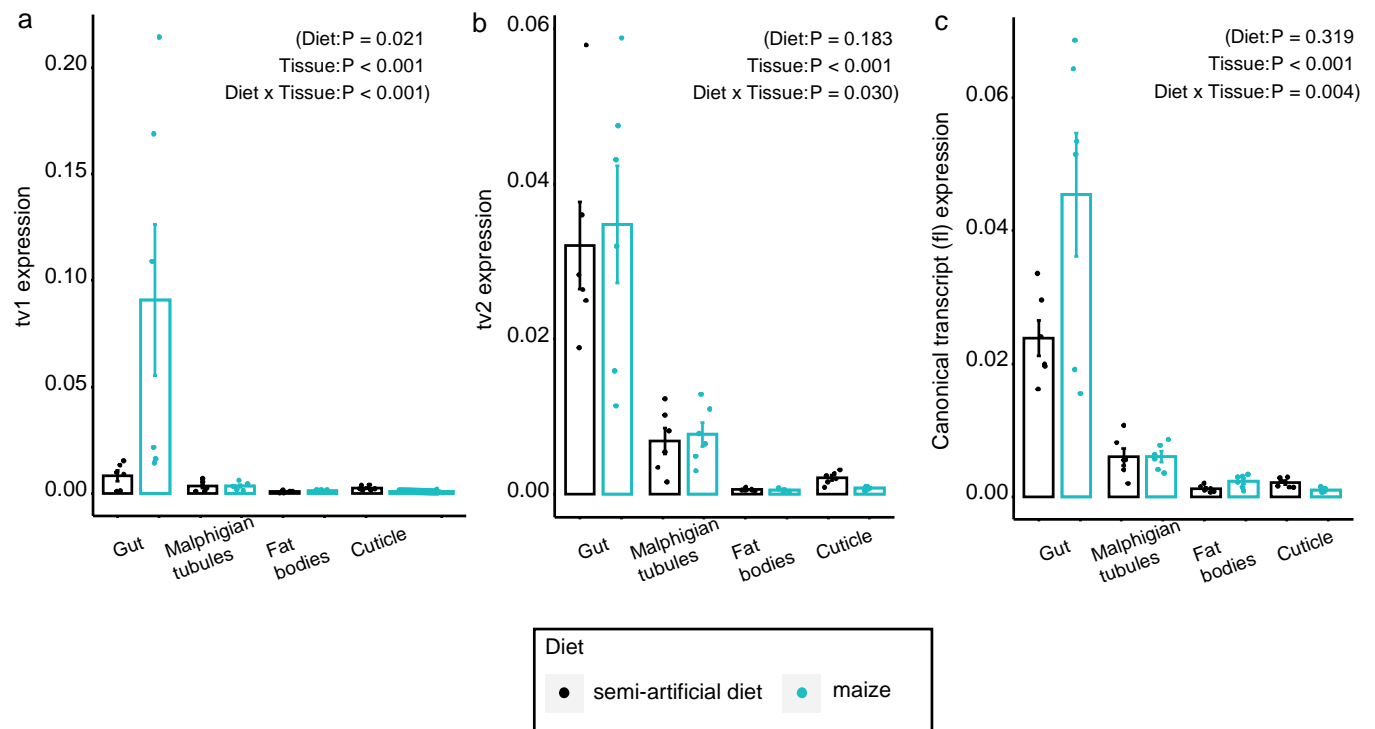


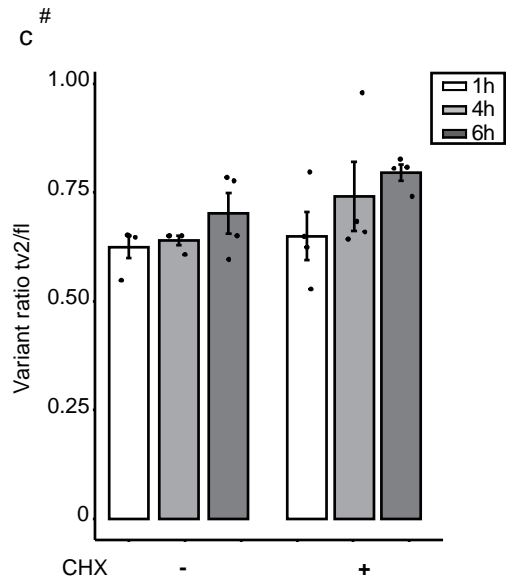
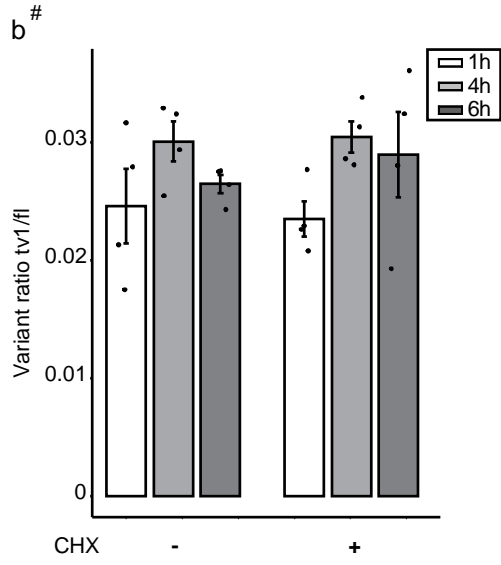
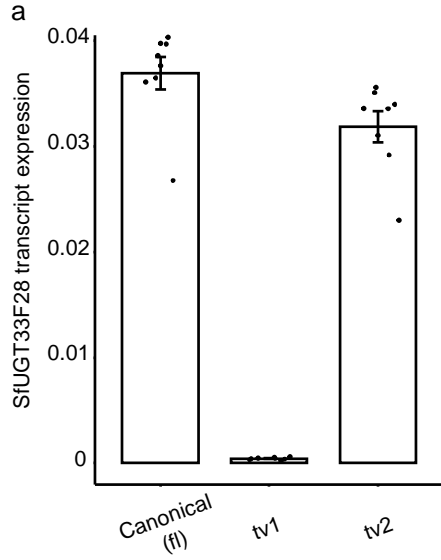
## SUPPLEMENTARY INFORMATION

### Alternative transcript splicing regulates UDP-glucosyltransferase-catalyzed detoxification of DIMBOA in the fall armyworm (*Spodoptera frugiperda*)

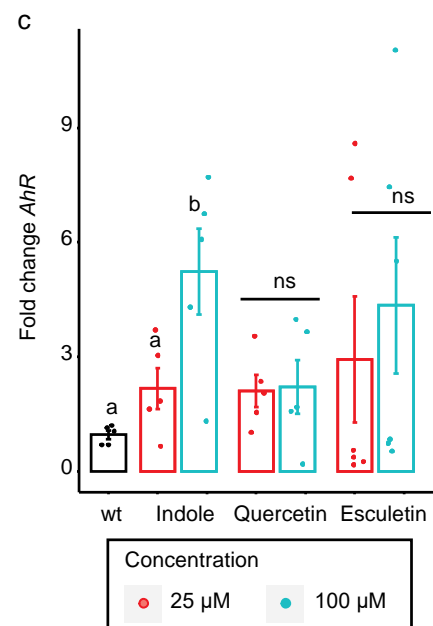
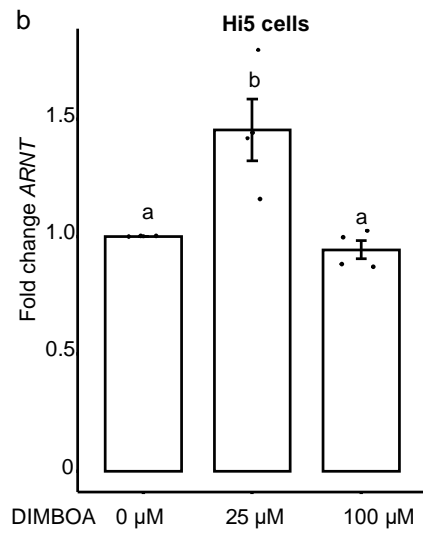
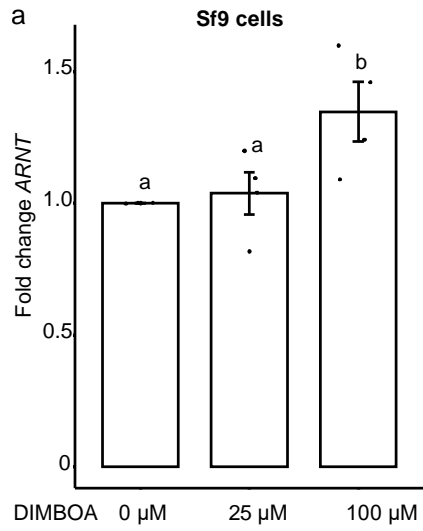
Bhawana Israni, Katrin Luck, Samantha C.W. Römhild, Bettina Raguschke, Natalie Wielsch, Yvonne Hupfer, Michael Reichelt, Aleš Svatoš, Jonathan Gershenzon, Daniel Giddings Vassão



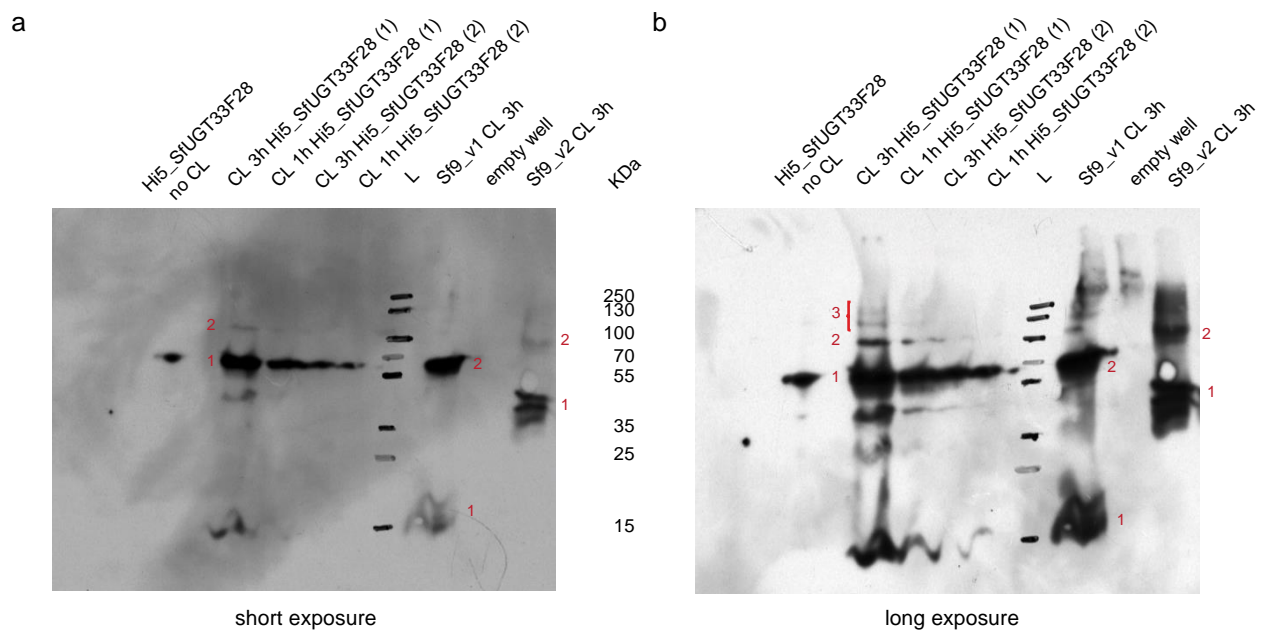
**Supplementary figure 1. Real time PCR quantification of expression of *SfUGT33F28* variants in various tissues from caterpillars maintained on semi-artificial (pinto bean-based) diet and maize leaves. Data are presented as mean  $\pm$  SEM (n=6). All values were log transformed to meet the criteria for normality. Two-way repeated measures ANOVA was performed on log transformed data, and Bonferroni t-test was applied to carry out all pairwise comparisons. tv1, transcript variant 1; tv2, transcript variant 2; fl, full-length.**



**Supplementary figure 2. Expression profile of *SfUGT33F28* transcript variants in insect derived Sf9 cells. a) Quantitation of *SfUGT33F28* transcript levels in untreated Sf9 cells by real time PCR, b) quantitation of *SfUGT33F28* variants tv1 and tv2 with respect to full length canonical transcript upon treatment of Sf9 cells with cycloheximide (CHX) over a period of 1- 6 hours. Data are presented as means  $\pm$  SEM (n=4-6). One-way ANOVA was performed, and Tukey's test was applied to carry out pairwise comparisons (b-c). # indicates no significant differences between the means. tv1, transcript variant 1; tv2, transcript variant 2; fl, full-length.**

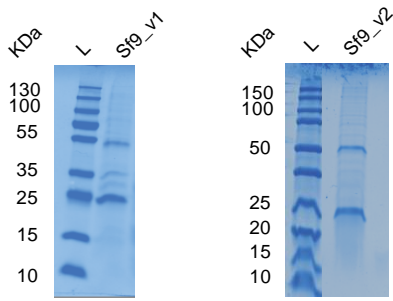


**Supplementary figure 3. *ARNT* expression in response to DIMBOA and specificity of *AhR* responses across a range of chemically diverse substrates. a-b) Quantitation of *ARNT* transcript levels in untreated and DIMBOA-treated Sf9 and Hi5 insect cells by real time PCR (n=4-5), c) quantitation of *AhR* transcript levels in insect cells upon treatment with indole, quercetin and esculetin respectively (n=4). All data are presented as mean  $\pm$  SEM. One-way ANOVA on ranks was performed and Tukey's test was applied to carry out all pairwise comparisons (a, b). One-way ANOVA was performed, and Tukey's test was applied to carry out all pairwise comparisons (c). Small letters on the bars indicate significant differences at  $P < 0.05$ . wt, untreated cells.**



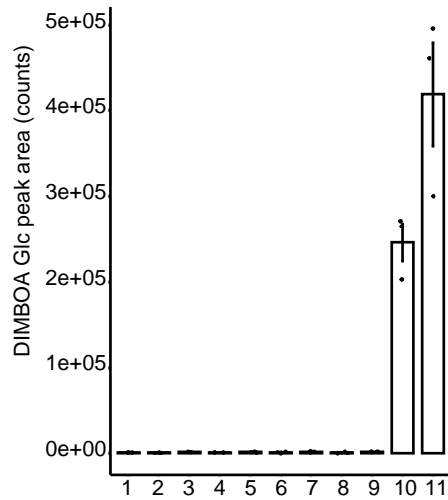
**Supplementary figure 4. Cross-linking assays to show the interaction of catalytically inactive UGT variants with other microsomal proteins. Cross-linking assays were performed with Hi5 insect cell microsomes producing UGT33F28 full-length protein, Sf9 insect cell microsomes producing UGT33F28 variants UGT33F28\_v1 and UGT33F28\_v2 (c) using dimethyl suberimide (DMS). The oligomers observed upon cross-linking are indicated by numbers in red (1: monomer, 2: dimer (homo/hetero), 3: higher order oligomer, L: pre-stained PAGE ruler).**

**a Affinity purified (anti V5) extracts**



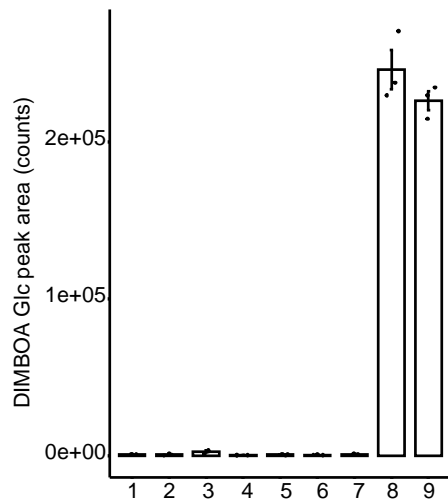
(L: pre-stained PAGE ruler)

**b**



- 1 Hi5 wt
- 2 Hi5 UGT33F28\_v1
- 3 Hi5 wt + UGT33F28\_v1
- 4 UGT40R12
- 5 UGT40R12 + UGT33F28\_v1
- 6 UGT40R13
- 7 UGT40R13 + UGT33F28\_v1
- 8 UGT40F19
- 9 UGT40F19 + UGT33F28\_v1
- 10 UGT33F28
- 11 UGT33F28 + UGT33F28\_v1

**c**



- 1 Hi5 wt
- 2 Hi5 UGT33F28\_v2
- 3 Hi5 wt + UGT33F28\_v2
- 4 UGT33T9
- 5 UGT33T9 + UGT33F28\_v2
- 6 UGT33S2
- 7 UGT33S2 + UGT33F28\_v2
- 8 UGT33F28
- 9 UGT33F28 + UGT33F28\_v2

**Supplementary figure 5. Interaction and catalytic activities of UGT33F28 variants with full-length UGT33F28 and other proteins predicted from proteomic analyses of bound fractions. a)**

Electrophoretic separation of V5 affinity purified fractions from Sf9 cells overexpressing UGT33F28 variants on a 4-20% Mini-PROTEAN tris glycine gel for proteomics and subsequent staining with Coomassie blue, b-c) *In vitro* enzymatic assays were performed using UGTs heterologously produced in Hi5 cells with DIMBOA as substrate. Data are represented as mean  $\pm$  SEM (n=3).

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
<i>SfUGT33F28</i> variant exon 1qPCR	CTAGCTAGAAGAGGGCTTACACCA AG	GATCCTGTCAGCTGGTAGTTTAG
<i>SfUGT33F28</i> canonical exon 1 qPCR	TGTTCCCGAACCCCTTCAATC	GACAGTAACTTCATGGCCTCTT
<i>SfUGT33F28</i> variant exon 3 qPCR	CAAACGTGGACCTGCAGCCA	CGTTTGGCTTTTTTAGGTTTAC
<i>SfUGT33F28</i> canonical exon 3 qPCR	GTCGTTCAACGCTGAGAGATAC	CGCAGTTTTACCATGTTCTCC
<i>SfRPL10</i> qPCR	ATTGGACAGCCCATCATGTC	AAGCCCCATTTCTTGGAGAC

Supplementary table 1. Primer pairs used for distinguishing between *SfUGT33F28* mRNA variants.

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
Exon 1-3 splicing variant	AACTCCCATGAGGGTAA GC	CTCTTATCATTAATAACAGTGTTGATAG C
Exon 1-2 splicing variant	AACTCCCATGAGGGTAA GC	GAGAAGATCAGGCTGTGGC

**Supplementary table 2. Primer pairs used for preparation of *SfUGT33F28* minigene.**

<b>Match name/ Accession</b>	<b>E- value</b>	<b>% Identity</b>	<b>Description</b>
XP_021187219.1	6.700e-279	99.38	aryl hydrocarbon receptor protein 1 [Helicoverpa armigera]
XP_022815375.1	6.700e-279	99.38	aryl hydrocarbon receptor isoform X2 [Spodoptera litura]
PCG63465.1	1.900e-278	100.00	hypothetical protein B5V51_12264, partial [Heliothis virescens]
XP_026735439.1	2.500e-278	99.17	aryl hydrocarbon receptor [Trichoplusia ni]
XP_026761162.1	1.300e-271	96.68	aryl hydrocarbon receptor [Galleria mellonella]
XP_028157966.1	3.700e-269	96.69	aryl hydrocarbon receptor isoform X1 [Ostrinia furnacalis]
XP_026323827.1	1.100e-268	95.44	aryl hydrocarbon receptor isoform X2 [Hyposmocoma kahamanoa]
XP_022815374.1	5.300e-268	99.35	aryl hydrocarbon receptor isoform X1 [Spodoptera litura]
XP_013165847.1	6.100e-264	94.19	PREDICTED: aryl hydrocarbon receptor, partial [Papilio xuthus]
KPJ03486.1	6.100e-264	94.19	Aryl hydrocarbon receptor [Papilio xuthus]

**Supplementary table 3. Description of top scoring BLAST hits using putative SfAhR protein sequence as a query against GenBank protein database.**

<b>Match name/ Accession</b>	<b>E- value</b>	<b>% Identity</b>	<b>Description</b>
----------------------------------	-----------------	-------------------	--------------------



XP_022837764.1	6.000e-291	98.63	aryl hydrocarbon receptor nuclear translocator-like protein 2 [Spodoptera litura]
XP_021188262.1	2.100e-280	94.72	aryl hydrocarbon receptor nuclear translocator-like protein 2 [Helicoverpa armigera]
XP_026734341.1	1.500e-273	92.37	aryl hydrocarbon receptor nuclear translocator-like protein 1 [Trichoplusia ni]
AZL94115.1	2.500e-273	92.56	juvenile hormone methoprene tolerant protein [Mythimna separata]
XP_026492150.1	8.800e-226	75.98	aryl hydrocarbon receptor nuclear translocator-like protein 1 [Vanessa tameamea]
KOB74415.1	1.300e-221	77.12	Methoprene tolerant protein 1 [Operophtera brumata]
ANZ54967.1	3.200e-220	77.17	juvenile hormone receptor methoprene-tolerant [Plodia interpunctella]
XP_013186860.1	5.500e-220	76.28	PREDICTED: aaryl hydrocarbon receptor nuclear translocator-like protein 1 [Amyelois transitella]
XP_028173519.1	3.600e-219	75.10	aryl hydrocarbon receptor nuclear translocator-like protein 2 [Ostrinia furnacalis]
XP_026761772.1	1.400e-215	73.62	aryl hydrocarbon receptor nuclear translocator-like

		protein 1 [Galleria mellonella]
--	--	---------------------------------

**Supplementary table 4. Description of top scoring BLAST hits using putative SfARNT protein sequence as a query against GenBank\_protein database.**

<b>Organism</b>	<b>Description</b>	<b>Predicted NLS</b>
<i>Homo sapiens</i> (NP_001659)	Aryl hydrocarbon receptor nuclear translocator isoform 1	RAIKRRPGLD (monopartite)
<i>Drosophila melanogaster</i> (AAB69695)	Aryl hydrocarbon receptor nuclear translocator-like protein	-
<i>Spodoptera frugiperda</i> (GSSPFG00020774001)	Aryl hydrocarbon receptor nuclear translocator	RFDSRKRRKVD, DSRKRRKVDC (monopartite)
<i>Trichoplusia ni</i> (XP_026734341.1)	Aryl hydrocarbon receptor nuclear translocator-like protein 1	RFDSRKRRKID, DSRKRRKIDC (monopartite)

**Supplementary table 5. ARNT proteins from select organisms with their predicted nuclear localization signal (NLS). Prediction of NLS in ARNT proteins from *H. sapiens*, *D. melanogaster*, *S. frugiperda* and *T.ni* was carried using the cNLS mapper. NLS prediction was carried out using a cut-off score of 5.**

<b>Gene</b>	<b>Forward primer (5'-3')</b>	<b>Reverse primer (5'-3')</b>
<i>SfAhR</i> / <i>TnAhR</i> qPCR	GCTTACCGCTTCCAACAAAG	GCGGTCGATGCGTACTAATA
<i>SfARNT</i> / <i>TnARNT</i> qPCR (1)	GTGTTATTGGAGAAGATTATGGA	CATCATTGGCTGCTTGTT
<i>SfARNT</i> / <i>TnARNT</i> qPCR (2)	GTATGCCTACCGCAAAGCAAG	GCACTATATCCGACAACAACGAC

**Supplementary table 6. Real time primer pairs used for estimation of *AhR* and *ARNT* transcript levels in insect cells.**

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
Canonical DNA probe 1 (wt) <i>SfUGT33F28</i>	TACGCGTGTTATTGAACTTC	GAAGTTCAATAACACGCGTA
Mutant DNA probe 2 ( $\Delta$ ahr/arnt) <i>SfUGT33F28</i>	TATTATTGAACTTCAAAACATA C	GTATGTTTTGAAGTTCAATAATA
Unlabeled canonical DNA probe 1 (wt) <i>SfUGT33F28</i>	TACGCGTGTTATTGAACTTC	GAAGTTCAATAACACGCGTA
<i>SfUGT33F28</i> : $\Delta$ ahr/arnt	AATAGATATTATTGAACTTCAA AACATACGG	TCAATAATATCTATTTGTAGTTAA ATGTTAGA

**Supplementary table 7. Probes used for EMSA for detection of AhR-ARNT binding to *SfUGT33F28* gene, and subsequent deletion mutagenesis in *SfUGT33F28* 5' UTR.**

Accession	Mascot DB	Mascot Description	Mascot Good Peptides	Mascot Score
GSSPFG000137 17001	Spodoptera_Frugiperda_PRO _20151204_X	Putative aryl hydrocarbon receptor	1	57
GSSPFG000067 64001	Spodoptera_Frugiperda_PRO _20151204_X	EH domain-containing protein isoform	4	267

GSSPFG000059 66001.4	Spodoptera_Frugiperda_PRO _20151204_X	kinetochore protein NDC80 homolog	4	236
GSSPFG000193 38001.3	Spodoptera_Frugiperda_PRO _20151204_X	Heat shock protein 70-2	3	182
GSSPFG000112 92001.1	Spodoptera_Frugiperda_PRO _20151204_X	ATP-dependent RNA helicase WM6	2	170
GSSPFG000036 02001.1	Spodoptera_Frugiperda_PRO _20151204_X	Protein disulfide isomerase	2	141
GSSPFG000185 84001	Spodoptera_Frugiperda_PRO _20151204_X	putative ATP- dependent RNA helicase DDX3X	2	137
GSSPFG000345 79001.2	Spodoptera_Frugiperda_PRO _20151204_X	Eukaryotic initiation factor 2%CE%B3	2	137
GSSPFG000257 50001.2	Spodoptera_Frugiperda_PRO _20151204_X	gene=putative pre- mRNA-splicing factor ATP- dependent RNA helicase DHX15- like	2	128
GSSPFG000345 79001.2	Spodoptera_Frugiperda_PRO _20151204_X	initiation factor	2	127
GSSPFG000250 91001.1	Spodoptera_Frugiperda_PRO _20151204_X	26S regulatory subunit	3	124
GSSPFG000100 67001	Spodoptera_Frugiperda_PRO _20151204_X	histone deacetylase Rpd3 partial [Anopheles gambiae]	3	117
GSSPFG000106 30001.2	Spodoptera_Frugiperda_PRO _20151204_X	putative loquacious	1	89
GSSPFG000085 03001	Spodoptera_Frugiperda_PRO _20151204_X	cyclin-K-like gene=GSSPFG000 08503001 [Spodoptera litura]	1	72
GSSPFG000359 47001.2	Spodoptera_Frugiperda_PRO _20151204_X	Suppressor of variegation 3-9	1	70
GSSPFG000271 94001	Spodoptera_Frugiperda_PRO _20151204_X	KH domain- containing RNA- binding signal transduction- associated protein 2-like isoform [Helicoverpa armigera] X1 [Spodoptera litura]	1	70

GSSPFG000128 08001.3	Spodoptera_Frugiperda_PRO _20151204_X	26S proteasome regulatory ATPase subunit 10b	1	68
GSSPFG000051 83001.3	Spodoptera_Frugiperda_PRO _20151204_X	gene=UDP- glycosyltransferase- 40-05 gene=UDP- glycosyltransferase- 40-11	1	61
XP_021172547. 1	NCBIprot_20200524	period circadian protein homolog 2- like [Fundulus heteroclitus]	5	397
XP_022815275. 1	NCBIprot_20200524	EH domain- containing protein 3 [Spodoptera litura]	4	267
AAF54856.2	NCBIprot_20200524	Putative achaete scute target 1, isoform B [Drosophila melanogaster]	1	73

**Supplementary table 8. Overview of protein hits predicted from the insect cell- nuclear protein fraction bound with biotinylated *SfUGT33F28* DNA probe using streptavidin-based affinity purification.**

<b>Accession</b>	<b>Mascot DB</b>	<b>Mascot Description</b>	<b>Mascot Good Peptides</b>	<b>Mascot Score</b>
GSSPFG0001823 7001.3	Spodoptera_Frugiperda_PRO_2 0151204_X	UDP- glycosyltransf erase-33-13	3	206
GSSPFG0000518 3001.3	Spodoptera_Frugiperda_PRO_2 0151204_X	UDP- glycosyltransf erase-40-05	2	69
GSSPFG0000826 8001.2	Spodoptera_Frugiperda_PRO_2 0151204_X	CYP9A30	3	192
GSSPFG0000435 2001	Spodoptera_Frugiperda_PRO_2 0151204_X	UDP- glycosyltransf erase-40-11	2	151
GSSPFG0001576 7001	Spodoptera_Frugiperda_PRO_2 0151204_X	Heat shock protein 60	5	319

GSSPFG0003465 5001	Spodoptera_Frugiperda_PRO_2 0151204_X	Protein disulfide isomerase	10	893
GSSPFG0000360 2001	Spodoptera_Frugiperda_PRO_2 0151204_X	Protein disulfide isomerase	10	893
GSSPFG0002232 7001	Spodoptera_Frugiperda_PRO_2 0151204_X	Niemann-Pick C1 protein isoform [Spodoptera litura]	7	243
GSSPFG0001525 3001	Spodoptera_Frugiperda_PRO_2 0151204_X	probable citrate synthase mitochondrial	14	712
GSSPFG0002610 6001.1	Spodoptera_Frugiperda_PRO_2 0151204_X	isocitrate dehydrogenas e [NADP] cytoplasmic- like [Spodoptera litura]	5	249
GSSPFG0000113 4001.1	Spodoptera_Frugiperda_PRO_2 0151204_X	cytochrome uncharacterize d oxidase protein subunit	3	183
GSSPFG0000762 6001	Spodoptera_Frugiperda_PRO_2 0151204_X	ATP PROTEIN synthase subunit	3	171
GSSPFG0003196 9001.5	Spodoptera_Frugiperda_PRO_2 0151204_X	molecular heat chaperone shock DnaK protein	8	458
GSSPFG0000869 1001.3	Spodoptera_Frugiperda_PRO_2 0151204_X	voltage- dependent anion- selective channel	2	157
GSSPFG0002422 2001.2	Spodoptera_Frugiperda_PRO_2 0151204_X	UGT41-03	2	181

**Supplementary table 9. Overview of protein hits predicted from V5 affinity purified fractions of Sf9 cells overexpressing SfUGT33F28\_v1.**

Accession	Mascot DB	Mascot Description	Mascot Good Peptides	Mascot Score
GSSPFG0001823 7001.3	Spodoptera_Frugiperda_PRO_2 0151204_X	UDP-glycosyltransferase-33-13	1	58
GSSPFG0001403 9001	Spodoptera_Frugiperda_PRO_2 0151204_X	CYP314A1	1	58
GSSPFG0001008 7001	Spodoptera_Frugiperda_PRO_2 0151204_X	carboxylesterase 021c	1	61
GSSPFG0001371 7001	Spodoptera_Frugiperda_PRO_2 0151204_X	Putative aryl hydrocarbon receptor	1	57
GSSPFG0000983 1001.2	Spodoptera_Frugiperda_PRO_2 0151204_X	CYP6AN4	1	57
GSSPFG0002680 0001.3	Spodoptera_Frugiperda_PRO_2 0151204_X	CYP9A27PARTIAL	1	56
GSSPFG0000036 5001.3	Spodoptera_Frugiperda_PRO_2 0151204_X	Serine protease	2	110
GSSPFG0001576 7001.3	Spodoptera_Frugiperda_PRO_2 0151204_X	Heat shock protein 60	1	63
GSSPFG0000183 4001.5	Spodoptera_Frugiperda_PRO_2 0151204_X	carboxylesterase 016d	1	58

**Supplementary table 10. Overview of protein hits predicted from V5 affinity purified fractions of Sf9 cells overexpressing SfUGT33F28\_v2.**

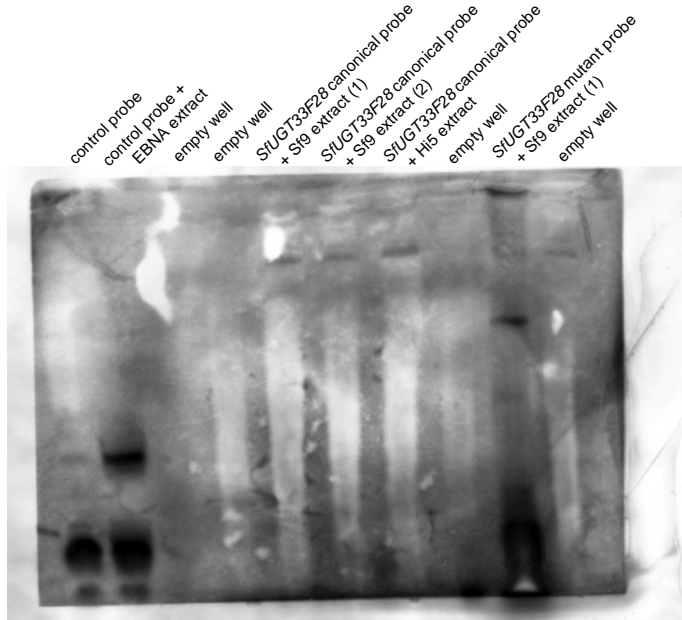
Gene	Accession	Forward primer (5'-3')	Reverse primer (5'-3')
<i>SfUGT33</i> <i>F28</i>	GSSPFG00 018237001. 3	ACCATGAAGATATTAATTTG TATAACTCTG	CGTTTGGCTTTTTTAGGTT TAC
<i>SfUGT33</i> <i>F28</i> tv 1	GSSPFG00 018237001. 3	ACCATGAAGATATTAATTTG TATAACTCTG	GGTCACAGTACCAGTCAC AGT
<i>SfUGT33</i> <i>F28</i> tv 2	GSSPFG00 018237001. 3	ACCATGAAGATATTAATTTG TATAACTCTG	GTTCGAGGTACTCTGCCC A

<i>SfUGT40</i> <i>R12</i>	GSSPFG00 004145001	GTGCCGATA <u>ATGGC</u> CTTAGT ATTA	ATTCCTTTTTTTACTATTT ACAATTTTCTT
<i>SfUGT40</i> <i>R13</i>	GSSPFG00 005183001	TTGTGCGTGCCAATA <u>ATGGC</u> G	ATTCTTCTTATCCTTTTTTC TACTATTTAC
<i>SfUGT40</i> <i>F19</i>	GSSPFG00 005185001	TTAATAATA <u>ATGG</u> AAAAGTT AATATGTTTT	ATTCTTCTTCTCTTTTATA TTCTTCTGA
<i>SfUGT33</i> <i>T9</i>	GSSPFG00 025012001	TCGATCACA <u>ATGT</u> CTGTGC	ACTCCGTTTAATCTTAAC ATCACTAA
<i>SfUGT33</i> <i>S2</i>	GSSPFG00 016675001	AGTGAGAAC <u>ATGT</u> TGCTGTG C	GTCTGCCTTAAGCTTCTTT AAAGT

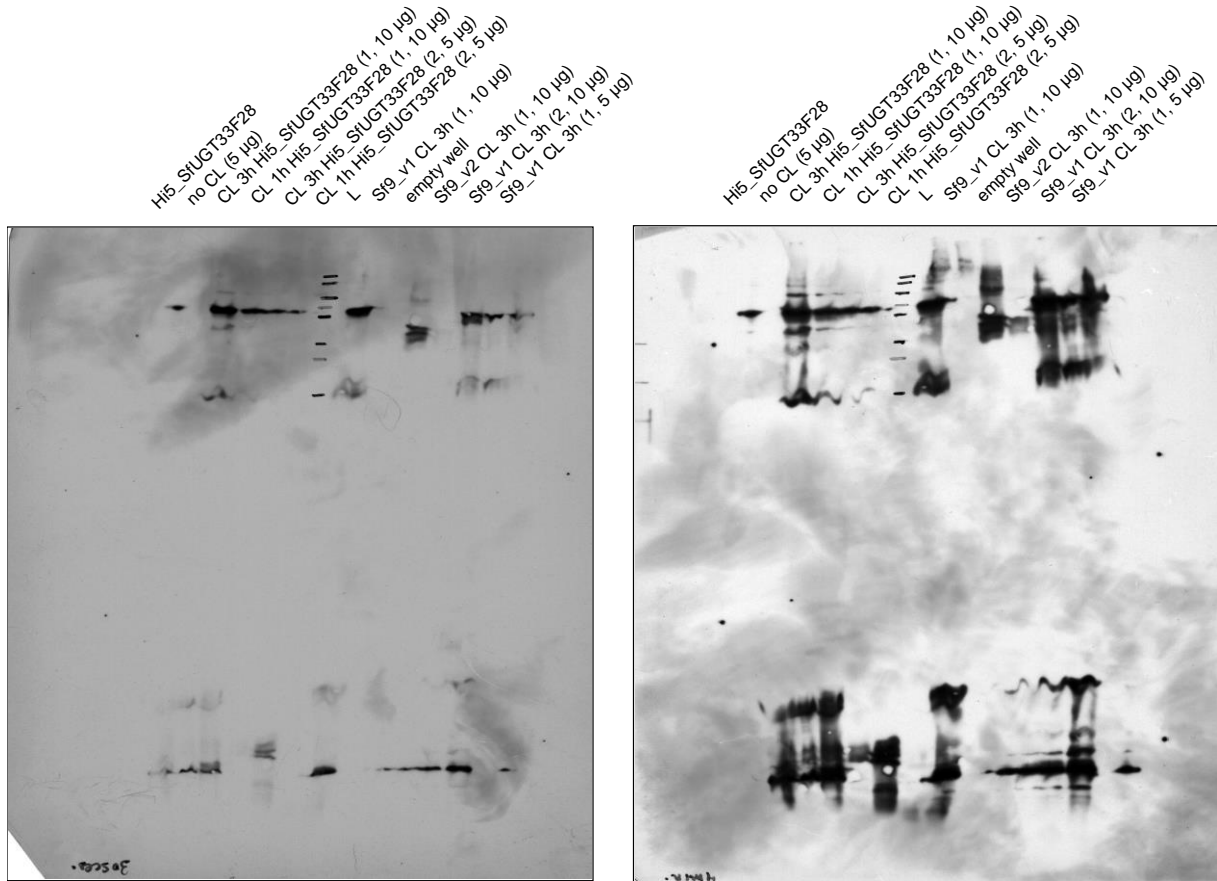
**Supplementary table 11. Primer pairs used for cloning and heterologous expression of SfUGTs.**



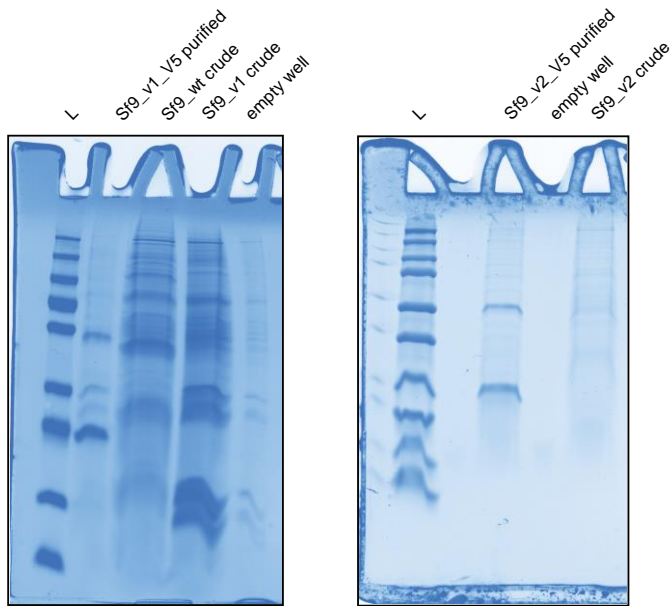
## Original blots and gels



**Electrophoretic mobility shift assay (EMSA) to demonstrate the binding of the AhR-ARNT protein complex from insect cells to the nucleotide sequence derived from *SfUGT33F28* 5' UTR.**



**Cross-linking assays to show the interaction of catalytically inactive UGT variants with other microsomal proteins. Cross-linking assays were performed with Hi5 insect cell microsomes producing UGT33F28 full-length protein, Sf9 insect cell microsomes producing UGT33F28 variants UGT33F28\_v1 and UGT33F28\_v2 (c) using dimethyl suberimide (DMS). L: pre-stained PAGE ruler.**



**Electrophoretic separation of V5 affinity purified fractions from Sf9 cells overexpressing UGT33F28 variants on a 4-20% Mini-PROTEAN tris glycine gel for proteomics and subsequent staining with Coomassie blue. L: pre-stained PAGE ruler.**