Indole-3-glycerolphosphate synthase, a branchpoint for the biosynthesis of tryptophan, indole, and benzoxazinoids in maize

Annett Richter[1], Adrian F. Powell[1], Mahdieh Mirzaei[1], Lucy J. Wang[1], Navid Movahed[1], Julia K. Miller[2], Miguel A. Piñeros[3], Georg Jander[1*]

[1]Boyce Thompson Institute, Ithaca, NY
[2]Plant Biology Section, School of Integrative Plant Science, Cornell University, Ithaca, NY
[3]Robert W. Holley Center for Agriculture and Health, USDA-ARS, Ithaca, NY

*To whom correspondence should be addressed
Georg Jander
Boyce Thompson Institute
533 Tower Road
Ithaca, NY 14850
Phone: 1 (607) 254-1365
Email: gj32@cornell.edu

Key words
maize, *Zea mays*, tryptophan, indole, benzoxazinoid, indole-3-glycerolphosphate, indole-3-glycerolphosphate synthase, metabolic channeling, defense

Abstract
The maize (*Zea mays*) genome encodes three indole-3-glycerolphosphate synthase enzymes (IGPS1, 2, and 3) catalyzing the conversion of 1-(2-carboxyphenylamino)-l-deoxyribulose-5-phosphate to indole-3-glycerolphosphate. Three further maize enzymes (BX1, benzoxazinoneless 1; TSA, tryptophan synthase α subunit; and IGL, indole glycerolphosphate lyase) convert indole-3-glycerolphosphate to indole, which is released as a volatile defense signaling compound and also serves as a precursor for the biosynthesis of tryptophan and defense-related benzoxazinoids. Phylogenetic analyses showed that IGPS2 is similar to enzymes found in both monocots and dicots, whereas maize IGPS1 and IGPS3 are in monocot-specific clades. Fusions of yellow fluorescent protein (YFP) with maize IGPS enzymes and indole-3-glycerolphosphate lyases were all localized in chloroplasts. In bimolecular fluorescence complementation assays, IGPS1 interacted strongly with BX1 and IGL, IGPS2 interacted primarily with TSA, and IGPS3 interacted equally with all three indole-3-glycerolphosphate lyases. Whereas *IGPS1* and *IGPS3* expression was induced by insect feeding, *IGPS2* expression was not. Transposon insertions in *IGPS1* and *IGPS3* reduced the abundance of both benzoxazinoids and free indole. *Spodoptera exigua* (beet armyworm) larvae show improved growth on *igps1* mutant maize plants. Together, these results suggest that IGPS1 and IGPS3 function mainly in the biosynthesis of defensive metabolites, whereas IGPS2 may be involved in the biosynthesis of tryptophan. This metabolic channeling is similar, though less exclusive than that proposed for the three maize indole-3-glycerolphosphate lyases.

Introduction

Grasses, including maize (*Zea mays*), produce a wide variety of constitutive and inducible defense compounds to protect themselves against herbivores and pathogens. Benzoxazinoids, a class of indole-derived antifeedant and insecticidal compounds in maize, provide resistance to numerous arthropod pests (McMullen *et al.*, 2009a, Meihls *et al.*, 2012, Meihls *et al.*, 2013, Tzin *et al.*, 2017, Bui *et al.*, 2018). The core benzoxazinoid biosynthesis pathway is well-studied and involves seven BX enzymes (BX1-BX5, BX8, and BX9) that catalyze the formation of 2,4-dihydroxy-1,4-benzoxazin-3-one glucoside (DIBOA-Glc) from indole-3-glycerol phosphate (IGP) (Figure 1; Frey *et al.*, 1997, Niemeyer, 2009, Wouters *et al.*, 2016). The first reaction in the pathway is catalyzed by BX1, an indole-producing indole-3-glycerolphosphate lyase. Based on global coexpression network analysis, Wisecaver *et al.* (2017) proposed that the committed step in the benzoxazinoid biosynthesis pathway may be upstream of BX1. *IGPS1* (GRMZM2G106950), a maize gene that is co-regulated with BX genes in maize inbred line B73, was predicted to encode indole-3-glycerolphosphate synthase (IGPS), which catalyzes the ring closure of 1-(2-carboxyphenylamino)-l-deoxyribulose-5-phosphate into IGP (Figure 1). Maize contains two additional predicted IGPS genes, *IGPS2* (GRMZM2G169516) and *IGPS3* (GRMZM2G145870), but the expression of these genes was not co-regulated with benzoxazinoid pathway genes (Wisecaver *et al.*, 2017).

To date, the only plant IGPS enzymes that have been characterized are AT2G04400 and AT5G48220 in *Arabidopsis thaliana* (Arabidopsis; Li *et al.*, 1995a). AT2G04400 enzymatic activity was verified by complementing an *Escherichia coli* mutant strain containing a missense mutation in the *trpC* gene, which encodes a bifunctional IGPS/*N*-(5´-phosphoribosyl)-anthranilate isomerase (PRAI) enzyme (Yanofsky *et al.*, 1971, Li *et al.*, 1995a). Based on the complementation assays using *E. coli* mutants deficient in either the IGPS or the PRAI activity of TrpC, it appears that plants only have monofunctional IGPS enzymes (Li *et al.*, 1995a, Li *et al.*, 1995b). Indole-3-glycerolphosphate synthase activity was also verified in a wheat tissue extract, though the actual enzyme was not functionally characterized (Singh and Widholm, 1974).

Indole-3-glycerol phosphate (IGP) is a precursor for indole and numerous other indole-derived plant metabolites, including tryptophan, indole alkaloids, glucosinolates, benzoxazinoids, and auxin (Rose and Last, 1994, Frey *et al.*, 2009). Arabidopsis has two

confirmed indole-3-glycerolphosphate lyases, TSA1 (tryptophan synthase alpha subunit; AT3G54640) and INS (indole synthase; AT4G02610), which convert IGP to indole. Whereas TSA1 is involved in the tryptophan biosynthesis, as part of a complex with TSB1 (tryptophan synthase beta subunit AT5G54810), INS catalyzes indole synthesis independent of TSB (Zhang *et al.*, 2008). Unlike TSA1, Arabidopsis INS does not have a chloroplast targeting sequence. Thus, the pathways for the biosynthesis of tryptophan and tryptophan-independent indole-containing metabolites may be separated in Arabidopsis subcellular compartments.

Three maize genes, *BX1* (*benzoxazinoneless 1*), *TSA* (*tryptophan synthase α subunit*) and *IGL* (*indole glycerolphosphate lyase*), encode enzymes that convert IGP to indole (Figure 1; Kriechbaumer *et al.*, 2008). TSA-GFP fusion proteins were chloroplast-localized in maize protoplast expression expriements (Kriechbaumer *et al.*, 2008) and assays with spinach chloroplasts expressing maize *BX1* also indicated a chloroplastic location for this protein (Stettner, 1998). A fourth protein, designated as TSA-like, was localized to the cytosol and did not have indole biosynthesis activity *in vitro* (Kriechbaumer *et al.*, 2008). *BX1*, *TSA*, and *IGL* genes can be differentiated by both their expression levels and their relative contributions to the biosynthesis of tryptophan, benzoxazinoids, or free indole. Although *BX1* is constitutively expressed in young maize seedlings, its expression also is induced by caterpillar feeding (Maag *et al.*, 2016, Tzin *et al.*, 2017). Similarly, *IGL* expression is upregulated in response to herbivory, consistent with the observed herbivory-induced increase in maize indole emission (Frey *et al.*, 2000). Mutations in *IGL*, but not *BX1*, led to an indole-deficient phenotype (Ahmad *et al.*, 2011). Conversely, only very low amounts of benzoxazinoids can be detected in *bx1* mutant maize lines (Hamilton, 1964, Melanson *et al.*, 1997, Betsiashvili *et al.*, 2015, Tzin *et al.*, 2017), indicating that the indole produced by IGL and/or TSA is not easily incorporated into benzoxazinoids. Nevertheless, exogenous indole application to *bx1* mutant maize plants leads to complementation of the benzoxazinoid-deficient phenotype (Frey *et al.*, 1997, Ahmad *et al.*, 2011).

Maize TSA forms a complex with TSB1 *in vitro*. Partial purification and proteomic analysis of the 160 kD tryptophan synthase complex from maize leaves identified TSA and TSB proteins, but not BX1, IGL, or TSA-like (Kriechbaumer *et al.*, 2008). To date, no maize *TSA* mutation has been reported, but double mutations in the tryptophan synthase ß-subunit genes *TSB1* and *TSB2* led to a seedling-lethal phenotype with high indole accumulation (Wright *et al.*,

1992). Maize plants with double mutations in *BX1* and *IGL1* are viable, indicating that neither of these enzymes makes an essential contribution to tryptophan biosynthesis (Ahmad *et al.*, 2011)

Similar to IGP catabolism by the *BX1*, *TSA*, and *IGL* protein products, IGP synthesis was predicted to be encoded by three maize genes, *IGPS1*, *IGPS2*, and *IGPS3*. Here we report experiments to confirm the function of maize IGPS enzymes and determine whether they represent an earlier branch point in the indole biosynthesis pathway. Bimolecular fluorescence complementation (BIFC) assays show differences in the interaction profiles. For instance, whereas IGPS1 strongly interacts with BX1 and IGL but not with TSA, IGPS2 has a stronger interaction with TSA. However, the analysis of *IGPS1* and *IGPS3* mutant plants suggests that the metabolic channeling of IGP is not as strong as that of indole produced by BX1, TSA, and IGL.

Results

Maize has three indole-3-glycerol phosphate synthase genes

Amino acid sequence comparisons showed greater identity between maize IGPS1 and IGPS3 (64%) than when comparing IGPS1 *vs.* IGPS2 (52%) and IGPS3 *vs.* IGPS2 (51 %).
We performed a phylogenetic analysis to investigate the evolutionary relationships of the IGPS family in monocots and dicots, including maize, Arabidopsis, *Triticum aestivum* (wheat) *Sorghum bicolor* (sorghum), *Oryza sativa* (rice), *Hordeum vulgare* (barley), *Setaria viridis* (green bristlegrass), *Populus trichocarpa* (poplar), and *Nicotiana attenuata* (coyote tobacco). Figure 2). Maize IGPS2 was in a cluster with both monocot and dicot enzymes, including the two Arabidopsis enzymes with confirmed IGP synthesis activity (Zhang *et al.*, 2008). By contrast, IGPS1 and IGPS3 were in separate clades that contain only monocot enzymes. Phylogenetically, the maize IGPS enzymes were most similar to those of a close relative, *S. bicolor* (Springer et al., 1989).

Aligning the B73 alleles of the three *IGPS* genes with the genomes of the 26 founder lines of the maize Nested Association Mapping (NAM) population (McMullen *et al.*, 2009b; www.gramene.org) showed high sequence conservation among members of this enzyme family, with 96-100% identity in the predicted protein sequence (Supplementary Table S2). The only exception was inbred line M162W, where *IGPS2* was partially aligned to two adjacent genes, which may result from an incorrect annotation of the M162W genome. The chromosome positions of the three *IGPS* genes were conserved in the 26 maize genome assemblies, with the

inbred line Tzi8 being the exception, having a second, almost-identical, *IGPS1* gene in an assembled scaffold that was not associated with a chromosome. This is possibly an assembly error, as none of the other NAM founder lines genomes had any additional predicted *IGPS* genes.

The IGPS enzyme family is known for catalyzing a carbon-carbon ring-forming reaction in which the carbon atoms become covalently bonded (Wierenga, 2001). The three maize IGPS proteins have the characteristic domains of this enzyme class, including a TIM (triosephosphate isomerase)-like beta/alpha barrel domain, which is one of the most common enzyme folds (Supplemental Figure S1; Chan et al., 2017; Nagano et al., 2002). The TIM barrel motif includes a phosphate-binding domain, which is a conserved protein fold consisting of eight α-helices and eight parallel β-strands that alternate along the peptide backbone (Farber and Petsko, 1990, Brändén, 1991, Wilmanns *et al.*, 1991). The well-organized scaffold contains predicted binding sites for indole, phosphate, ribulose/trilose, and 1-(2-carboxyphenylamino)-1-deoxy-D-ribulose-5-phosphate (Supplemental Figure S1).

To confirm the enzymatic activity of IGPS1, IGPS2, and IGPS3, we complemented the tryptophan-auxotrophic phenotype of the *E. coli trpC9800⁻* strain (Yanofsky *et al.*, 1971), which is deficient in the IGPS but not the PRAI enzymatic activity of TrpC. The coding regions of the three maize *IGPS* genes, including the signal sequences, were cloned into *E. coli* plasmid vectors for expression in *E. coli trpC9800*. Whereas the empty vector control strain grew only when the minimal medium was supplemented with tryptophan, all three maize *IGPS* genes complemented the tryptophan-auxotrophic phenotype (Figure 3).

Maize *IGPS1* and *IGPS3* expression is induced by herbivory

Previous experiments showed that 24 h of *S. exigua* caterpillar feeding on maize inbred line B73 resulted in a 100-fold induction of *IGPS1* expression (p = < 0.001, Dunnett's test) and a 5-fold induction of *IGPS3* (p = < 0.001, Dunnett's test), but no significant changes in *IGPS2* (Tzin *et al.*, 2017, Wisecaver *et al.*, 2017; Supplemental Figure S2A). To determine whether the gene expression response is similar in the W22 inbred line, we infested maize seedlings with *S. exigua* for 24 h and measured the transcript levels by quantitative real-time PCR (qRT-PCR). Similar to the previous results with inbred line B73 (Supplemental Figure S2A), this showed significant induction of *IGPS1* (p < 0.001, *t*-test) and *IGPS3* (p < 0.001, *t*-test) transcripts (Figure 4A) and no significant transcriptional changes for *IGPS2*. Methyl jasmonate treatment, which regulates

many insect defense responses in plants (Howe and Jander 2008; Martin et al., 2002; Richter et al., 2015), also increased *IGPS1* and *IGPS3* but not *IGPS2* expression (Figure 4B). We measured *IGPS* expression by 3'RNAseq (Kremling *et al.*, 2018) in a time course with 0 h, 2 h, 8 h, and 24 h of *R. maidis* feeding. *IGPS1* expression showed a significant increase after 8 h of aphid feeding W22 ($p < 0.05$, *t*-test; Figure 4C), which is comparable to the significant induction at 4 h after *R. maidis* infestation in inbred line B73 (Tzin *et al.*, 2015; Figure S2B). Although *IGPS3* expression was not significantly changed in response to *R. maidis* feeding, the Illumina 3'RNAseq read counts of *IGPS3* (in average: 188) in W22 plants were higher than those of *IGPS1* (average: 2.5) and *IGPS2* (average: 21), even in the absence of insect feeding. These results are different from the 3'RNAseq results with inbred line B73, which showed an average of 30 read counts for *IGPS1*, 18 read counts for *IGPS3*, and 12 read counts for *IGPS2* (Tzin *et al.*, 2015)*. Given that benzoxazinoids are both constitutively synthesized and induced by insect feeding, both *IGPS1* and *IGPS3* could function in W22 benzoxazinoid biosynthesis. In contrast to *IGPS1* and *IGPS3*, *IGPS2* expression was unchanged or even slightly decreased by insect feeding in both W22 and B73 (Figure 4; Figure S2; Tzin *et al.*, 2015, Tzin *et al.*, 2017).

Transposon insertions in *IGPS1* and *IGPS3* reduce benzoxazinoid levels

A search of publicly available transposon insertions (McCarty *et al.*, 2005, Settles *et al.*, 2007, Williams-Carrier *et al.*, 2010) identified mu-02540, a likely *IGPS1* knockout line with a *Mu* insertion in the second exon (Chr.7: 82883253 bp to 82883261 bp; maize W22 genome v2). We used quantitative reverse transcriptase-PCR (QRT-PCR) to compare homozygous, heterozygous, and wildtype plants after methyl jasmonate induction (Figure 5A). Heterozygous plants showed lower expression levels than wildtype, and no *IGPS1* transcripts above background levels could be amplified from the homozygous mutant plants. The expression of *IGPS3*, but not *IGPS2*, was significantly increased in the *igps1* mutant line (Figure 5B).

We identified a likely *IGPS3* knockout mutation in the *Mu*-Illumina collection (mutant line 228894.6; chr10: 69445669 - 69446382; Williams-Carrier *et al.*, 2010). Plants were induced with methyl jasmonate to activate *IGPS3* transcription, and qRT-PCR was used to verify homozygous mutants and wildtype plants. *IGPS3* expression was significantly reduced ($p < 0.001$) in the mutant line, but the other two *IGPS* genes showed no significant changes in expression levels (Figure 5C).

To investigate IGPS2 involvement in benzoxazinoid or indole formation, we characterized several maize lines with transposon insertions upstream of the *IGPS2* open reading frame: *Mu*-Illumina_235844.6 (chr.10: 124084078-124084606; v3), *Mu*_Illumina_252733.6 (chr.10: 124083942-124084621; v3), *Mu*_Illumina_226348.6 (chr.10: 124084282-124084696; v3); *Mu*_Illumina_243642.6 (chr.10: 124084273-124084590; v3), UfMu_mu1030978 (Chr.10: 124084219-124084227; v3) and UfMu_mu1013033 (Chr.10: 124084306-124084314; v3) (McCarty *et al.*, 2005, Settles *et al.*, 2007, Williams-Carrier *et al.*, 2010). However, none of these transposon insertions reduced *IGPS2* transcript levels (data not shown). No *Mu* or *Ds* transposon insertions in the coding region of *IGPS2* were available in public collections.

Both heterozygous and homozygous *igps1* mutant plants had significantly lower levels of DIMBOA (2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one), HMBOA (2-hydroxy-7-methoxy-2H-1,4-benzoxazin-3(4H)-one), DIMBOA-Glc (2-O-β-D-glucopyranosyloxy-4-hydroxy-7-(2H)-methoxy-1,4-benzoxazin-3(4H)-one), DIM2BOA-Glc (2-O-β-D-glucopyranosyloxy-4-hydroxy-7,8-dimethoxy-(2H)-1,4-benzoxazin-3(4H)-one) and HM2BOA-Glc (Glc-2-O-β-D-glucopyranosyloxy-7,8-dimethoxy-(2H)-1,4- benzoxazin-3(4H)-one) than wildtype plants after methyl jasmonate treatment (Figure 6A). The abundance of all measured benzoxazinoids was decreased in homozygous *igps3* mutant plants (Figure 6B), though less so than in the *igps1* mutant line. The level of methyl jasmonate-induced indole accumulation was lower in both *igps1* and *igps3* mutant lines (Figure 6C). By contrast, there were no significant changes in the tryptophan content in either of the mutant lines relative to wildtype W22 ($p > 0.05$, *t*-test; Figure 6D,E).

Spodoptera exigua is a generalist lepidopteran herbivore that grows better on *bx1::Ds* and *bx2::Ds* mutant plants, which are almost completely devoid of benzoxazinoids, than on wildtype W22 maize (Tzin *et al.*, 2017). We conducted bioassays to determine whether the benzoxazinoid decrease in the homozygous *igps1* mutant line (Figure 6A) also affects *S. exigua* caterpillar growth. There was a significant increase in caterpillar body mass, with an average weight of 20 mg on homozygous *igps1* mutant plants compared to 13.6 mg average weight on wildtype plants after nine days of feeding ($p < 0.05$, *t*-test; Figure 6F). We did not expect to see a measureable effect on caterpillar growth due to the relatively small decrease in benzoxazinoids in the *igps3* mutant (Figure 6B) and therefore did not conduct experiments with this line.

IGPSs enzymes are localized in maize chloroplasts

In addition to gene expression levels, another factor that can influence the function of IGPS enzymes is their sub-cellular localization. The TargetP 1.1 software program (http://www.cbs.dtu.dk/services/TargetP/), which predicts subcellular localization and cleavage sites for signal peptides, indicated a chloroplast location for all three maize IGPS proteins. To experimentally confirm this, we cloned the *IGPS* open reading frames (B73-alleles) into the pEXSG vector (Bethke *et al.*, 2009), to introduce an *C*-terminal yellow fluorescent protein (YFP) tag. These constructs were transiently expressed in maize protoplasts and analyzed by confocal microscopy. Positive transformed cells containing the *C*-terminal tagged protein showed a YFP signal after excitation with a 514 nm laser. Protoplast transformation of *pEXSG::IGPS1-YFP*, *pEXSG::IGPS2-YFP*, and *pEXSG::IGPS3-YFP* revealed similar YFP signals in maize chloroplasts (Figure 7A), with the YFP fluorescence overlapping the red chlorophyll autofluorescence of the plastids. Previous research with control constructs expressed in protoplasts showed that YFP by itself was located in the cytosol (Richter *et al.*, 2016).

In addition to the IGPS proteins, we determined the sub-cellular localization of IGL, TSA, BX1, and BX2 (Figure 1, Figure 7A). As has been reported previously for TSA-GFP (Kriechbaumer *et al.*, 2008), the protein products of *pENSG::YFP-BX1*, *pENSG::YFP-IGL*, and *pENSG::YFP-TSA* were located to the plastids of transformed maize protoplasts (Figure 7A). By contrast, BX2 carrying an *N*-terminal mCherry tag (Karimi *et al.*, 2016) was co-localized with an endoplasmic reticulum marker, rather than to the chloroplasts (Figure 7B).

Interactions of IGPS and indole-3-glycerolphosphate lyases

To determine whether the three IGPS enzymes interact with BX1, TSA, and IGL, we performed BiFC interaction studies for all nine possible two-protein combinations. *IGPS1*, *IGPS2*, and *IGPS3* were cloned into p2YC (Yang *et al.*, 2007), such that their C-termini were fused to a C-terminal YFP fragment (amino acids 158-238). The open reading frames for *BX1*, *TSA*, and *IGL* were cloned into p2YN (Yang *et al.*, 2007), fusing an *N*-terminal YFP fragment (amino acids 1-159) at the *C*-termini. All constructs were transformed into *Agrobacterium tumefaciens* and infiltrated into *Nicotiana benthamiana* plants in pairs. Two days after infiltration, protoplasts from the infiltrated leaf area were isolated and analyzed by confocal microscopy. Strong YFP signals were localized in the chloroplasts for the combinations of IGPS1-C/Bx1-N and IGPS1-C/IGL-N (Figure 8A). On the other hand, coexpression of IGPS1-C and TSA-N did not result in reconstitution of YFP fluorescence. By contrast, coexpression of IGPS2 with TSA resulted in a

strong fluorescence signal, and weaker signals were observed when IGPS2 was coexpressed with BX1 or IGL (Figure 8B). IGPS3 consistently showed reconstitution of the fluorescence when co-expressed with all three indole-3-glycerolphosphate lyase partners (BX1, IGL, and TSA; Figure 8C). Additionally, by cloning *IGPS2* and *IGPS3* cDNA sequences into p2YN, we were able to test whether IGPS1 could interact with the other two IGPS enzymes. The co-expression of IGPS1 and IGPS3 could be detected, suggesting a possible complex that could bind to the BX1 and IGL proteins (Figure 8D). However, no interaction signal was detected for the IGPS1 - IGPS2 combination.

Discussion

Both complementation of an *E. coli trpC* mutation and analysis of maize transposon insertion lines indicated that maize *IGPS1*, *IGPS2*, and *IGPS3* encode functional indole-3-glycerolphosphate synthases. In a phylogenetic analysis, IGPS2 clustered with previously characterized Arabidopsis IGPS proteins (Li *et al.*, 1995a; Figure 2), suggesting a role in primary metabolism. By contrast, IGPS1 and IGPS3 are in separate phylogenetic clades, consistent with their functions in defensive benzoxazinoid metabolism that is not found in Arabidopsis and other dicots that are included in the dendrogram. It is possible that the IGPS1 and IGPS3 enzyme clades in other monocots (rice, wheat, sorghum, barley and *S. viridis*; Figure 2) have additional functions in the production of defense-related metabolites in these species. A similar phylogenetic analysis of maize indole-synthesizing enzymes showed that BX1 and IGL, which contribute to defensive metabolism, cluster separately from the Arabidopsis and maize TSA proteins (Frey *et al.*, 2000, Dutartre *et al.*, 2012). Together, these observations are consistent with the hypothesis that pathways for the production of defense-related benzoxazinoids and indole arose through the duplication and repurposing of ancestral IGPS and TSA enzymes that contribute to primary metabolism in maize and other plants.

Our results are consistent with the observations of Wisecaver *et al.* (2017), who used global co-expression network analysis to propose that *IGPS1* is part of the benzoxazinoid biosynthesis pathway in maize. Both B73 and W22 inbred lines show a greater than a 100-fold increase of *IGPS1* transcripts in response to caterpillar feeding (Figure 4A, Figure S2A, Tzin *et al.*, 2017). Consistent with this induction, *igps1* maize mutants have lower levels of benzoxazinoids (Figure 6A), and *S. exigua* caterpillars grow better on *igps1* mutant than on wildtype W22 plants (Figure 6F). *IGPS3* did now show up in co-expression studies with *BX*

genes (Wisecaver *et al.*, 2017), perhaps due to the lower level of defense-related induction of gene expression, which may have been a factor in the integrative comparison of multiple gene expression studies.

Unlike in the case of IGPS1, IGPS2, IGPS3, BX1, IGL, and TSA, which localize to the chloroplasts when overexpressed in maize protoplasts, we were not able to verify a chloroplastic location for BX2, a cytochrome P450 that catalyzes the first oxygenation reaction in the pathway from indole to DIBOA-Glc (Frey *et al.*, 1995). Consistent with localization of the BX2 enzymatic activity in microsomes after cell separation (Glawischnig *et al.*, 1999), our protein fusion studies indicated that BX2 is localized in the endoplasmic reticulum (Figure 7B). Given that both *bx1* and *bx2* mutations result in an almost complete loss of benzoxazinoids in maize inbred line W22 (Betsiashvili *et al.*, 2015, Tzin *et al.*, 2015), there must be a mechanism by which the indole that is produced by BX1 in the chloroplast is passed to BX2 in the endoplasmic reticulum. Indole produced by TSA or IGL appears to be in a separate metabolic pool and is not used as a substrate by BX2. There are direct interactions between maize TSA and TSB enzymes (Kriechbaumer *et al.*, 2008), which may prevent the release of free indole in this pathway. However, as BX1 and BX2 are in different cellular compartments (chloroplast and endoplasmic reticulum, respectively), it is harder to imagine a direct interaction that allows passage of indole from one protein to the other, without including free indole that is produced by IGL.

Indole production highly induced in maize after herbivore feeding (Frey *et al.*, 2000, Erb *et al.*, 2015). As in the case of benzoxazinoids, free indole was less abundant in both *igps1* and *igps3* mutant lines. Whereas *bx1* mutations cause an almost complete loss of benzoxazinoids (Tzin *et al.*, 2017), and IGL is primarily involved in indole biosynthesis (Frey *et al.*, 2000), IGPS1 and IGPS3 seem to be involved in both metabolic pathways. Consistent with a function in both benzoxazinoid and free indole biosynthesis, IGPS1 and IGPS3 appear to interact with both BX1 and IGL (Figure 8A,C), and interactions between IGPS1 and IGPS3 (Figure 8D) may also contribute to the overlapping functions of these enzymes. An *igps1 igps3* double mutant would be required to determine whether benzoxazinoids and free indole in maize are derived from these two enzymes alone or whether IGPS2 also functions in the production of defensive maize metabolites.

Among the three maize indole-3-glycerolphosphate lyases, IGPS2 had the strongest interaction with TSA (Figure 8B), which is thought to function in tryptophan biosynthesis

(Figure 1). Consistent with the hypothesis that tryptophan is synthesized primarily by an IGPS2-dependent pathway, neither *igps1* nor *igps3* mutations significantly decreased the tryptophan content of maize leaves (Figure 6D,E). Maize *tsb1 tsb2* double mutations caused a seedling lethal phenotype (Wright *et al.*, 1992), and it has been hypothesized that TSA, the only maize indole-3-glycerolphosphate lyase that interacts with TSB1, is therefore also essential (Kriechbaumer *et al.*, 2008). By extension, if IGPS2 has a primary function in providing IGP to TSA, it may also be essential. When conducting our experiments, we were not able to find any *IGPS2* transposon knockout mutations of *IGPS2* in public collections. However, there are predicted *IGPS2* coding-region insertions in the recently published BonnMu collection (Marcon *et al.*, 2020). If these insertions can be confirmed, it will be possible to determine whether IGPS2 has an essential function in maize metabolism.

BX1, IGL, and TSA channel their metabolic products to particular pathways. The free indole-producing function of maize IGL (Frey *et al.*, 2000, Ahmad *et al.*, 2011) is not provided by the BX1 or TSA enzyme activity. Similarly, a *bx1* knockout phenotype cannot be complemented by the activity of IGL or TSA, only by supplementing free indole (Frey *et al.*, 1997, Melanson *et al.*, 1997). To this metabolic channeling, we can now add the effects of IGPS1, IGPS2, and IGPS3, which have distinct, though partially overlapping, functions in channeling IGP to specific indole pools that are formed by BX1, IGL, and TSA. The protein interactions that we have observed (Figure 8) provide new insight into the regulation of indole metabolism and how, within the chloroplast, separate pools of this metabolite can be used for primary metabolism (tryptophan) and specialized metabolism (volatile indole and benzoxazinoids).

Materials & methods

Plant material and methyl jasmonate treatment

Maize (*Zea mays* L.) seeds from the inbred lines B73, W22, and transposon insertion lines Ufmu-02540, UfMu_mu1030978, UfMu_mu1013033, UfMu_mu1030978 (http://maizecoop.cropsci.uiuc.edu/) and *Mu*-Illumina_228854.6, *Mu*-Illumina_235844.6, *Mu*-Illumina_252733.6, *Mu*-Illumina_226348.6, *Mu*-Illumina_235844.6, *Mu*-Illumina_252733.6, *Mu*-Illumina_226348.6, *Mu*-Illumina_243642.6, UfMu_mu1030978 (Barkan lab, http://molbio.uoregon.edu/barkan/) were germinated in water-filled Petri dishes for three days

and were then transferred into plastic pots filled with a soil mix (0.16 m3 Metro-Mix 360 from Scotts, Marysville, OH, USA; 0.45 kg finely ground lime; 0.45 kg Peters Unimix from Griffin Greenhouse Supplies, Auburn, NY, USA; 68 kg Turface MVP from Banfield-Baker Corp., Horseheads, NY, USA; 23 kg coarse quartz sand, and 0.018 m3 pasteurized field soil). Plants were grown for about 2 weeks in climate-controlled chambers (16 h light:8 h dark cycle, 180 mmol photons m^{-2} s^{-1} light intensity at constant 23 °C and 60% humidity). For methyl jasmonate induction experiments using a previously described method (Richter *et al.*, 2015, Tamiru *et al.*, 2017), the third leaf of each plant was cut off and incubated in 2 ml tap water with 250 μM methyl jasmonate (Sigma-Aldrich; 392707) for 24 h, after which the leaves were snap-frozen in liquid nitrogen and stored at -80 °C until they were used for assays. Control leaves were harvested directly into liquid nitrogen without methyl jasmonate treatment.

Gene expression induction by insect feeding

Maize inbred line W22 gene expression assays after *R. maidis* infestation were performed as described for inbred line B73 by Tzin et al. (2015), but we used 10-day-old seedlings and samples were harvested at different time points: 0, 2, 8, and 24 h. The control plants received empty cages without aphids for 24 h and were harvested and frozen in liquid nitrogen at the same time as the experimental samples. To induce gene expression changes in maize line W22 with *S. exigua* caterpillars, we followed the protocol of Tzin *et al.* (2017). Leaf tissue for analysis was harvested after 24 h of caterpillar feeding.

Total RNA extraction and gene expression analysis

Harvested leaf material was frozen in liquid nitrogen, and ground to a fine powder for RNA isolation with on-column DNA digestion, using the SV Total RNA isolation kit (Promega, Madison, WI, USA). The total RNA concentration was measured with a NanoDrop instrument (2000c; ThermoFisher Scientific Inc., Waltham, MA, USA), and 500 ng total RNA was incorporated for first-strand cDNA synthesis using the M-MLV reverse transcriptase kit (TaKaRa Bio USA, Mountain View, CA, USA), and the library was used as a template for qRT-PCR analysis. Thereafter, 5 μl SYBR Green Mix (Universal SYBR Green Supermix, Bio-Rad Laboratories, Hercules, CA), 1 μl gene-specific forward primer, 1 μl gene-specific reverse primer, 0.5 μl template, and 2.5 μl PCR-grade water were mixed. The PCR reaction was performed in a QuantStudio™ 6 Flex Real-Time PCR system instrument (384-well, Applied Biosystems, Foster City, CA) with an initial incubation at 95°C for 10 min. The following cycle

was repeated 40 times: 95°C for 30 s, 60°C for 15 s, and 72°C for 30 s. Maize adenine phosphate transferase 1 (*APT1*) was used as a reference gene. For each analyzed gene, a cDNA pool from all plants was diluted from 1 times to 80 times to generate a standard curve. With Applied Biosystems software, the ΔC_T (Δ-threshold cycle) for each gene was calculated relative to the reference gene. For the CT values an average from at least three biological replicates was calculated. The primer sets used to amplify the genes are given in Supplementary Table S1.

Transcriptome sequencing and RNAseq data analysis

The transcriptome of the W22 aphid time course experiment was sequenced using the 3´RNAseq method (Kremling *et al.*, 2018). RNA was isolated as described above with a total of 8 biological replicates for each treatment. The purity of all RNA samples was proofed with a NanoDrop2000 instrument (Thermo Scientific). The 3´RNA-seq libraries were prepared from 6 µg total RNA at the Cornell Genomics facility (http://www.biotech.cornell.edu/brc/genomics-facility; Kremling et al. 2018). Trimmomatic version 0.36 (Bolger *et al.*, 2014) was used to remove the first 12 bp from each read. STARaligner v.2.4.2 (Dobin *et al.*, 2013) was used to align the reads against the W22 genome (Springer *et al.*, 2018), with a maximum of one location to map to (–outFilterMultimapNmax 1). Another step in the pre-processing was increasing the region size of each gene by 500 bp to decrease the number of genes mapped to a region with no feature from > 120000 to < 60000 on average across the libraries. For this, we used the slop command from bedtools (https://bedtools.readthedocs.io/en/latest/) with –r 500 (-r the number of base pairs to add to the end coordinate). The program Tablet (https://ics.hutton.ac.uk/tablet/) was used to visualize the annotated genes on each chromosome. HTSeq32 v.0.6.1 were used to obtain gene-level counts from the resulting BAM files (htseq-count -r name -s yes -a 10 -f bam). To normalize the read counts, the CPM (count per million) was calculated by dividing the read counts for each gene by the total counts of the RNAseq sample. Illumina sequencing data were submitted to the Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) as data set SUB8291142 (Submission ID), PRJNA669140 (Bioproject ID).

Isolation and cloning of cDNA

To isolate the open reading frames of *IGPS1*, *IGPS2*, *IGPS3, TSA, IGL, BX1,* and *BX2* from B73 cDNA, specific primers (Supplementary Table S1) were created based on the B73_v4 maize genome. iproof high fidelity polymerase (Bio-Rad) was used for PCR amplification. The

fragments were cloned into the specific vectors mentioned for each experiment and sequenced for their accuracy.

Complementation assays in *E. coli*

The open reading frames of *IGPS1*, *IGPS2* and *IGPS3*, including their respective signal sequences, were cloned into the bacterial expression vector pASKIBA37+ (IBAGmbH, Göttingen, Germany). To amplify these genes, special primers with a *Bsa*I restriction site were created with the ''Primer D´ Signer'' (IBA GmbH) (Supplementary Table S1). After sequencing the constructs, plasmids were used for complementation assays in the *E. coli trpC9800⁻* tryptophan-auxotrophic mutant background. The plasmids were transformed into *trpC9800⁻* competent cells and plated out on 100 µg/ml ampicillin LB-agar plates. To test their activity in the mutant background, a starter overnight culture of 2 ml LB medium with 100 µg/ml ampicillin was prepared for each construct, including the empty vector as a negative control. The next morning, the starter culture was used to inoculate 5 ml of tryptophan-deficient Vogel and Bronner (Vogel and Bonner, 1956) minimal medium containing 2 % dextrose and 100 µg/ml ampicillin. For positive controls, 20 µg/ml *L*-tryptophan was added. All cultures started with an optical density at 600 nm (OD600) = 0.2 in the time-course experiment. OD600 was measured six times over 11 hours, until the bacteria had reached stationary phase.

Indol-3-glycerolphosphate lyase phylogenetic tree

A dendrogram analysis was conducted using MegaX software (Kumar *et al.*, 2018). For the alignment we used the MUSCLE Codon algorithm (gap open −2.9, gap extend 0, hydrophobicity multiplier 1.2, clustering method, upgma). The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model (Jones *et al.*, 1992). The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.7732)). In an analysis involved 27 amino acid sequences, a bootstrap resampling analysis with 1,000 replicates was performed to evaluate the tree topology. All positions with less than 80% site coverage were eliminated, i.e., fewer than 20% alignment gaps, missing data, and ambiguous

bases were allowed at any position (partial deletion option). There were a total of 358 positions in the final dataset. The maize TSA protein was used as outgroup for the phylogeny.

Metabolite analysis *via* LC-MS and data processing

Frozen, powdered leaf material was weighed (~100 mg), three volumes of 100% methanol were added to each sample, and samples were incubated for 45 min shaking at 4 °C. After a 10 min centrifugation step (15,000 x g) the samples were filtered in a 96-well MultiScreen filter plate (0.65 µm, Millipore, Burlington, MA). Reverse-phase liquid chromatography was performed using a DionexUltimate 3000 Series LC system (HPG-3400 RS High-Pressure pump, TCC-3000RScolumn compartment, WPS-3000TRS autosampler) controlled by ChromeleonSoftware (Thermo Fisher Scientific, Waltham, MA) and coupled to Orbitrap Q-Exactive mass spectrometer controlled by the Xcalibur software (Thermo Fisher Scientific). A Titan™ C18 UHPLC Column, 1.9 µm 100 mm × 2.1 mm at 40C and the flow rate of 0.5 ml/min of mobile phases A (H_2O:0.1% formic acid) and B (AcN:0.1% formic acid) was used for the separation of target molecular features. The gradient starting condition was 0% B at 0 min, rising to 95% B in 15 min, then was held for 0.5 min, and followed by 0.5 min re-equilibration to the starting condition. A heated electrospray ionization source (HESI-II) in negative mode was used for the ionization, with the parameters of spray voltage at 3.5 kV, the capillary temperature at 380 C, sheath gas and auxiliary gas flow at 60 and 20 arbitrary units, respectively, S-Lens RF Level 50, and probe heater temperature 400°C. The data were acquired in the m/z range of 100–900, 140,000 FWHM resolution (at m/z 200), AGC target 3e6, maximum injection time of 200ms, in profile mode. Metabolite quantification was estimated with ThermoScientific Xcalibur™ Version 4.1.31.9 (Quan/Processing). For data analysis, the raw mass spectrometry files were converted to mzxml formats using the MSConvert tool (Chambers *et al.*, 2012). The raw mzxml files were further processed using the XCMS (http://metlin.scripps.edu/download/, Tautenhahn *et al.*, 2012) and the CAMERA software packages for R (http://www.bioconductor.org/packages/release/bioc/html/CAMERA.html). Data were analyzed using the XCMS-CAMERA mass scan data processing pipeline (Tautenhahn *et al.*, 2008, Benton *et al.*, 2010, Kuhl *et al.*, 2012).

To quantify tryptophan content of maize leaf samples, a stock solution was made by diluting 1 M solution of *L*-tryptophan (Sigma-Aldrich, St. Louis, MO) in water 1:10 with methanol. Further dilutions were made using 100% methanol and a standard curve was prepared

by LC-MS analysis, as described above, of samples containing to 0.3 to 10 ng/µl tryptophan (Supplemental Figure S3A). The concentration of tryptophan in maize tissue samples (µg/g) was calculated based on comparisons to this standard curve.

Indole quantification by GC-MS

For the quantification of indole, leaf material was harvested, ground with a mortar and pestle in liquid nitrogen, and stored at -80°C until sample preparation. The volatile indole released from 50 mg powdered plant material was collected on a solid-phase microextraction (SPME) fiber (100 µm polydimethylsiloxane, fused silica 24 Ga; Supelco, Bellefonte, PA) for 45 min. Samples were analyzed using an Agilent 7890B gas chromatograph coupled to an Agilent 7000D triple quadrupole mass spectrometer (Agilent Technologies, Santa Clara, CA). Collection of volatiles from frozen and macerated tissue in this manner produces similar results to volatile collection from intact maize leaves (Köllner *et al.*, 2004). To release volatiles from the fiber, an injection temperature of 220°C was used. Helium served as a carrier gas with a flow rate of 2.3 ml min^{-1} and nitrogen as collision gas with a flow rate of 1.5 ml min^{-1}. To separate the volatiles, an EC5-MS column (30 m length, 0.25 mm diameter., and 0.25-µm film; Agilent J&W GC columns) was used under the following conditions: 60°C for 3 min, first ramp of 10°C min^{-1} to 150°C, second ramp of 100°C min^{-1} to 300°C, and final 2-min hold. Indole was identified and quantified by running purified indole (Supelco) as a standard. Indole was dissolved in dimethylsulfoxide (DMSO), diluted in water 1:10 and added to 50 mg powdered leaf tissue from uninduced maize inbred line W22, which contains no detectable indole. A standard curve of indole concentrations ranging from 0.58 µg to 5.8 µg indole added to 50 mg powdered leaf material (Supplemental Figure S3B), was used to quantify indole abundance in methyl jasmonate-induced leaf samples.

Subcellular localization

Maize mesophyll protoplasts were isolated as described by Richter *et al.* (2016), with few modifications. After washing the protoplasts, the cells were transferred into the wash buffer instead of the MMG buffer, with a final concentration of ~ 1-2 x 10^6/ ml.

The open reading frames of *IGPS1, IGPS2, IGPS3, BX1, BX2, TSA* and *IGL* were amplified without stop codons and cloned into pDONR 201 by BP clonase recombination (Invitrogen, Carlsbad, CA), before being transferred with a second recombination reaction (LR, Invitrogen) into the vector YFP-pENSg (Dissmeyer et al, 2017). To isolate high concentrations of each plasmid a Midi preparation (Qiagen, Hilden, Germany) was performed. All of the fusion

proteins were transformed into maize mesophyll protoplasts as described above. After incubating maize protoplasts overnight, they were analyzed using the Leica TCS SP5 spectral imaging system (Leica Microsystems, Wetzler, Germany). The YFP fluorescence was excited with a 514 nm laser and the emission spectra were collected with a hybrid detector (HyD) at 523 nm to 591 nm. The autofluorescence of chloroplasts was excited with an argon laser and detected at 675 to 736 nm. Protoplasts from each transformation were imaged with a 20x water immersion objective. For localization of BX2 in the endoplasmic reticulum (ER), the open reading frame was cloned from pDONR 207 into pK7m34GW by LR Gateway recombination. For this reaction, the destination vector, pDONR, pENTR4-1 containing the 35S double promotor and plasmid pENTR2-3 with an mCherry tag were fused. Finally, the construct 35S:mCherry:BX2 was transformed into *A. tumefaciens* (GV3101) and grown overnight in LB medium containing appropriate antibiotics. Likewise, an ER-CFP (cyan fluorescent protein) marker construct (plasmid #953, Nelson *et al.*, 2007) was transformed into *A. tumefaciens*. After mixing *A. tumefaciens* cultures of both constructs with the same OD600, they were co-infiltrated into *N. benthamiana* leaves. Two days after infiltration, protoplasts were isolated from the infiltrated leaf tissue area and used for microscopic analysis. CFP fluorophore excitation was performed using a 458 nm argon laser and detected at 468 to 496 nm. For co-localization, CFP and mCherry were scanned sequentially.

For interaction studies, the open reading frames of *IGPS1, IGPS2, IGPS3, TSA, IGL,* and *BX1* were cloned without stop codons into the BIFC vectors p2YC and p2YN (Kong *et al.*, 2014). In addition, *IGPS1* was cloned into p2YN to study the interaction with IGPS2 and IGPS3. The constructs used for protein-protein interactions studies were co-infiltrated into *N. benthamiana* leaves, as described previously (DeBlasio *et al.*, 2018). To detect the YFP signal and chloroplast autofluorescence, we used the microscope settings described above.

Caterpillar growth assays

Eleven-day-old homozygous *igps1* mutant and wildtype W22 plants were used for caterpillar feeding assays. *Spodoptera exigua* eggs were purchased from Benzon Research (Carlisle, PA) and kept for 52 h in a 28°C incubator on artificial diet (Beet Armyworm Diet, Southland Products Inc., Lake Village, AR). One 10 x 15 cm organza mesh bag (amazon.com, item B073J4RS9C) containing one second instar *S. exigua* caterpillar was placed on each 11-day-old

maize plant. The bags were gently closed with bobby pins without wounding the plant. After one week of feeding on maize plants, the caterpillars were weighed.

Acknowledgements

References

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

Benton, H.P., Want, E.J. and Ebbels, T.M. (2010) Correction of mass calibration gaps in liquid chromatography-mass spectrometry metabolomics data. *Bioinformatics*, **26**, 2488-2489.

Bethke, G., Unthan, T., Uhrig, J.F., Poschl, Y., Gust, A.A., Scheel, D. and Lee, J. (2009) Flg22 regulates the release of an ethylene response factor substrate from MAP kinase 6 in Arabidopsis thaliana via ethylene signaling. *Proc Natl Acad Sci U S A*, **106**, 8067-8072.

Betsiashvili, M., Ahern, K.R. and Jander, G. (2015) Additive effects of two quantitative trait loci that confer *Rhopalosiphum maidis* (corn leaf aphid) resistance in maize inbred line Mo17. *J Exp Bot*, **66**, 571-578.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114-2120.

Brändén, C.-I. (1991) The TIM barrel—the most frequently occurring folding motif in proteins: . *Curr Opin Struct Biol*, **1**, 978-983.

Bui, H., Greenhalgh, R., Ruckert, A., Gill, G.S., Lee, S., Ramirez, R.A. and Clark, R.M. (2018) Generalist and specialist mite herbivores induce similar defense responses in maize and barley but differ in susceptibility to benzoxazinoids. *Front Plant Sci*, **9**, 1222.

Chambers, M.C., Maclean, B., Burke, R., Amodei, D., Ruderman, D.L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T.A., Brusniak, M.Y., Paulse, C., Creasy, D., Flashner, L., Kani, K., Moulding, C., Seymour, S.L., Nuwaysir, L.M., Lefebvre, B., Kuhlmann, F., Roark, J., Rainer, P., Detlev, S., Hemenway, T., Huhmer, A., Langridge, J., Connolly, B., Chadick, T., Holly, K., Eckels, J., Deutsch, E.W., Moritz, R.L., Katz, J.E., Agus, D.B., MacCoss, M., Tabb, D.L. and Mallick, P. (2012) A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotechnol*, **30**, 918-920.

DeBlasio, S.L., Rebelo, A.R., Parks, K., Gray, S.M. and Heck, M.C. (2018) Disruption of chloroplast function through downregulation of phytoene desaturase enhances the

systemic accumulation of an aphid-borne, phloem-restricted virus. *Mol Plant Microbe Interact*, **31**, 1095-1110.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15-21.

Dutartre, L., Hilliou, F. and Feyereisen, R. (2012) Phylogenomics of the benzoxazinoid biosynthetic pathway of Poaceae: gene duplications and origin of the *Bx* cluster. *BMC Evol Biol*, **12**, 64.

Erb, M., Veyrat, N., Robert, C.A., Xu, H., Frey, M., Ton, J. and Turlings, T.C. (2015) Indole is an essential herbivore-induced volatile priming signal in maize. *Nat Commun*, **6**, 6273, DOI: 6210.1038/ncomms7273.

Farber, G.K. and Petsko, G.A. (1990) The evolution of alpha/beta barrel enzymes. *Trends Biochem Sci*, **15**, 228-234.

Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grun, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R.B., Briggs, S.P., Simcox, K. and Gierl, A. (1997) Analysis of a chemical plant defense mechanism in grasses. *Science*, **277**, 696-699.

Frey, M., Kliem, R., Saedler, H. and Gierl, A. (1995) Expression of a cytochrome P450 gene family in maize. *Mol Gen Genet*, **246**, 100-109.

Frey, M., Schullehner, K., Dick, R., Fiesselmann, A. and Gierl, A. (2009) Benzoxazinoid biosynthesis, a model for evolution of secondary metabolic pathways in plants. *Phytochemistry*, **70**, 1645-1651.

Frey, M., Stettner, C., Pare, P.W., Schmelz, E.A., Tumlinson, J.H. and Gierl, A. (2000) An herbivore elicitor activates the gene for indole emission in maize. *Proc Natl Acad Sci U S A*, **97**, 14801-14806.

Glawischnig, E., Grun, S., Frey, M. and Gierl, A. (1999) Cytochrome P450 monooxygenases of DIBOA biosynthesis: specificity and conservation among grasses. *Phytochemistry*, **50**, 925-930.

Hamilton, R.H. (1964) A corn mutant deficient in 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one with an altered tolerance of atrazine. *Weeds*, **12**, 27-30.

Jones, D.T., Taylor, W.R. and Thornton, J.M. (1992) The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci*, **8**, 275-282.

Karimi, S., Ahl, D., Vagesjo, E., Holm, L., Phillipson, M., Jonsson, H. and Roos, S. (2016) In vivo and in vitro detection of luminescent and fluorescent *Lactobacillus reuteri* and application of red fluorescent mCherry for assessing plasmid persistence. *PloS one*, **11**, e0151969.

Köllner, T.G., Schnee, C., Gershenzon, J. and Degenhardt, J. (2004) The sesquiterpene hydrocarbons of maize (*Zea mays*) form five groups with distinct developmental and organ-specific distributions. *Phytochemistry*, **65**, 1895-1902.

Kong, L., Wu, J., Lu, L., Xu, Y. and Zhou, X. (2014) Interaction between Rice stripe virus disease-specific protein and host PsbP enhances virus symptoms. *Mol Plant*, **7**, 691-708.

Kremling, K.A.G., Chen, S.Y., Su, M.H., Lepak, N.K., Romay, M.C., Swarts, K.L., Lu, F., Lorant, A., Bradbury, P.J. and Buckler, E.S. (2018) Dysregulation of expression correlates with rare-allele burden and fitness loss in maize. *Nature*, **555**, 520-523.

Kriechbaumer, V., Weigang, L., Fiesselmann, A., Letzel, T., Frey, M., Gierl, A. and Glawischnig, E. (2008) Characterisation of the tryptophan synthase alpha subunit in maize. *BMC Plant Biol*, **8**, 44.

Kuhl, C., Tautenhahn, R., Bottcher, C., Larson, T.R. and Neumann, S. (2012) CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Anal Chem*, **84**, 283-289.

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*, **35**, 1547-1549.

Li, J., Chen, S., Zhu, L. and Last, R.L. (1995a) Isolation of cDNAs encoding the tryptophan pathway enzyme indole-3-glycerol phosphate synthase from Arabidopsis thaliana. *Plant Physiol*, **108**, 877-878.

Li, J., Zhao, J., Rose, A.B., Schmidt, R. and Last, R.L. (1995b) Arabidopsis phosphoribosylanthranilate isomerase: molecular genetic analysis of triplicate tryptophan pathway genes. *Plant Cell*, **7**, 447-461.

Maag, D., Köhler, A., Robert, C.A., Frey, M., Wolfender, J.L., Turlings, T.C., Glauser, G. and Erb, M. (2016) Highly localized and persistent induction of *Bx1*-dependent herbivore resistance factors in maize. *Plant J*, **88**, 976-991.

Marcon, C., Altrogge, L., Win, Y.N., Stocker, T., Gardiner, J.M., Portwood, J.L., Opitz, N., Kortz, A., Baldauf, J., Hunter, C.T., McCarty, D.R., Koch, K.E., Schoof, H. and Hochholdinger, F. (2020) BonnMu: a sequence-indexed resource of transposon-induced maize mutations for functional genomics studies. *Plant Physiol*, **epub ahead of print, DOI:10.1104/pp.20.00478**.

McCarty, D.R., Settles, A.M., Suzuki, M., Tan, B.C., Latshaw, S., Porch, T., Robin, K., Baier, J., Avigne, W., Lai, J., Messing, J., Koch, K.E. and Hannah, L.C. (2005) Steady-state transposon mutagenesis in inbred maize. *Plant J*, **44**, 52-61.

McMullen, M., Frey, M. and Degenhardt, J. (2009a) Genetics and biochemistry of insect resistance in maize. In *Handbook of Maize: its Biology* (Bennetzen, J.L. and Hake, S. eds). New York: Springer, pp. 587.

McMullen, M.D., Kresovich, S., Villeda, H.S., Bradbury, P., Li, H., Sun, Q., Flint-Garcia, S., Thornsberry, J., Acharya, C., Bottoms, C., Brown, P., Browne, C., Eller, M., Guill, K., Harjes, C., Kroon, D., Lepak, N., Mitchell, S.E., Peterson, B., Pressoir, G., Romero, S., Oropeza Rosas, M., Salvo, S., Yates, H., Hanson, M., Jones, E., Smith, S., Glaubitz, J.C., Goodman, M., Ware, D., Holland, J.B. and Buckler, E.S. (2009b) Genetic properties of the maize nested association mapping population. *Science*, **325**, 737-740.

Meihls, L.N., Handrick, V., Glauser, G., Barbier, H., Kaur, H., Haribal, M.M., Lipka, A.E., Gershenzon, J., Buckler, E.S., Erb, M., Köllner, T.G. and Jander, G. (2013) Natural variation in maize aphid resistance is associated with 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one glucoside methyltransferase activity. *Plant Cell*, **25**, 2341-2355.

Meihls, L.N., Kaur, H. and Jander, G. (2012) Natural variation in maize defense against insect herbivores. *Cold Spring Harbor Symp Quant Biol*, **77**, 269-283.

Melanson, D., Chilton, M.D., Masters-Moore, D. and Chilton, W.S. (1997) A deletion in an indole synthase gene is responsible for the DIMBOA-deficient phenotype of bxbx maize. *Proc Natl Acad Sci U S A*, **94**, 13345-13350.

Nelson, B.K., Cai, X. and Nebenfuhr, A. (2007) A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. *Plant J*, **51**, 1126-1136.

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

Richter, A., Schaff, C., Zhang, Z., Lipka, A.E., Tian, F., Köllner, T.G., Schnee, C., Preiss, S., Irmisch, S., Jander, G., Boland, W., Gershenzon, J., Buckler, E.S. and Degenhardt, J. (2016) Characterization of the (herbivore-induced) biosynthetic pathways of the volatile homoterpenes 3,8-dimethyl-1,4,7-nonatriene (DMNT) and 4,8,12-trimethyltrideca-1,3,7,11-tetraene, in *Zea mays*. *Plant Cell*, **28**, 2651-2665.

Richter, A., Seidl-Adams, I., Kollner, T.G., Schaff, C., Tumlinson, J.H. and Degenhardt, J. (2015) A small, differentially regulated family of farnesyl diphosphate synthases in maize (Zea mays) provides farnesyl diphosphate for the biosynthesis of herbivore-induced sesquiterpenes. *Planta*, **241**, 1351-1361.

Rose, A.B. and Last, R.L. (1994) Molecular Genetics of Amino Acid, Nucleotide, and Vitamin Biosynthesis. *Cold Spring Harbor Monograph Archive*, **27**, 835-879.

Settles, A.M., Holding, D.R., Tan, B.C., Latshaw, S.P., Liu, J., Suzuki, M., Li, L., O'Brien, B.A., Fajardo, D.S., Wroclawska, E., Tseung, C.W., Lai, J., Hunter, C.T., 3rd, Avigne, W.T., Baier, J., Messing, J., Hannah, L.C., Koch, K.E., Becraft, P.W., Larkins, B.A. and McCarty, D.R. (2007) Sequence-indexed mutations in maize using the UniformMu transposon-tagging population. *BMC Genomics*, **8**, 116.

Singh, M. and Widholm, J.M. (1974) Measurement of 5 enzymes which convert chorismate to tryptophan in wheat plants (*Triticum aestivum* cv Kalyansona). *Physiologia Plantarum*, **32**, 240-246.

Springer, N.M., Anderson, S.N., Andorf, C.M., Ahern, K., Bai, F., Barad, O., Brad Barbazuk, W.B., Bass, H.W., Baruch, K., Ben-Zvi, G., Buckler, E.S., Bukowski, R., Campbell, M.S., Cannon, E.K.S., Chomet, P., Dawe, R.K., Davenport, R., Dooner, H.K., Du, L.H., Du, C., Easterling, K.A., Gault, C., Guan, J.-C., Jander, G., Hunter, C.T., Jiao, Y., Koch, K.E., Kol, G., Kudo, T., Li, Q., Lu, F., Mayfield-Jones, D., Mei, W., McCarty, D.R., Noshay, J., Portwood II, J.L., Ronen, G., Settles, M.A., Shem-Tov, D., Shi, J., Soifer, I., Stein, J.C., Suzuki, M., Vera, D.L., Vollbrecht, E., Vrebalov, J.T., Ware, D., Wei, X., Wimalanathan, K., Woodhouse, M.R., Xiong, W. and Brutnell, T.P. (2018) The W22 genome: a foundation for maize functional genomics and transposon biology. *Nat Genet*, **50**, 1282-1288.

Stettner, C. (1998) DIMBOA-Biosynthese in Mais: Isolierung und funktionelle Charakterisierung des *Bx1*-Gens. In *Lehrstuhl für Genetik*. Munich: Technische Universität München, pp. 95.

Tamiru, A., Bruce, T.J.A., Richter, A., Woodcock, C.M., Midega, C.A.O., Degenhardt, J., Kelemu, S., Pickett, J.A. and Khan, Z.R. (2017) A maize landrace that emits defense volatiles in response to herbivore eggs possesses a strongly inducible terpene synthase gene. *Ecol Evol*, **7**, 2835-2845.

Tautenhahn, R., Bottcher, C. and Neumann, S. (2008) Highly sensitive feature detection for high resolution LC/MS. *Bmc Bioinformatics*, **9**, 504.

Tautenhahn, R., Patti, G.J., Rinehart, D. and Siuzdak, G. (2012) XCMS Online: a web-based platform to process untargeted metabolomic data. *Anal Chem*, **84**, 5035-5039.

Tzin, V., Fernandez-Pozo, N., Richter, A., Schmelz, E.A., Schoettner, M., Schaefer, M., Ahern, K.R., Meihls, L.N., Kaur, H., Huffaker, A., Mori, N., Degenhardt, J., Mueller, L.A. and Jander, G. (2015) Dynamic maize responses to aphid feeding are

revealed by a time series of transcriptomic and metabolomic assays *Plant Physiol*, **169**, 1727-1743.

Tzin, V., Hojo, Y., Strickler, S.R., Bartsch, L.J., Archer, C.M., Ahern, K.R., Zhou, S.Q., Christensen, S.A., Galis, I., Mueller, L.A. and Jander, G. (2017) Rapid defense responses in maize leaves induced by *Spodoptera exigua* caterpillar feeding. *J Exp Bot*, **68**, 4709-4723.

Vogel, H.J. and Bonner, D.M. (1956) Acetylornithinase of *Escherichia coli*: partial purification and some properties. *J Biol Chem*, **218**, 97-106.

Wierenga, R.K. (2001) The TIM-barrel fold: a versatile framework for efficient enzymes. *FEBS Lett*, **492**, 193-198.

Williams-Carrier, R., Stiffler, N., Belcher, S., Kroeger, T., Stern, D.B., Monde, R.A., Coalter, R. and Barkan, A. (2010) Use of Illumina sequencing to identify transposon insertions underlying mutant phenotypes in high-copy Mutator lines of maize. *Plant J*, **63**, 167-177.

Wilmanns, M., Hyde, C.C., Davies, D.R., Kirschner, K. and Jansonius, J.N. (1991) Structural conservation in parallel beta/alpha-barrel enzymes that catalyze three sequential reactions in the pathway of tryptophan biosynthesis. *Biochemistry*, **30**, 9161-9169.

Wisecaver, J.H., Borowsky, A.T., Tzin, V., Jander, G., Kliebenstein, D.J. and Rokas, A. (2017) A global coexpression network approach for connecting genes to specialized metabolic pathways in plants. *Plant Cell*, **29**, 944-959.

Wouters, F.C., Blanchette, B., Gershenzon, J. and Vassao, D.G. (2016) Plant defense and herbivore counter-defense: benzoxazinoids and insect herbivores. *Phytochem Rev*, **15**, 1127-1151.

Wright, A.D., Moehlenkamp, C.A., Perrot, G.H., Neuffer, M.G. and Cone, K.C. (1992) The maize auxotrophic mutant orange pericarp is defective in duplicate genes for tryptophan synthase beta. *Plant Cell*, **4**, 711-719.

Yang, X., Baliji, S., Buchmann, R.C., Wang, H., Lindbo, J.A., Sunter, G. and Bisaro, D.M. (2007) Functional modulation of the geminivirus AL2 transcription factor and silencing suppressor by self-interaction. *J Virol*, **81**, 11972-11981.

Yanofsky, C., Horn, V., Bonner, M. and Stasiowski, S. (1971) Polarity and enzyme functions in mutants of the first three genes of the tryptophan operon of Escherichia coli. *Genetics*, **69**, 409-433.

Zhang, R., Wang, B., Ouyang, J., Li, J. and Wang, Y. (2008) Arabidopsis indole synthase, a homolog of tryptophan synthase alpha, is an enzyme involved in the Trp-independent indole-containing metabolite biosynthesis. *J Integr Plant Biol*, **50**, 1070-1077.

Figures Legends

Figure 1. Biosynthetic pathways from indole-3-glycerolphosphate to DIBOA-Glc, tryptophan, and free indole in maize. Protein localization and confirmation of IGPS1, IPGS2, and IGPS3 enzymatic activity is described in this manuscript. IGPS = indole-3-glycerolphosphate synthase, BX = benzoxazinoneless, TSA = tryptophan synthase alpha subunit, TSB = tryptophan synthase beta subunit, IGL = indole-3-glycerolphosphate lyase, DIBOA-Glc = 2,4-dihydroxy-1,4-benzoxazin-3-one glucoside, ER = endoplasmic reticulum.

Figure 2. Phylogenetic analysis of indole-3-glycerol synthase (IGPS) proteins from monocots and dicots, including *Zea mays* (maize), *Triticum aestivum* (wheat) *Sorghum bicolor* (sorghum), *Oryza sativa* (rice), *Hordeum vulgare* (barley), *Setaria viridis* (green bristlegrass), *Arabidopsis thaliana* (mouse-ear cress), *Populus trichocarpa* (poplar), and *Nicotiana attenuata* (coyote tobacco). The enzymes cluster into phylogenetic groups that include the three maize enzymes, IGPS1 (orange cluster); IGPS2 (grey cluster); and IGPS3 (blue cluster). The MUSCLE Codon algorithm was used to perform a sequence alignment, which was used to reconstruct the tree using the Maximum Likelihood method and JTT matrix-based model. The percentage of trees in which the associated taxa clustered together is shown next to the branches (1,000 bootstrap trials). Maize tryptophan synthase alpha subunit (TSA) was used as an outgroup for the analysis. All protein sequences were retrieved from GenBank (https://www.ncbi.nlm.nih.gov/genbank/) using the indicated identifiers.

Figure 3. Identification of maize indole-3-glycerol-phosphate synthase (IGPS) activity. Maize IGPS enzyme activity was verified by cloning the cDNA and complementing an *E. coli trpC9800⁻* mutation, which causes tryptophan auxotrophy through loss of indole-3-glycerolphosphate synthase activity. Empty vector controls were grown with or without the addition of 20 µg/ml tryptophan to the minimal medium. Optical density at 600 nm (OD600) was measured over 11 h.

Figure 4. Transcript abundance of *IGPS1*, *IGPS2*, and *IGPS3* in maize inbred line W22 leaves. Transcript levels of maize *IGPS* genes were determined after (A) 24 h after initiation of *Spodoptera exigua* feeding (mean +/- s.e. of n = 5), (B) 24 h after methyl jasmonate induction (mean +/- s.e. of N = 3), and (C) *Rhopalosiphum maidis* infestation (mean +/- s.e. of N = 8). Values are fold change relative to uninfested controls (n = 5). *P < 0.05, ***P < 0.001 by Student´s *t*-test (panels A and B) or Dunnett's test relative to 0 h control (panel C).

Figure 5. Effects of *igps1* and *igps3* knock-out mutations on gene expression.
(A) *IGPS1* expression in a segregating population after induction with methyl jasmonate (mean +/- s.e. of N = 6). (B) Expression of *IGPS* genes in wildtype W22 and a homozygous *igps1* mutant line after methyl jasmonate induction (mean +/- s.e of N = 6). (C) *IGPS* gene expression in wildtype W22 and a homozygous *igps3* mutant line after induction with methyl jasmonate (mean +/- s.e. of N = 4). Different letters indicate P < 0.05, ANOVA followed by Tukey's HSD test. *P < 0.05, ***P < 0.001 by Student´s *t*-test.

Figure 6. Effects of *igps1* and *igps3* knock-out mutations on metabolites and caterpillar growth. (A) Comparison of the five most abundant benzoxazinoids in an *igps1* segregating population (mean +/- s.e of N = 4-8). (B) The five most abundant benzoxazinoids in *igps3* mutant and wildtype plants (mean +/- s.e. of N = 4). (C) Tryptophan abundance in an *igps1* segregating population (mean +/- s.e. of N = 4-8). (D) Tryptophan in *igps3* mutant and wildtype plants (mean +/- s.e. of N = 4). (E) Free indole in wildtype, *igps1*, and *igps3* plants (mean +/- s.e. of N = 4). (F) Growth of *Spodoptera exigua* larvae on homozygous *igps1* mutant and wildtype W22. *P < 0.05, **P < 0.01, ***P < 0.001 by Student´s *t*-test. Different letters on bars indicate P < 0.05, ANOVA and Fischer´s LSD test. No significant differences were detected for panels D and E (P > 0.05). Samples in panels A-E were treated methyl jasmonate to induce production of defensive metabolites.

Figure 7. Subcellular localization of IGP synthases, BX1, BX2, TSA, and IGL.
(A) Images of maize mesophyll protoplasts transiently expressing *35S-IGPS1-YFP*, *35S-IGPS2-YFP*, *35S-IGPS3-YFP, 35S-YFP-BX1*, *35S-YFP-IGL*, and *35S-YFP-TSA*, showing YFP

fluorescence in chloroplasts, chlorophyll auto-fluorescence, and merged images. **(B)** Co-localization of *35S-BX2-mcherry* with *ER-CFP*. The white scale bars indicate 25 µm.

Figure 8. Interactions of IGPS proteins with BX1, TSA, and IGL by BiFC assays in *N. benthamiana*. The overlap of YFP (yellow) and the chloroplast auto-fluorescence (red) images is shown in the column labeled merged. IGPSs were fused with C-terminus of YFP (c-YFP) and each construct was co-transformed into *N. benthamiana* leaves, with BX1, TSA, and IGL, which were fused with the N-terminus of YFP (n-YFP). (A) Interactions of IGPS1 with BX1, TSA, and IGL, (B) interactions of IGPS2 with BX1, TSA, and IGL, (C) interactions of IGPS2 with BX1, TSA, and IGL, and (D) interactions of IGPS1 N-terminal YFP fusions with IGPS2 and IGPS3 C-terminal YFP fusions. The white scale bars indicate 25 µm.

Supplemental Figure S1. Alignment of the three maize IGPS proteins. Relevant protein features are highlighted.

Supplemental Figure S2. Expression of IGPS genes in response to insect feeding on maize inbred line B73. (A) Feeding by *Spodoptera* exigua (beet armyworm), mean +/- s.e. of N=4 (B) Feeding by *Rhopalosiphum maidis,* mean +/- s.e. of N = 5. *P < 0.05, **P < 0.01, ***P < 0.001, Dunnett's test relative to 0-hr control. These figures are produced from RNAseq data presented in Tzin et al (2017) and Tzin et al (2015), respectively.

Supplemental Figure S3. Standard curves for quantifying tryptophan and indole. (A) Tryptophan, as measured by LC-MS, (B) Indole, as measured by GC-MS.

Figure 1. Biosynthetic pathways from indole-3-glycerolphosphate to DIBOA-Glc, tryptophan, and free indole in maize. Protein localization and confirmation of IGPS1, IPGS2, and IGPS3 enzymatic activity is described in this manuscript. IGPS = indole-3-glycerolphosphate synthase, BX = benzoxazinoneless, TSA = tryptophan synthase alpha subunit, TSB = tryptophan synthase beta subunit, IGL = indole-3-glycerolphosphate lyase, DIBOA-Glc = 2,4-dihydroxy-1,4-benzoxazin-3-one glucoside, ER = endoplasmic reticulum.
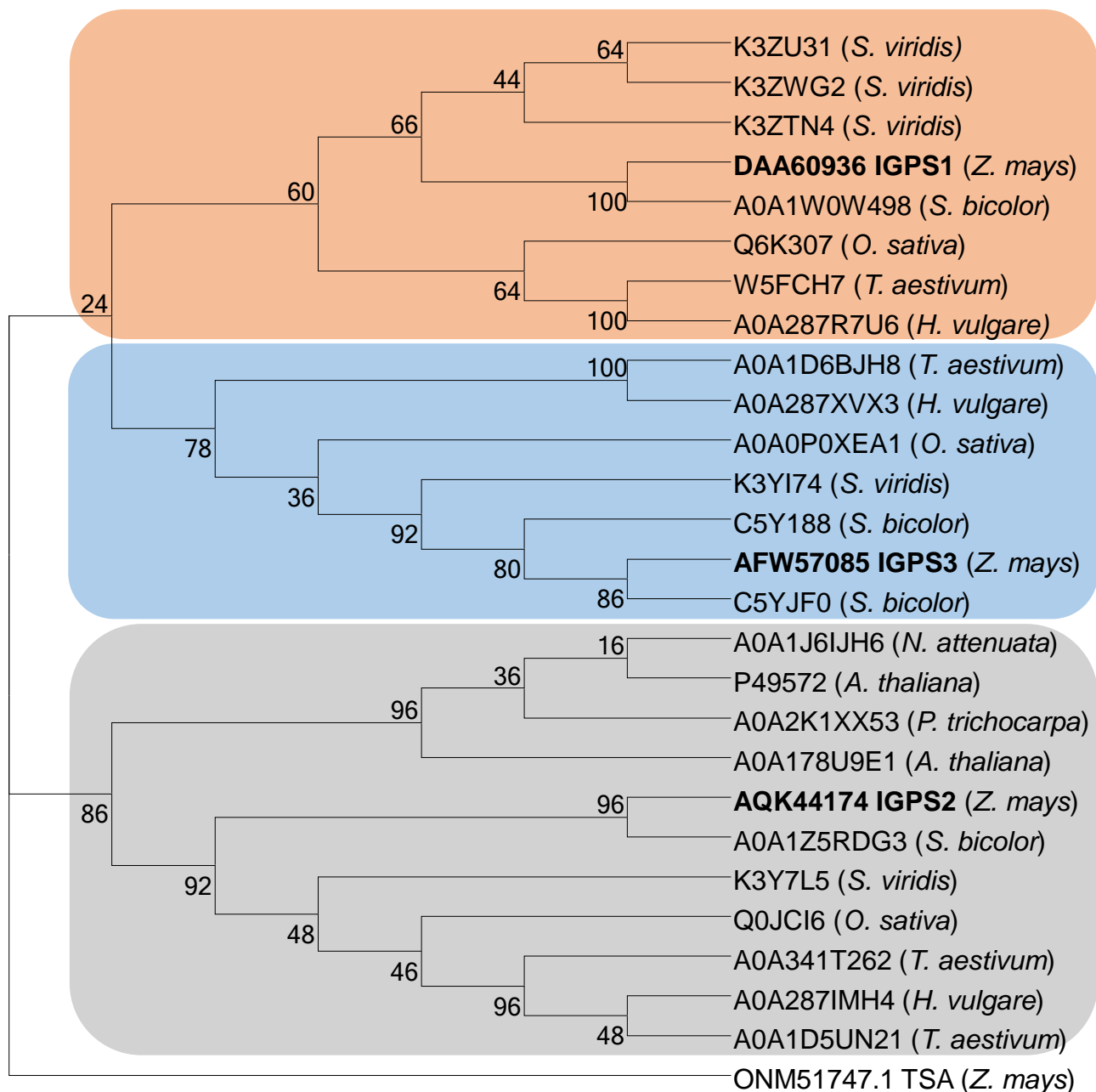
Figure 2. Phylogenetic analysis of indole-3-glycerol synthase (IGPS) proteins from monocots and dicots, including *Zea mays* (maize), *Triticum aestivum* (wheat) *Sorghum bicolor* (sorghum), *Oryza sativa* (rice), *Hordeum vulgare* (barley), *Setaria viridis* (green bristlegrass), *Arabidopsis thaliana* (mouse-ear cress), *Populus trichocarpa* (poplar), and *Nicotiana attenuata* (coyote tobacco). The enzymes cluster into phylogenetic groups that include the three maize enzymes, IGPS1 (orange cluster); IGPS2 (grey cluster); and IGPS3 (blue cluster). The MUSCLE Codon algorithm was used to perform a sequence alignment, which was used to reconstruct the tree using the Maximum Likelihood method and JTT matrix-based model. The percentage of trees in which the associated taxa clustered together is shown next to the branches (bootstrap trials 1,000). Maize tryptophan synthase alpha subunit (TSA) was used as an outgroup for the analysis. All protein sequences were retrieved from GenBank (https://www.ncbi.nlm.nih.gov/genbank/) using the indicated identifiers.

Figure 3. Identification of maize indole-3-glycerol-phosphate synthase (IGPS) activity. Maize IGPS enzyme activity was verified by cloning the cDNA and complementing an *E. coli trpC9800*[-] mutation, which causes tryptophan auxotrophy through loss of indole-3-glycerolphosphate synthase activity. Empty vector controls were grown with or without the addition of 20 µg/mL tryptophan to the minimal medium. Optical density at 600 nm (OD600) was measured over 11 h.
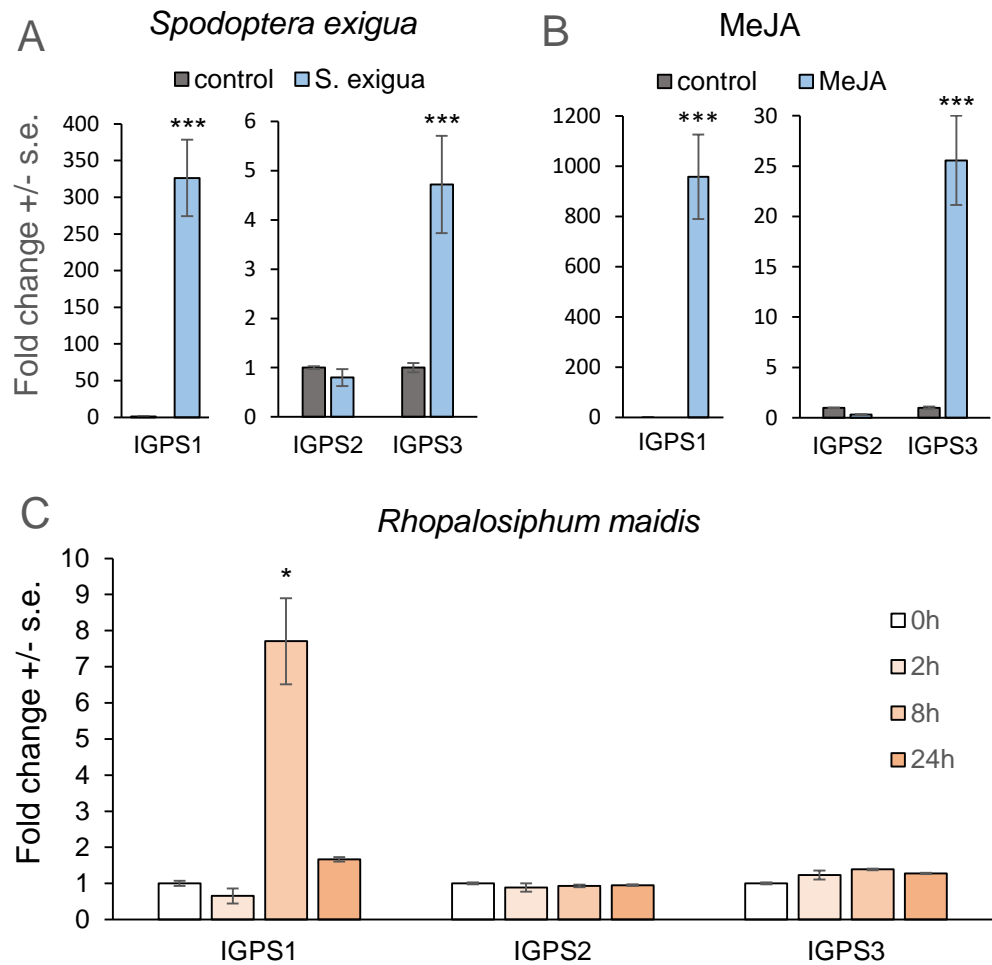
Figure 4. Transcript abundance of *IGPS1*, *IGPS2,* and *IGPS3* in maize inbred line W22 leaves. Transcript levels of maize *IGPS* genes were determined after (A) 24 h after initiation of *Spodoptera exigua* feeding (mean +/- s.e. of n = 5), (B) 24 h after methyl jasmonate (MeJA) induction (mean +/- s.e. of N = 3), and (C) *Rhopalosiphum maidis* infestation (mean +/- s.e. of N = 8). Values are fold change relative to uninfested controls (n = 5). *P < 0.05, ***P < 0.001 by Student´s *t*-test (panels A and B) or Dunnett's test relative to 0 h control (panel C).

Figure 5. Effects of *igps1* and *igps3* knock-out mutations on gene expression.
(A) *IGPS1* expression in a segregating population after induction with methyl jasmonate
(mean +/- s.e. of N = 6). (B) Expression of *IGPS* genes in wildtype W22 and a homozygous
igps1 mutant line after methyl jasmonate induction (mean +/- s.e of N = 6). (C) *IGPS* gene
expression in wildtype W22 and a homozygous *igps3* mutant line after induction with methyl
jasmonate (mean +/- s.e. of N = 4). Different letters indicate P < 0.05, ANOVA followed by
Tukey's HSD test. *P < 0.05, ***P < 0.001 by Student's *t*-test.

Figure 6. Effects of *igps1* and *igps3* knock-out mutations on metabolites and caterpillar growth. (A) Comparison of the five most abundant benzoxazinoids in an *igps1* segregating population (mean +/- s.e of N = 4-8). (B) The five most abundant benzoxazinoids in *igps3* mutant and wildtype plants (mean +/- s.e. of N = 4). (C) Tryptophan abundance in an *igps1* segregating population (mean +/- s.e. of N = 4-8). (D) Tryptophan in *igps3* mutant and wildtype plants (mean +/- s.e. of N = 4). (E) Free indole in wildtype, *igps1*, and *igps3* plants (mean +/- s.e. of N = 4). (F) Growth of *Spodoptera exigua* larvae on homozygous *igps1* mutant and wildtype W22. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ by Student´s *t*-test. Different letters on bars indicate $P < 0.05$, ANOVA and Fischer´s LSD test. No significant differences were detected for panels D and E ($P > 0.05$). Samples in panels A-E were treated methyl jasmonate to induce production of defensive metabolites.
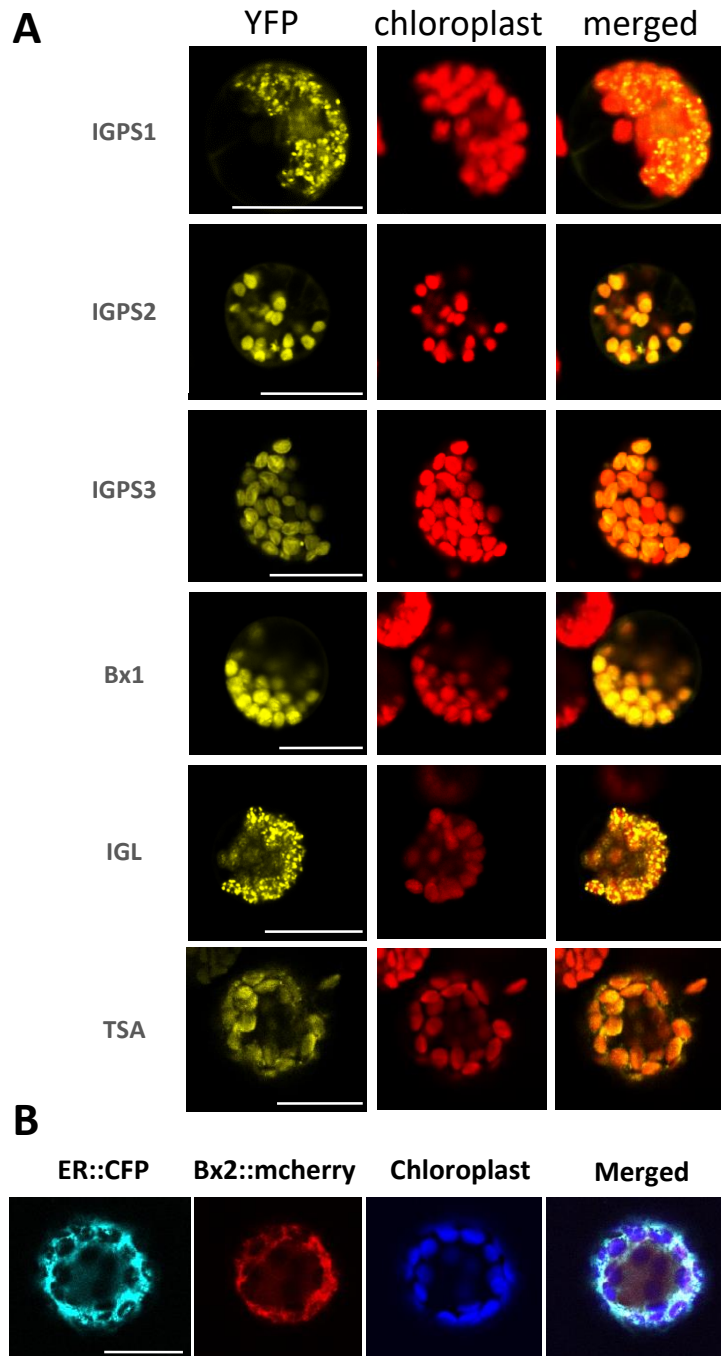
Figure 7. Subcellular localization of IGP synthases, BX1, BX2, TSA, and IGL.
(**A**) Images of maize mesophyll protoplasts transiently expressing *35S-IGPS1-YFP*, *35S-IGPS2-YFP*, *35S-IGPS3-YFP, 35S-YFP-BX1*, *35S-YFP-IGL*, and *35S-YFP-TSA*, showing YFP fluorescence in chloroplasts, chlorophyll auto-fluorescence, and merged images. (**B**) Co-localization of *35S-BX2-mcherry* with *ER-CFP*. The white scale bars indicate 25 µm.
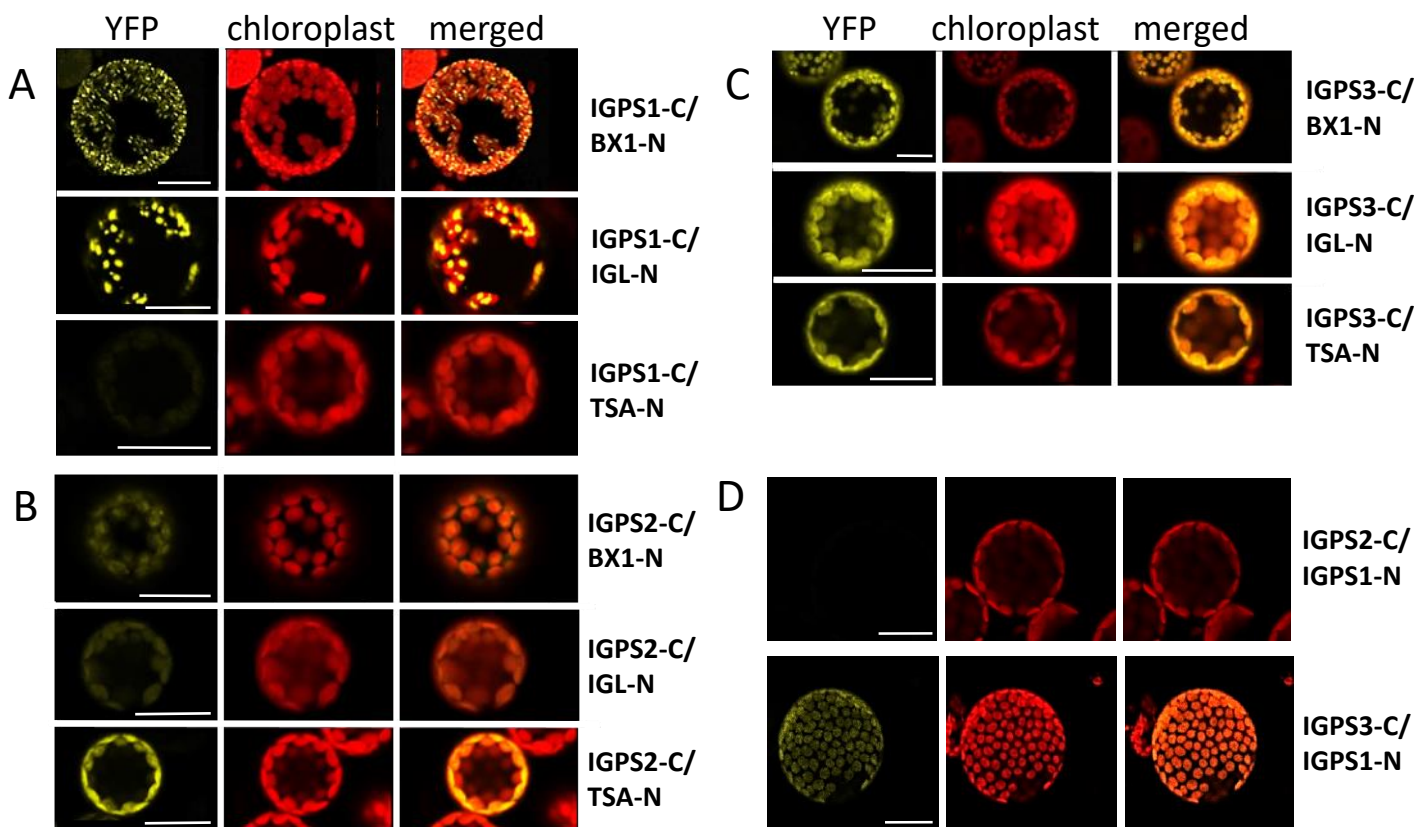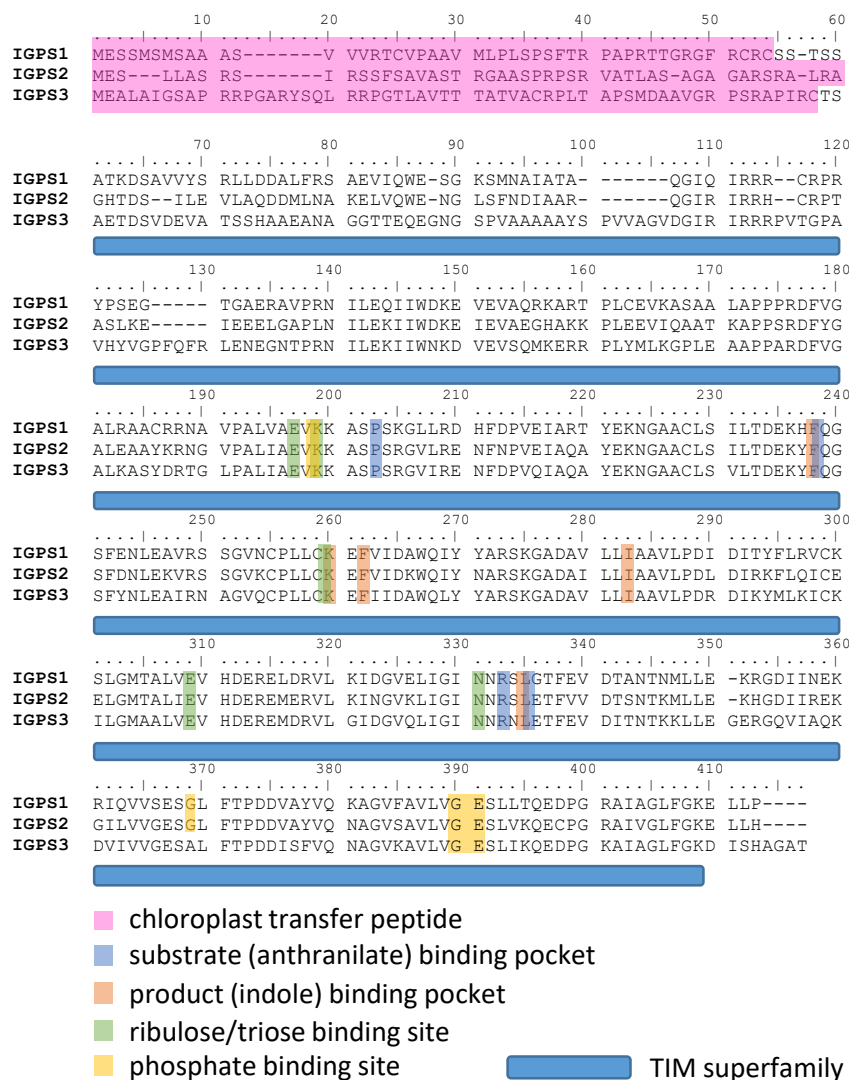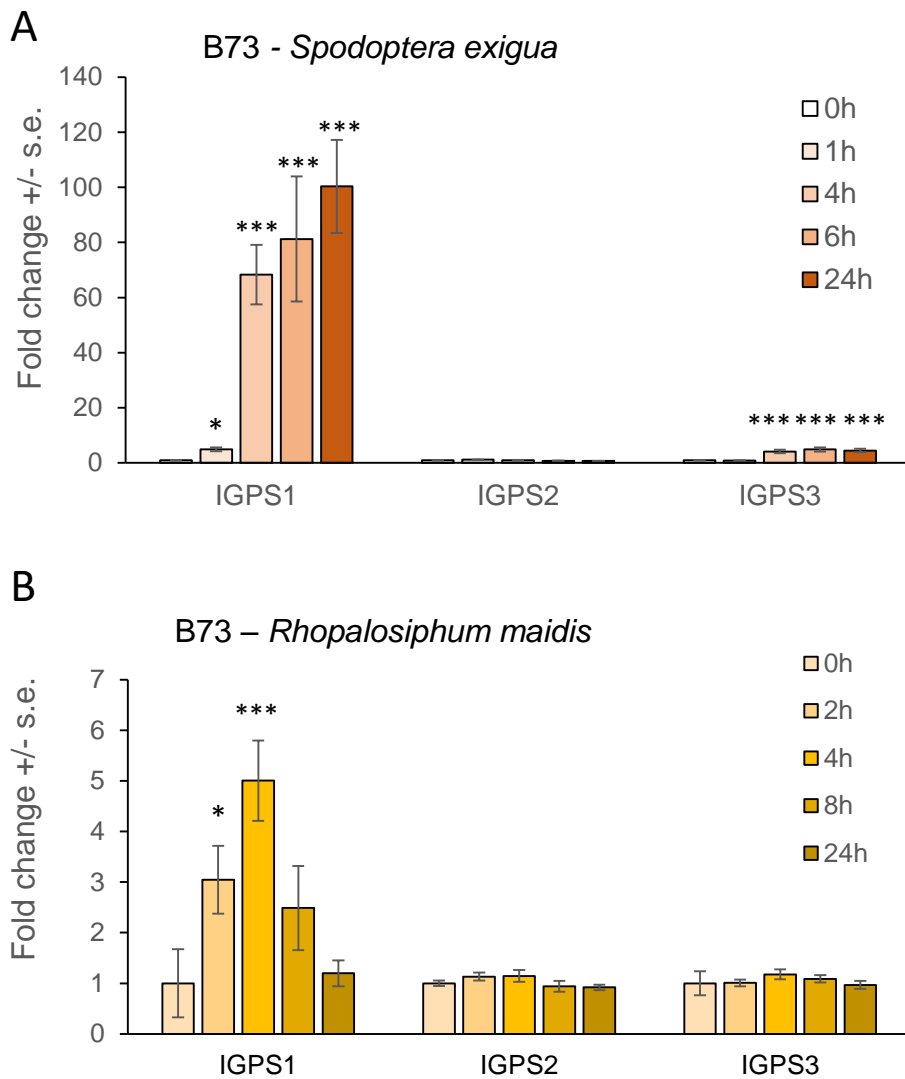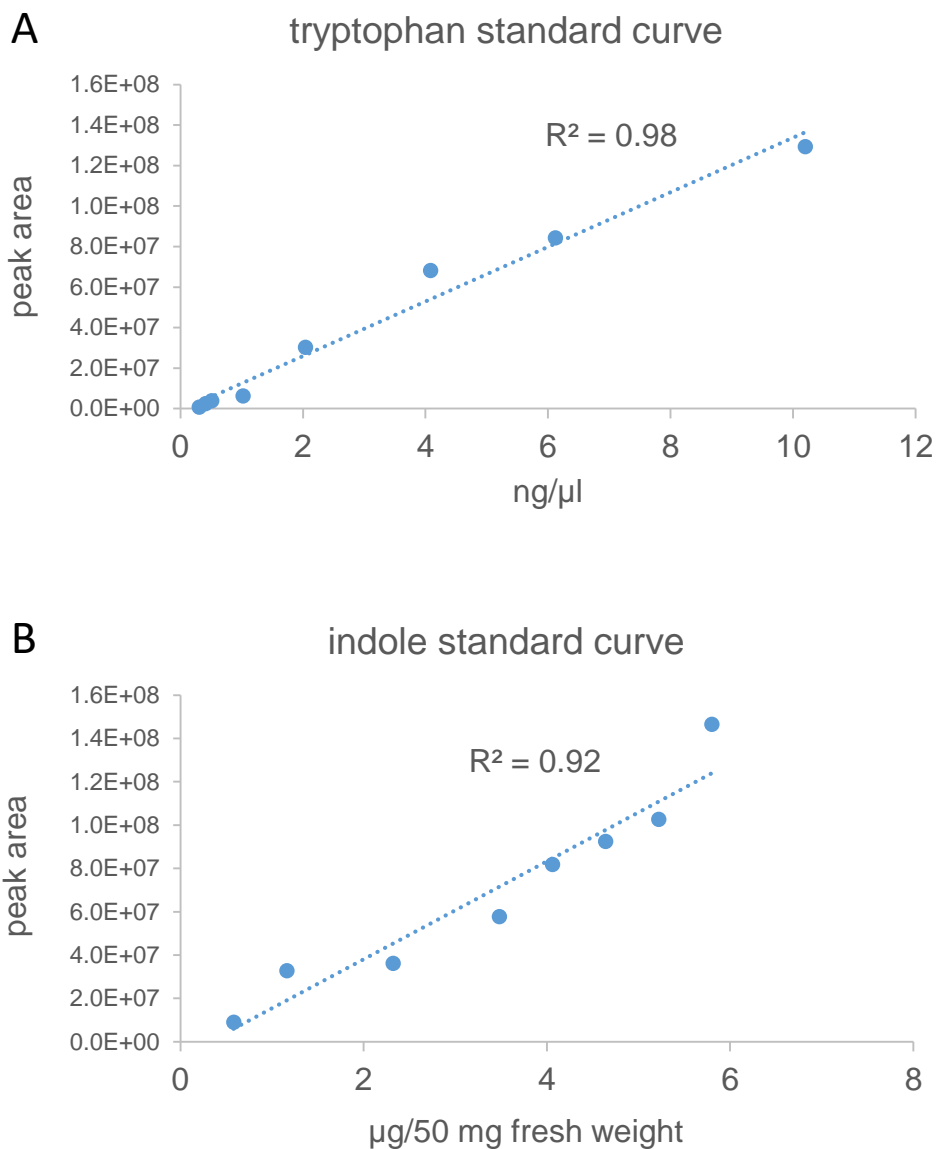
Figure 8. Interactions of IGPS proteins with BX1, TSA, and IGL by BiFC assays in *N. benthamiana*. The overlap of YFP (yellow) and the chloroplast auto-fluorescence (red) images is shown in the column labeled merged. IGPSs were fused with C-terminus of YFP (c-YFP) and each construct was co-transformed into *N. benthamiana* leaves, with BX1, TSA, and IGL, which were fused with the N-terminus of YFP (n-YFP). (A) Interactions of IGPS1 with BX1, TSA, and IGL, (B) interactions of IGPS2 with BX1, TSA, and IGL, (C) interactions of IGPS2 with BX1, TSA, and IGL, and (D) interactions of IGPS1 N-terminal YFP fusions with IGPS2 and IGPS3 C-terminal YFP fusions. The white scale bars indicate 25 µm.

Supplemental Figure S1. Alignment of the three maize IGPS proteins. Relevant protein features are highlighted.

Supplemental Figure S2. Expression of IGPS genes in response to insect feeding on maize inbred line B73. (A) Feeding by *Spodoptera* exigua (beet armyworm), mean +/- s.e. of N=4 (B) Feeding by *Rhopalosiphum maidis,* mean +/- s.e. of N = 5. *P < 0.05, **P < 0.01, ***P < 0.001, Dunnett's test relative to 0-hr control. These figures are produced from RNAseq data presented in Tzin et al (2017) and Tzin et al (2015), respectively.

Supplemental Figure S3. Standard curves for indole and tryptophan. (A) Tryptophan, as measured by LC-MS. (B) Indole, as measured by GC-MS.