SUPPLEMENTARY INFORMATION

The fall armyworm *Spodoptera frugiperda* utilizes specific UDP-glycosyltransferases to inactivate maize defensive benzoxazinoids

Bhawana Israni, Felipe C. Wouters, Katrin Luck, Elena Seibel, Seung-Joon Ahn, Christian Paetz, Maximilian Reinert, Heiko Vogel, Matthias Erb, David G. Heckel, Jonathan Gershenzon, Daniel Giddings Vassão



Supplementary figure 1. Recovery of BXD metabolites in *S. frugiperda* **frass.** Overall rates of recovery calculated after oral administration of DIMBOA, MBOA and (2*R*)-DIMBOA-Glc (n=5-6, mean \pm SEM).

	signal pentide cleavage site	
SfUGT33F28 SfUGT33B25 SfUGT42C10 SfUGT40R13 SfUGT40R13 SfUGT40L8 SfUGT40L8 SfUGT40L9 SfUGT40F19 SfUGT33V6 HaUGT47A2 HsUGT2B7	1 KILL C. TILL - S. FANYDE - AARI LAVEPNESI SHOVYERPI TOELARRGHEVTVI TTDPVEPEG-KAPANL TELDVH - DI SY - 1 STALS - L-FIL- S. SI MYSE - AARI LAVEPNESI SHOVYERPI TOELARRGHEVTVI TTDPVEPEG-KAPANL TELDVH - DESY - 1 KKENNI - II - LEF - SALESQU - S. NI LGVEPYGEKSHEFVEAPY KELARRGHOVTVI TPDPAFPKG-GTEANL TELDVH - DESY - 1 MKENNI - II - LEF - SALESQU - S. NI LGVEPYGEKSHEFVEAPY KELARRGHOVTVI TPDPAFPKG-GTEANL TELDVH - DESY - 1 MKENNI - II - LEF - SALESQU - S. NI LGVEPYGEKSHEFVEAPY KELARRGHOVTVI STE DIK - DIFPONIEGO NAM - SKED - 1 MKENNI - II - LEF - SALESQU - SANIEG FEHTGKSHO WE DPLLKTLAERGHOVTVI SFE PILK - NPENNIEGO NAM - SKED - 1 MRLA - VL - V - A FAANVQ - SARIEG FEHTGKSHO WE DPLLKTLAERGHOVTVI SFE PILK - NPENNIEGO NAM - SKED - 1 MRLA - VL - V - A FAANVQ - SARIEG FEHTGKSHO WE DPLLKTLAERGHOVTVI SFE PILK - NPENNIEGO NAM - SKED - 1 MRLA - VL - V - LAVAASA - AARI LAVEPNISSHSI LGYG INNULEAGHEVYHI TSFE PIK - NPENNIEGO A GTI P - 1 MEKL C. FLECA - VISLCSCU - A KKI LAVEPNISSHG LGDN II KHILLAGHEVTMI TPV ASKOKSIPKLHI DVT - DHHS - 1 MSLVKL - TVVASLA SI DTSG - GKI LGEPEPSI SHOVYERPLI GLARRGHEVTVI TPDPAFRG-GTEANL TELDVH - DASY - 1 NY T	
STUGT33F28 STUGT42C10 STUGT42C10 STUGT40R13 STUGT40R13 STUGT40L8 STUGT40L8 STUGT40F19 STUGT40F19 STUGT33V6 HaUGT47A2 HsUGT2B7	81 D.WRKI, F.LASTKGN-KDDVVQATAL, TA, H.V.DVQLKDE-KVLNL, RD-KSK, FDLLLVEACV 80 E.WQSFYKVTST-G-EKDLNQQTTAFTLLVDIVEMQLKLD-AVQCLLKD-AVQCLLVCEE-FDLLLVEACV 77 I.LEDVPLY	
SfUGT33F28 SfUGT42C10 SfUGT40R13 SfUGT40R13 SfUGT40R15 SfUGT40L8 SfUGT40L8 SfUGT40F19 SfUGT33V6 HaUGT47A2 HsUGT2B7	143 RP- ALVESHI Y KVPVLQ SSEMATEDN- YAN GAPI HELYEA. TRORI ON TMLEKI TEEFINDEMLN	
SfUGT33F28 SfUGT33B25 SfUGT42C10 SfUGT40R13 SfUGT46A15 SfUGT46A15 SfUGT47A12 SfUGT40F19 SfUGT33V6 HaUGT47A2 HsUGT2B7	← N-terminal domain Interformain linker C-terminal domain → 220 EXTRSHE-KDLEPI SELINNIVDULFINNPAFEEI REVPERSVIYLGELHQVP- HKPLPADLQTYDSS-KNGVI YVSFGTN/DPTKLEAD 2216 EVAKLEGPD/FTI/SELKNIVDULFINNAPHFEEI REVPERSVIYLGELHQVP- HKPLPADLKQVLDSS-KNGVI YVSFGTN/DPTKLEAD 2216 EVAKLEGPD/FTI/SELKNIVDULFINNAVSSESSESSESSESSESSESSESSESSESSESSESSESSE	
SfUGT33F28 SfUGT33B25 SfUGT42C10 SfUGT40R13 SfUGT46A15 SfUGT46A15 SfUGT46A15 SfUGT46F19 SfUGT33V6 HaUGT47A2 HsUGT2B7	01 Signature motif 01 Signature motif 03 R EV. VKT SQLPYD L WKVNGOVL PGOTPNI KLAKW POPDLLI HPK KLFI TO QL QST DEALTAGVP LVAL PINF GOSF MAERY 04 RVOMAKVFSQLPYD LVKVNKDEL PGAFKNI YI SKVLPOSDLLRIPK KLFI TO QL QST DEALTAGVP LVAL PINF GOOPGVAAA 03 KV QAL DANSELPCR LWKVEKNL PGAFKNI YI SNVLPONELLAHPK VLAF THGQLQTTEAL HYGVPTI ALPVF GOOPGVAAA 14 I KHD LK FEELKOTV WKFEES-L PN PKNVH INKWAPOPS LAHPK VLAF THGQLQTTEAL HYGVPTI ALPVF GOOPGVAAA 16 GEELINAALSKLKORV WKYESDEE GT TGN LKVKVLPOYELLAHPK VLAF THGQLQTTEAL HYGVPTI ALPVF GOOF INVKKA 316 KEELINAALSKLKORV WKYESDEE DN PRAVHI VKWAPOPS LAHPS KLFFI THGQLQTSEAL HYGI PL VGI PY ADOF INVKKA 317 MOKS LN FSKYKOTV WKFESDMK DN PRAVHI VKWAPOSLLAHPS KLFFI THGQLQTSEAL HYGI PL VGI PY ADOF INVKKA 318 LKGRV WKYESDEE GT TGN LKVKVL POYELLCHPN KLFI THGQLQTSEAL HYGI PL VGI PY ADOF INVKA 319 LKGRV LV WKFESDMK DN PRAVHI VKWAPOSLLAHPS KLFFI THGQLQTSEAL HYGI PL VGI PY ADOF INVAN 310 LKGRV LEF GGLQTV WKFESDMK DN PRAVHI VKWAPOSLLAHPS KLFFI THGQLQTSEAL HYGI PL VGI PY ADOF INVAN 310 LKGRV LEF GGLQTV WKFESDMK DN PRAVHI VKWAPOSLLAHPS KLFFI THGQLQTSEAL HYGI PL VGI PY ADOF INVAN 310 LKGRV LEF GGLQTV WKFESDMK DN PRAVHI VKWPOSLLAHPS KLFFI THGQLQT SEAL HYGI PL VGI PY ADOF INVAN 310 LKGRV LEF GGLQTV WKFESDMK DN PRAVHI VKWPOSLLAHPS KLFFI THGQLQT SEAL HYGI PL VGI PY ADOF INVAN 310 LKGRV LVKYEKN HY VKVPOSL PGRTDNI KI SKWLPOSDLLKHPK KAFI TGGGLQT AVACOVPN LIVP YADOF INVAN 310 KLAMF I RAFSRLPYDV VKVNKDEL PGRTDNI KI SKWLPOSDLLKHPK KAFI TGGGLQT AVACOVPN LIVP YADOF INVAN 320 RANV ASA AQI PCKVLWREDGNKP DT GLT RI YKW PONDLLGHPK RAFI THGGANG YAN YCH YHO PL ADOPDNA AHM	rE)
SfUGT33F28 SfUGT33B25 SfUGT42C10 SfUGT40R13 SfUGT40R13 SfUGT40R18 SfUGT40F19 SfUGT40F19 SfUGT33V6 HaUGT47A2 HsUGT2B7	394 EYFK GKK STEK TYEDFTNA NTVI ND-KSYREN VKURTU LODEPLSPLERAVWTEHVLRIGGARHLRGPAAN SVAEYLELEL F 391 VQIKI GLO DELSTNEEEFKNAI NTU ND-ESYRKN YKURSVMDOROPPLERAVWTEHVLRIGGARHLRGPAAN SVAEYLELEL YL 388 EENEFGYOLH KO TKEULEKLII VLN- PEFRRIKERSK WHOOROOPPLERAVWTEHVLRIGGARHLRGPAAN SVAEYLELEL YL 389 VARGFALEVKI SYSI AADUKLAI GELINN- PKYRORYKELSY YHOORY ORGAELRHWO IV NTIRGAPHLRSPALA VPLYOR YD LAA 400 VARGFALEVKI SYSI AADUKLAI GELINN- PKYRORYKELSY YHOPP ORGAELRHWO IV NTIRGAPHLRSPALA VPLYOR YD LAA 401 TILG GKAI SYAD SEKSULEG OSVIS- PEMMISAARASKI VOORI ADPLDTAVYWE WI FWCHOPP I STSRD. GFI EINIL DI VAA 404 TILG GKAI SYAD SEKSULEG OSVIS- PEMMISAARASKI VOORI ADPLDTAVYWE WI FWCHOP I STSRD. GFI EINIL DI VAA 404 ENKGFGYKVTISEDM PELDAAVKWI TD- DAYRKKSKEI SA FHORVI TPGAVPW EYVYRTHRGARHLRSPADD VPLYOK YD LAA 404 ENKGFGYKVTISEDM PELDAAVKWI TD- DAYRKKSKEI SA FHORVI TPGAVPW EYVYRTHRGARHLRSPADD VPLYOK YD LAA 402 GHKR GEKVDLSHNL FKUKVKADDKU OD PATAKIKSEI SA FHORVI TPGAVPW EYVYRTHRGARHLRSPADD VPLYOK YD LAA 403 20 KK G KULEUDTI TEEL ENA FKU DD- ORYKNA FYN KYKE ALN WE EYVRTHRGAAHLRSPALD TYP OYNLLDUVA 404 20 KK G KULEUDTI TEEL ENA FKU DD- DSYRRN EKLRSVADEPI APLERAVWTEHVLRHGGARHLRSPALD TYP OYNLLDUVA 407 KARGAAVRVDFNT SSDUL AVRASLAUTIN- ORYKNNAI DIKN FLORPIKPL DMGVYW EYVLRHRGAAHLRSPALD TYP OYNLLDUVA 407 KARGAAVRVDFNT SSDUL AALKWI DD- DSYRRN EKLRSVADEPI APLERAVWTEHVLRHGGARHLRSPALD TYP OYNLLDUVA 407 KARGAAVRVDFNT SSDUL AALKWI DD- PYYRNEN KKLSR OHOD VKULDAVWE EYVLRHRGAAHLRSPALD TYP OYNLLDUVA 407 KARGAAVRVDFNT SSDUL AALKWI DD- PYYRNEN KKLSR OHOD VKULDAVWE EYVLRHRGAAHLRSPALD TYP OYNLLDUVA 407 KARGAAVRVDFNT SSDUL AALKWI DD- PYYRNEN KKLSR OHOD VKULDAVWE EYVLRHRGAAHLRSPALD TYP OYNLLDUVA 407 KARGAAVRVDFNT SSDUL AALKWI DD- PYYNNEN KKLSR OHOD VKULDAVWE EYNRHKGAKHLRSPALD TYP OYNLLDUVA 407 KARGAAVRVDFNT SSDUL AALKWI DD- PYYNEN KKLSR OHOD VKULDAVWE EYNRHKGAKHLRSPALD TYP OYNLLDUVA	
SfUGT33F28 SfUGT33B25 SfUGT42C10 SfUGT40R13 SfUGT40R13 SfUGT40L8 SfUGT40L8 SfUGT47A12 SfUGT3V6 HaUGT47A2 HsUGT2B7	483 TULIGUATISVLEV LRS_YKW/FGSTVTG-TVTS	

Supplementary figure 2. Multiple sequence alignment of functionally characterized *S. frugiperda* UGTs active towards benzoxazinoids. The primary protein structures of SfUGTs with potential function in benzoxazinoid detoxification are compared to the human UGT2B7 and the insect UGT47A2 from *H. armigera*. The predicted signal peptide, UGT signature motif, transmembrane domain and cytoplasmic tail are shown as bars on top of the alignment. The catalytic residues, putative sheets as well as regions (donor binding regions/ DBR1 and 2, yellow box) and residues (*) responsible for sugar-donor binding in the C-terminal domain were predicted based on the crystal structure and mutational analysis reported for UGT2B7.



Supplementary figure 3. Activity profiles of SfUGT33F28 and SfUGT40L8. (A) SfUGT33F28 and SfUGT40L8 enzyme activities towards HMBOA, (B)-(D) SfUGT33F28 and SfUGT33F28-H34L site-directed mutant enzyme activities towards DIMBOA, HMBOA and MBOA respectively. Data are presented as mean \pm SEM (n=3). T-tests were performed to differentiate between the means (*P < 0.05, ***P < 0.001 according to the t-test used in the present study). #- no significant difference between the means.



Supplementary figure 4. Comparison of glucosylation activities of SfUGT33F28, SfUGT33B25 and SfUGT40L8 towards different classes of UGT substrates. The three DIMBOA-active UGT enzymes were additionally tested for activities towards a small panel of structurally diverse aglucone substrates classified generally as (A) phenolics, (B) flavonoid, (C) terpenoids, (D) coumarin and (E) alkaloid. Activities are expressed as raw peak areas from extracted LC-MS chromatograms.



Supplementary figure 5. Relative expression profile of genes encoding MBOA glycosylating enzymes across various *S. frugiperda* tissues. Relative transcript abundances of *SfUGT33B25* (A), *SfUGT42C10* (B), *SfUGT40R13* (C) and *SfUGT46A15* (D) were measured across different tissues of *S. frugiperda* by quantitative PCR (n=3, mean \pm SEM). One way repeated measures ANOVA was performed and Holm-Sidak method was applied to carry out all pairwise comparisons. The differences between individual means are denoted as small letters (a-c). #- One way ANOVA shows no significant difference in expression levels among the tissue samples.



Supplementary figure 6. Expression profile of *SfUGT33F28* and comparative account of UGT33F28 mediated glycosylation activity in two *Spodoptera* species. (A), Relative expression levels of *SfUGT33F28* across larval stages of corn and rice strain FAW caterpillars maintained on bean-based artificial diet, (B) Relative expression levels of *SfUGT33F28* during the L2 larval stage in *S. frugiperda* corn and rice strain caterpillars fed on artificial diet and maize, (C) *in vitro* enzyme activity assays towards DIMBOA from midgut extracts prepared from *S. frugiperda* corn strain, rice strain and *S. littoralis* caterpillars fed on artificial diet and corn leaves, (D) specific activity measured towards DIMBOA from crude microsomal extracts prepared from *T. ni* insect cell cultures heterologously producing *S. frugiperda* corn strain UGT33F28 and *S. littoralis* UGT33F28 homologs (SlUGT33F28 a and b). Data are presented as mean \pm SEM (n=3). One way ANOVA was performed and Holm-Sidak method was applied to carry out all pairwise comparisons. The differences between individual means are denoted as small letters (a-b).

SIUGT33F28a	1	MEI FIYLILOLSI FASHDEAARI LAVFPNPSI SHQVVFRPLTQELAK RGHEVTVI TTDPV
SIUGT33F28b	1	MEI FI YLI LOLSI FASHDEAARI LAVFPNPSI SHQVVFRPLTQELAKRGHEVTVI TTDPV
SfUGT33F28 corn strain	1	MKI LI CI TLLLSI FANYDEAARI LAVFPNPSI SHQVVFRPLTQELARRGHEVTVI TTDPV
SfUGT33F28 rice strain	1	MKILICITLLSI FANHDEAARI LAVFPNPSI SHQVVFRPLTQELARRGHEVTVI TTDPV
		1:
SIUGT33F28a	61	FPEGKAPANLTEI DVHDI SYDI WKKI FLASTKGNKDDVFMAMSAI VNALTAVVDVQLKDE
SIUGT33F28b	61	FPEGKAPANLTEI DVHDI SYDI WKKI FLASTKGNKDDV MAMSAI VNALTAVVDVQLKDE
SfUGT33F28 corn strain	61	FPEGKAPANLTEI DVHDI SYDI WRKI FLASTKONKODVVCAMIAI VTALTAVVDVQLKDE
SfUGT33F28 rice strain	61	FPEGKAPANLTET DVHDT SYDT WRKTFLAST KGNKDDVVLAMITAT VTALTAVVDVQLKDE
SIUGT33E28a	121	RVONLERDKSKOEDLELLEACVRPALVESHLYKEPVLOLSSEMATEDNYANI GAPLIHPEL
SIUGT33F28b	121	RVONLI RDKSKOFDLLLLEACVRPALVESHI YKI PVI QI SSEMATEDNYANI GAPI HPFL
SfUGT33E28 corn strain	121	KVLNLI RDKSKKFDLLLVEACVRPALVFSHI YKVPVI QI SSFMATFDNYANI GAPI HPFL
SfLIGT33E28 rice strain	121	KVLNLI RDKSKKFDLLL VEACVRPALVESHI YKVPVI QVSSEMATEDNYANI GAPI HPEL
510/91/301/20 Hite audiin		
SIUGT33F28a	181	YPALTRORLONLTM EKI KEFYNDFMLNRLY <mark>AN</mark> NMEKENEMLRSHFKDI PPI SELNNNVD
SIUGT33F28b	181	YPALTRORLONLTM EKI KEFYNDFMLNRLYANNMEKENEMLRSHFKDI PPI SELNNNVD
SfUGT33F28 corn strain	181	YPALTRORI ONLTMIEKI TEFYNDFMLNRLYDSNMEKENEMLRSHFKDIPPISELNNNVD
SfUGT33F28 rice strain	181	YPALTRORIONLTMIEKITEFYNDFMLNRLYDSNMEKENEMLRSHFKDIPPISELNNNVD
SILIGT33E28a	241	MLFLNVNPMFEGI RPVPPSVVYLGGLHOVPHKPLPKDLOTYMDSSKNGVLYVSEGTNVDP
SIUGT33F28b	241	MLFLNVNPMFEGIRPVPPSVVYLGGLHQVPHKPLPKDLQTYMDSSKNGVIYVSFGTNVDP
SfUGT33F28 corn strain	241	MLFLNVNPMFEGIRPVPPSVVYLGGLHQVPHKPLPKDLQTYMDSSKNGVIYVSFGTNVDP
SfUGT33F28 rice strain	241	MLFLNVNPMFEGI RPVPPSVVYLGGLHQVPHKPLPKDLQTYMDSSKNGVI YVSFGTNVDP
0110733520	204	
SIUG 133F208	301	TKEPADRI EVEVKTESQEPTDI EVKWINGDVEPGATINI RI AKVEPQPDELI HPKEKEPT TKEPADRI EVEVKTESQEPTDI EVKWINGDVEPGATINI RI AKVEPQPDELI HPKEKEPT
OfficT33E28 com strain	201	TKEPADRI EVEVITE SQEPTDI EVITINGOVEPONTENI NI AKWEPOPDI LI HPKEKEPIT
SIUG 133F20 com suam	301	TKEPADRI EVEVITESQEPTOFENNINGOVERGETENI KI AKVEPQEDLI HPKI KLEFT
SIUG 133F20 fice sualif	301	
SIUGT33F28a	361	QAGLQSTDEAI TAGVPLVAI PMFGDQFFNAERYENFKI GKRLSMDSLTVGEFTNAI NTVI
SIUGT33F28b	361	QAGLQSTDEAI TAGVPLVAI PMFGDQFFNAERYENFKI GKRLSMDSLTVGEFTNAI NTVI
SfUGT33F28 corn strain	361	QAGLQSTDEAI TAGVPLVAI PMFGDQSFNAERYEYFKI GKKLSMEKLTVEDFTNAI NTVI
SfUGT33F28 rice strain	361	QAGLQSTDEAI TAGVPLVAI PMFGDQSFNAERYEYIFKI GKKLSMEKLTVEEFTNAI NTVI

SIUGT33F28a	421	NDESYREN WKLRTLIODEPMSPLERAVWWTEHVLRHGGARHLRGPAANMSWAEYLELEL
SIUGT33F28b	421	NDESYRENIVKLRTLIQDEPMSPLERAVWWTEHVLRHGGARHLRGPAANMSWAEYLELEL
SfUGT33F28 corn strain	421	NDKKSYREN MYKLRTLI QDEP MSPLERAV WWTEHVLRHGGARHLRGPAAN MSWAEYLELEL
SfUGT33F28 rice strain	421	NDKSYRENMYKLRTLIQDEPMSPLERAVWWTEHVLRHGGARHLRGPAANMSWAEYLELEL
		;***
0110733508-	45.4	
SIUG 133F268	401	VETLELGELAVESVEPELERSEYN FEGSTVTSKPKKAKRS
SIUGI 33F200	401	
SIUG 133F20 COM SUBIN	401	VETLELGELAVI SVETVERSE KNIVEGSTVTSTVTSKEKKAKRS
SIDG 133F20 HOE SCAIN	401	

Supplementary figure 7. Protein sequence alignment of SfUGT33F28 from *S. frugiperda* host differentiated strains with UGT33F28 homolog from *S. littoralis* using Clustal omega.



Supplementary figure 8. Optimization of RNAi-mediated knock-down of *SfUGT33F28* in cultured Sf9 cells showing endogenous glycosylation activity. (A-B) time course quantitative analysis of silencing on Sf9 cells treated with 15 μ g and 30 μ g of dsRNA against *SfGAPDH*, 24 and 48 hours post transfection (TR: controls treated with transfection reagent), (C) quantitative analysis of *SfUGT33F28* silencing on Sf9 cells transfected with 15 μ g and 30 μ g dsRNA specific to *UGT33F28* and an unrelated dsRNA control (TR: controls treated with transfection reagent), 48 hours post transfection, (D) quantitative *SfUGT40L8* expression analysis in control and ds*UGT33F28* transfected Sf9 cells to rule out conditional functional compensation under lack of *SfUGT33F28* gene product, (E) cell morphology of Sf9 control and ds*UGT33F28* treated cells, (F) MTT cytotoxicity assay for determination of cell viability in control and dsRNA transfected Sf9 cells. Data are presented as mean ± SEM (n= 3-5 for RNAi treatments in Sf9 cells). One way

ANOVA was performed and Tukey's test was applied for all pairwise comparisons (A, B, C). The differences between the individual means are denoted as small letters (a-b). T-tests were performed to differentiate between the means (D). #- no significant difference in expression levels among the samples.

Supplementary table 1. *UGT* expression levels in Sf9 cells. Illumina reads from the Sf9 transcriptome (Shu et al., 2017) were downloaded from GenBank SRA Archives SRR5892097 on November 6, 2019. CLC Genomics Workbench 12.0.4 was used for assembly of the 87 Mio paired-end (2 x 100 bp) reads, mapping, and TPM/RPKM expression calculations. For the assembly, the graph parameters were set for automatic word size, a bubble size of 150, mismatch cost of 2, and insertion and deletion cost of 3. Mappings were performed with default parameter settings except for the length fraction (0.8), the similarity fraction (0.92) and the maximum number of hits for a read (5). RPKM indicates Reads per Kilobase of transcript per Million mapped reads, a normalized index of transcript expression. n.d., not enough reads to assemble a contig for RPKM calculations. When this was the case, the number of reads was estimated by inspection of the output of NCBI blastn against the Sequence Read Archive SRR5892097. Three housekeeping genes with different expression levels are shown for comparison: *LDH* (lactate dehydrogenase), *eIF4A* (eukaryotic initiation factor 4A) and *RpL5* (ribosomal protein L5).

Name	Contig RPKM	Contig Reads
SfUGT33B24	4.1	1,181
SfUGT33B25	4.1	1,181
SfUGT33F28	125.5	115,108
SfUGT33F29	4.8	824
SfUGT33F31	0.8	226
SfUGT33F34	2.5	774
SfUGT33J12	3.3	1,075
SfUGT33S2	1.7	404
SfUGT33T9	1.0	146
SfUGT33V6	0.2	22
SfUGT33V7	25.5	23,693
SfUGT39B21	n.d	0
SfUGT40D16	2.6	854
SfUGT40D17	3.8	1,119
SfUGT40F17	n.d	11
SfUGT40F19	1.8	810
SfUGT40L8	n.d	13
SfUGT40M10	n.d	4
SfUGT40M11	n.d	0
SfUGT40M8	1.6	656
SfUGT40Q8	6.9	2,488
SfUGT40R11	8.2	2,946
SfUGT40R12	19.4	7,997

G(1) C(T) (0) 10		
SfUGT40R13	30.2	11,726
SfUGT41B7	6.7	578
SfUGT41D8	219.4	80,133
SfUGT42B13	n.d	6
SfUGT42C10	n.d	1
SfUGT44A18	0.3	17
SfUGT46A15	0.4	34
SfUGT46B5	0.5	47
SfUGT47A12	3.6	375
SfUGT48D4	n.d	27
SfUGT50A15	n.d	16
SfeIF4A	28.2	15,061
SfLDH	228.2	110,961
SfRpL5	1605.4	342,697

Supplementary table 2. Summary of SfUGT33F28 glycosylation activities observed across a spectrum of substrates. ANOVA was employed based on the net amounts of glucoside produced by UGT33F28 after subtracting what was produced by the corresponding amount of non-transfected control (NTC). When the glucoside formation did not differ significantly from NTC, the enzyme was considered not active (-) towards the substrate. The activity was scored on a scale of 1 to 3, 1 corresponding to the lowest and 3 corresponding to the highest activity.

	SfUGT33F28 enzyme
Substrate	activity
Benzoxazinoid-related	
Indole	-
IAA	3
BOA	-
6-OH-BOA	-
MBOA	-
DIMBOA	3
HMBOA	3
Phenolic compounds	
4-nitrophenol	1
1-naphthol	-
Salicyl aldehyde	-
Vanillin	1
Catechol	-
L-Tyrosine	-
L-DOPA	-
Dopamine	-
Flavonoids	
Quercetin	-
Terpenoids	
Menthol	-
Gossypol	-
Coumarins	
Esculetin	1
Alkaloids	
Capsaicin	-

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
SfUGT33V7	AGGACACGTGCA <u>ATG</u> TCGTTC	ACTCCTTTTAACCTTACCAGCG
SfUGT40R1	GTGCCGATA <u>ATG</u> GCCTTAGTAT	ATTCCTTTTTTTACTATTTACAATTT
SfUGT33F2	TAGAAA <u>ATG</u> AAGATATTAATTT	ACTACGTTTGGCTTTTTTAGGTT
SfUGT40D1	ATTGAGAAT <u>ATG</u> GAAAGAATG	GTTGTTCTTCTTTTGGCTTCC
SfUGT50A1	GCAAAA <u>ATG</u> CATAGGTGGA	CAATTTCTTCTTTGTTGCCG
SfUGT40F1	CTGTCAAAA <u>ATG</u> GAGATAATT	ATTCTTCTTAAGTTTTGGGCTCT
SfUGT41D8	ACA <u>ATG</u> AAGCTAAGTATCCTGT	ATTCCTTTTTCTTTTTACTAGATTTCT
SfUGT41B7	CCGTTCGAA <u>ATG</u> AAGGCT	ATTCTTCTTCTTATGCGCTAGTCT
SfUGT33B2	TTGCGTGTAATC <u>ATG</u> TCTACCG	ATTGTCTTTTAACTTACTATATTTCA
SfUGT33F3	CGCGTTACA <u>ATG</u> TCGGTATTAC	ATTCCGTTTCTTCTTCACACTAGTTT
SfUGT42C1	AATAAA <u>ATG</u> AAGTTTAATAAT	CTTAGTCTTGAACTCATAATAACGA
SfUGT44A1	GCATTCAAA <u>ATG</u> ACAAAACAG	TTCGAACTTCACTCTTTTTGAGGA
SfUGT40Q8	AGCAATAAA <u>ATG</u> AACAAATGG	ACTCCTTTTCTTCTTCTTTTATCTTCAC
SfUGT33F3	ATAAAA <u>ATG</u> TTATCTTTTGTGT	ATTATTTTTAACTTTAGGAGTAGAA
SfUGT40R1	TTGTGCGTGCCAATA <u>ATG</u> GCG	ATTCTTCTTATCCTTTTTCTTACTAT
SfUGT40M1	TGTGTCACA <u>ATG</u> AGGTTGCCA	ATTAGTTTTCTTCTGATCATATTTTG
SfUGT48D4	GTAGCTGACATGAGGCGG	TTCGTTTTTCTTGTCTTCAAAATA
SfUGT39B2	AAA <u>ATG</u> GACCCCCTAAAAATA	ATGTTTCTTCTGTTTTTTCGTTG
SfUGT33J1	CTCAAG <u>ATG</u> GTTGGGTTT	ACTACTCTTCACTTTAGTCTTAATT
SfUGT40L8	TTTGTAAAT <u>ATG</u> AAGCATAAA	GTGACTCTTCTCTTTTTTATTAGTCG
SfUGT33T9	TCGATCACA <u>ATG</u> TCTGTGC	ACTCCGTTTAATCTTAACATCACTA
SfUGT42B1	TCCATA <u>ATG</u> AAGCCCGGC	CGTTCGCTTCGCCTTTTT
SfUGT46A1	TTGCCCACG <u>ATG</u> CG	ATGCAACTTCTCCTTCTTATTCG
SfUGT33F2	AGCCAGACA <u>ATG</u> AGGTATTTT	ACTACGTTTTGATTTTGATTTTTAA
SfUGT40R1	ATTTGTAAA <u>ATG</u> GCATTAGCAA	ATTCCTCTTCTTACTGTCTACAATTT
SfUGT40M8	AGAACAAAA <u>ATG</u> GCGGATT	ATTCTTCTTCTCTTTTATATTCTTCT
SfUGT40D1	TCAGCCATC <u>ATG</u> CAACG	ATTCTTCTTCTTTTTTTCCTCCTGTAT
SfUGT40M1	TGTGCCGTA <u>ATG</u> AAGTGG	CATTATTTTCTTTTTGCATATTTTCA
SfUGT33B2	CGTTCGATA <u>ATG</u> TTCTTCCC	ATTCTTCTTAACTTTCTTCTGCTTTG
SfUGT46B5	TATCTCAAG <u>ATG</u> CTACTCCCG	ACTCTGTTTAAGTTTCTCAGTCTTTT
SfUGT47A1	AGCACAGAGAGATGAGGCCG	ATTGGGTTTCACTTTGTACTTCC
SfUGT40F1	TTAATAATA <u>ATG</u> GAAAAGTTA	ATTCTTCTTCTCTTTTATATTCTTCT
SfUGT33S2	AGTGAGAACATGTTGCTGTGC	GTCTGCCTTAAGCTTCTTTAAAGT
SfUGT33V6	GAAACCAACATGTCTCTAGTA	TGCGCGCTTAATTTTAATAATAG

Supplementary table 3. List of primers used for the amplification of full length SfUGTs described in the study.

Supplementary table 4. List of primers used for site-directed mutagenesis in the catalytic region of *SfUGT33F28*.

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
SfUGT33F28	ACCCTTCAATCAGCC <u>TG</u> CA	TCGGAAAACCACTTG <u>CA</u> GGCTGATT
AT-TG	AGTGGTTTTCCGA	GAAGGGT

Supplementary table 5. List of primers used for quantitative real time PCR (OGS2.2 gene annotation in brackets).

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
SfUGT33F28	GTCGTTCAACGCTGAGAGATA	CGCAGTTTTACCATGTTCT
(GSSPFG0001823700	С	CC
1/ UGT33-13)		
SfUGT33B25	CAACGAATGAAGAGGAATTCA	AGGCTGGTCGTACATCAC
(GSSPFG0002458200	AG	G
1/ UGT33-19)		
SfUGT42C10	CCGCTGCTATCGAAGAAAAC	TCATGCCATACTTTGGATC
(GSSPFG0003536900		G
1/UGT42-01-hap)		
SfUGT40L8	AGCAAACATCGCTGGTTACC	TAAATGACGCCGTGTTTCG
(GSSPFG0000846600		С
1/UGT40Lb)		
SfUGT40R13	TCGCTAGAGGCTTTGCTTTG	TCGGTCGTGGTAGATGTAG
(GSSPFG0001460100		G
1/		
GSSPFG00005183001		
/ UGT40-05)		
SfUGT46A15	TCGAAGGATTGCAGAGTGTG	TGCTGTGTCCAAAGGATCT
(GSSPFG0003214900		G
1/		
GSSPFG00021909001		
/ UGT46-01)		
SfGAPDH	TGGATAACTTTGGCGAGAGG	TGTCAACCTTGACGCTTAC
(GSSPFG0002813900		G
1/		
GSSPFG00007681001		
)		
SfRPL10	ATTGGACAGCCCATCATGTC	AAGCCCCATTTCTTGGAGA
(GSSPFG0001220500		С
1)		

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
SfUGT33F	TAATACGACTCACTATAGGGC	CCCTATAGTGAGTCGTATTAGTCCA
28 RNAi	CAAGTACCTCATAAACCAT	GGCAGAACATCA
SfGAPDH	TAATACGACTCACTATAGGGG	CCCTATAGTGAGTCGTATTATTGGTC
RNAi	AAGGGCATTCTTGACTA	TGGATGTACTTG

Supplementary table 6. List of	f primers used for the	e preparation of RNAi constructs.
--------------------------------	------------------------	-----------------------------------

Shu, B., Zhang, J., Sethuraman, V., Cui, G., Yi, X., and Zhong, G. (2017). Transcriptome analysis of Spodoptera frugiperda Sf9 cells reveals putative apoptosis-related genes and a preliminary apoptosis mechanism induced by azadirachtin. *Scientific reports* 7, 1-13.