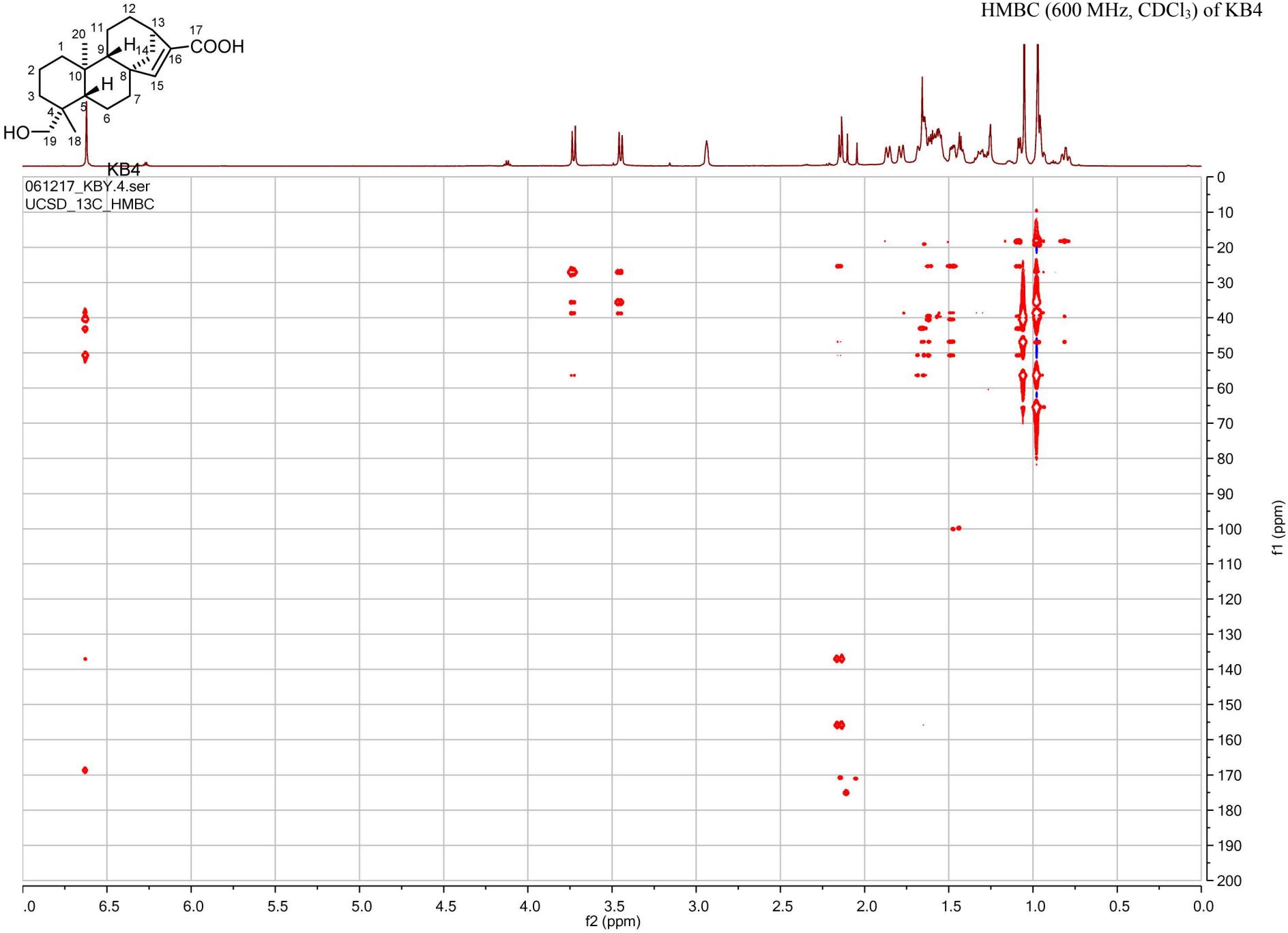


KB4

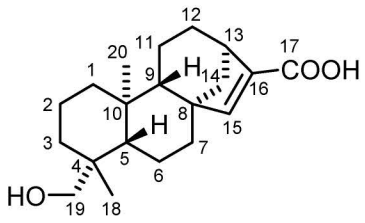
HSQC (600 MHz, $CDCl_3$) of KB4



HMBC (600 MHz, $CDCl_3$) of KB4



NOESY (600 MHz, $CDCl_3$) of KB4

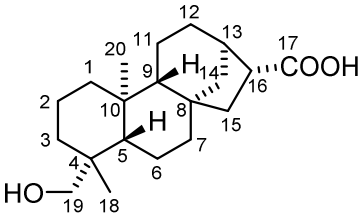


KB4

Optical rotations

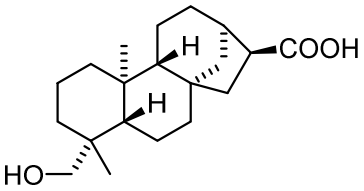
KA4: $[\alpha]_D = -34$ (c = 0.15, CH_2Cl_2)
KA4: $[\alpha]_D = -30$ (c = 0.15, EtOH)

KB4: $[\alpha]_D = -26$ (c = 0.10, CH_2Cl_2,)



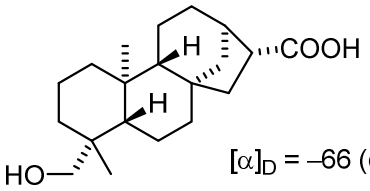
KA4

KNOWN COMPOUNDS:



(ent-kaurane configuration)

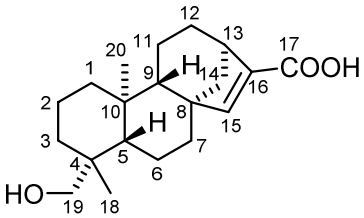
Chen, et al. *Journal of Natural Products* **2000,** *63*, 1000-1003
Hsieh, et al. *Journal of the Chinese Chemical Society* **2004,** *51*, 869-876



$[\alpha]_D = -66$ (c = 3.2, EtOH)

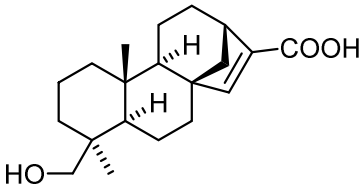
(ent-kaurane configuration)

Henrick, et al. *Australian Journal of Chemistry* **1964,** *17*, 915-933



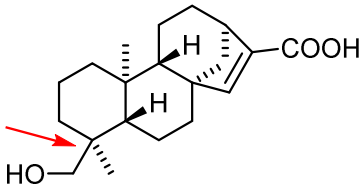
KB4
(relative configuration)

KNOWN COMPOUNDS:



(kaurane configuration)

no publications



pseudolaric acid D
(ent-kaurane configuration)

no publications