Table S1. List of new GH5_2 proteins from Cerambycidae, including information on species and accession numbers.

Gene name	Species	subfamily	Genbank accession
PLE1	*Phymatodes lengi*	Cerambycinae	OM585467
PLE2	*Phymatodes lengi*	Cerambycinae	OM585468
PLE3	*Phymatodes lengi*	Cerambycinae	OM585469
PTE1	*Phymatodes testaceus*	Cerambycinae	OM585470
PTE2	*Phymatodes testaceus*	Cerambycinae	OM585471
PTE4	*Phymatodes testaceus*	Cerambycinae	OM585472
PTE6	*Phymatodes testaceus*	Cerambycinae	OM585473
XCO1	*Xylotrechus colonus*	Cerambycinae	OM585502
AAE1	*Acanthocinus aedilis*	Lamiinae	OM585420
AAE2	*Acanthocinus aedilis*	Lamiinae	OM585421
AAE3	*Acanthocinus aedilis*	Lamiinae	OM585422
AAE4	*Acanthocinus aedilis*	Lamiinae	OM585423
AAE5	*Acanthocinus aedilis*	Lamiinae	OM585424
AVI1	*Agapanthia villosoviridescens*	Lamiinae	OM585431
AVI2	*Agapanthia villosoviridescens*	Lamiinae	OM585432
AVI3	*Agapanthia villosoviridescens*	Lamiinae	OM585433
AVI4	*Agapanthia villosoviridescens*	Lamiinae	OM585434
AVI5	*Agapanthia villosoviridescens*	Lamiinae	OM585435
EAD1	*Exocentrus adspersus*	Lamiinae	OM585446
EAD2	*Exocentrus adspersus*	Lamiinae	OM585440
EAD3	*Exocentrus adspersus*	Lamiinae	OM585441
EAD4	*Exocentrus adspersus*	Lamiinae	OM585442
EAD5	*Exocentrus adspersus*	Lamiinae	OM585443
EAD6	*Exocentrus adspersus*	Lamiinae	OM585444
EAD7	*Exocentrus adspersus*	Lamiinae	OM585445
EAD8	*Exocentrus adspersus*	Lamiinae	OM585439
MNE1	*Mesosa nebulosa*	Lamiinae	OM585452
MNE2	*Mesosa nebulosa*	Lamiinae	OM585453
MNE3	*Mesosa nebulosa*	Lamiinae	OM585454
MNE4	*Mesosa nebulosa*	Lamiinae	OM585455
SSC1	*Saperda scalaris*	Lamiinae	OM585495
SSC2	*Saperda scalaris*	Lamiinae	OM585496
SSC3	*Saperda scalaris*	Lamiinae	OM585497
SSC4	*Saperda scalaris*	Lamiinae	OM585498
SSC5	*Saperda scalaris*	Lamiinae	OM585499
SSC6	*Saperda scalaris*	Lamiinae	OM585500
SSC8	*Saperda scalaris*	Lamiinae	OM585501
CVI1	*Carilia virginea*	Lepturinae	OM585436
CVI2	*Carilia virginea*	Lepturinae	OM585437
CVI3	*Carilia virginea*	Lepturinae	OM585438
LAU1	*Leptura aurulenta*	Lepturinae	OM585447
LAU2	*Leptura aurulenta*	Lepturinae	OM585448
LAU3	*Leptura aurulenta*	Lepturinae	OM585449
LAU4	*Leptura aurulenta*	Lepturinae	OM585450
LAU5	*Leptura aurulenta*	Lepturinae	OM585451
OCU1	*Oxymirus cursor*	Lepturinae	OM585462
OCU2	*Oxymirus cursor*	Lepturinae	OM585463
OCU3	*Oxymirus cursor*	Lepturinae	OM585464
OCU4	*Oxymirus cursor*	Lepturinae	OM585465
OCU5	*Oxymirus cursor*	Lepturinae	OM585466
RBI1	*Rhagium bifasciatum*	Lepturinae	OM585474
RBI2	*Rhagium bifasciatum*	Lepturinae	OM585475
RBI3	*Rhagium bifasciatum*	Lepturinae	OM585476

RBI4	*Rhagium bifasciatum*	Lepturinae	OM585477
RBI5	*Rhagium bifasciatum*	Lepturinae	OM585478
RBIC1	*Rhamnusium bicolor*	Lepturinae	OM585479
RBIC2	*Rhamnusium bicolor*	Lepturinae	OM585480
RBIC3	*Rhamnusium bicolor*	Lepturinae	OM585481
RBIC4	*Rhamnusium bicolor*	Lepturinae	OM585482
RBIC5	*Rhamnusium bicolor*	Lepturinae	OM585483
RBIC6	*Rhamnusium bicolor*	Lepturinae	OM585484
RBIC9	*Rhamnusium bicolor*	Lepturinae	OM585485
NMA1	*Necydalis major*	Necydalinae	OM585456
NMA2	*Necydalis major*	Necydalinae	OM585457
NMA3	*Necydalis major*	Necydalinae	OM585458
NMA4	*Necydalis major*	Necydalinae	OM585459
NMA5	*Necydalis major*	Necydalinae	OM585460
NMA6	*Necydalis major*	Necydalinae	OM585461
ABA1	*Anisarthron barbipes*	Spondylidinae	OM585425
ABA2	*Anisarthron barbipes*	Spondylidinae	OM585426
ABA3	*Anisarthron barbipes*	Spondylidinae	OM585427
ABA4	*Anisarthron barbipes*	Spondylidinae	OM585428
ABA5	*Anisarthron barbipes*	Spondylidinae	OM585429
ABA6	*Anisarthron barbipes*	Spondylidinae	OM585430
SBU1	*Spondylis buprestoides*	Spondylidinae	OM585486
SBU2	*Spondylis buprestoides*	Spondylidinae	OM585487
SBU3	*Spondylis buprestoides*	Spondylidinae	OM585488
SBU4	*Spondylis buprestoides*	Spondylidinae	OM585489
SPI1	*Saphanus piceus*	Spondylidinae	OM585490
SPI2	*Saphanus piceus*	Spondylidinae	OM585491
SPI3	*Saphanus piceus*	Spondylidinae	OM585492
SPI4	*Saphanus piceus*	Spondylidinae	OM585493
SPI5	*Saphanus piceus*	Spondylidinae	OM585494
TFU1	*Tetropium fuscum*	Spondylidinae	OM810302
TFU2	*Tetropium fuscum*	Spondylidinae	OM810303
TFU3	*Tetropium fuscum*	Spondylidinae	OM810304
TFU4	*Tetropium fuscum*	Spondylidinae	OM810305
TFU5	*Tetropium fuscum*	Spondylidinae	OM810306
TFU6	*Tetropium fuscum*	Spondylidinae	OM810307
TFU7	*Tetropium fuscum*	Spondylidinae	OM810308
TFU8	*Tetropium fuscum*	Spondylidinae	OM810309
Reconstructed ancestral GH5_2			
ANC_GH5_2	Cerambycidae		OM810301

Table S2. List of GH5_2 sequences collected from Genbank.

Gene name	*Species*	Subfamily	Family	Genbank accession
ACH	*Anoplophora chinensis*	Lamiinae	Cerambycidae	AFN89566.1
AGE	*Apriona germari*	Lamiinae	Cerambycidae	AAX18655.1
AGL1	*Anoplophora glabripennis*	Lamiinae	Cerambycidae	XP_018565006.1
AGL2	*Anoplophora glabripennis*	Lamiinae	Cerambycidae	XP_018565008.1
AGL3	*Anoplophora glabripennis*	Lamiinae	Cerambycidae	XP_018564984.1
AGL4	*Anoplophora glabripennis*	Lamiinae	Cerambycidae	XP_018564982.1
AGL5	*Anoplophora glabripennis*	Lamiinae	Cerambycidae	XP_018562606.1
AGL6	*Anoplophora glabripennis*	Lamiinae	Cerambycidae	XP_018574045.1
AJA1	*Apriona japonica*	Lamiinae	Cerambycidae	AHI15746.1
AJA2	*Apriona japonica*	Lamiinae	Cerambycidae	AHI15747.1
AJA3	*Apriona japonica*	Lamiinae	Cerambycidae	AHI15748.1
AJA4	*Apriona japonica*	Lamiinae	Cerambycidae	AHI15749.1
AJA5	*Apriona japonica*	Lamiinae	Cerambycidae	AHI15750.1
AJA6	*Apriona japonica*	Lamiinae	Cerambycidae	AHI15751.1
MMYA	*Mesosa myops*	Lamiinae	Cerambycidae	*AMA76414.1*
MMYB	*Mesosa myops*	Lamiinae	Cerambycidae	AMA76415.1
MMYC	*Mesosa myops*	Lamiinae	Cerambycidae	AMA76416.1
MMYD	*Mesosa myops*	Lamiinae	Cerambycidae	AMA76418.1
MMYE	*Mesosa myops*	Lamiinae	Cerambycidae	AMA76417.1
MMYF	*Mesosa myops*	Lamiinae	Cerambycidae	AMA76419.1
OAL	*Oncideres albomarginata chamela*	Lamiinae	Cerambycidae	ADI24131.1
PHI	*Psacothea hilaris*	Lamiinae	Cerambycidae	BAB86867.1
Outgroup				
AAD45868	*Meloidogyne incognita*		Nematoda	AAD45868.1
AAK21881	*Meloidogyne incognita*		Nematoda	AAK21881.1
BAA31712	*Bacillus sp. 5H*		Bacteria	BAA31712.1

Table S3. Results of the branch site model selection and BEB analyses.

	lnL alt model	lnL null model	delta L	p-value chi-square	p-value (FDR adjusted)	
branch 1	-55839.038	-55856.15506	34.23413	4.8864E-09	2.44E-08	19 L 0.958* 50 F 0.995 ** 89 S 0.954* 106 K 0.998** 133 D 0.962* 138 E 0.952* 172 H 0.955* 214 S 0.954* 226 A 0.974* 227 E 0.984* 230 K 0.998** 278 N 0.984* 323 F 0.998** 373 K 0.991**
branch 2	-55854.98657	-55863.26441	16.55569	4.7242E-05	7.79E-05	69 Q 0.971*
branch 3	-55854.93527	-55863.07695	16.28337	5.454E-05	7.79E-05	
branch M	-55853.05763	-55860.11715	14.11905	0.0001716	2.14E-04	244 W 0.996** 311 G 0.967*
branch M1	-55859.59279	-55859.6947	0.203824	0.65165177	6.52E-01	
branch M2	-55862.26284	-55864.26771	4.009726	0.0452385	5.03E-02	91 G 0.972* 334 W 0.973*
branch C	-55852.82598	-55862.75256	19.853154	8.3625E-06	1.67E-05	343 A 0.995** 353 G 0.996**
branch Xyl	-55837.69894	-55851.89263	28.38740	9.931E-08	2.48E-07	244 W 0.980* 245 S 0.993** 273 T 0.950*
branch R9	-55852.47082	-55868.54361	32.14558	1.4304E-08	4.77E-08	50 F 0.991** 73 S 0.960* 214 S 0.963*
branch gt	-55820.03293	-55840.6665	41.26715	1.3278E-10	1.33E-09	19 L 0.954* 71 V 0.953* 172 H 0.975* 174 A 0.999* 180 Q 0.959* 181 A 0.963* 211 V 0.999** 248 V 0.999** 253 A 0.994** 254 N 0.950* 275 H 1.000** 276 K 0.954* 277 Q 0.999** 278 N 0.999** 292 K 0.982* 299 T 0.960* 305 D 1.000** 316 S 0.961* 319 E 0.960* 326 S 0.989* 328 K 0.999**

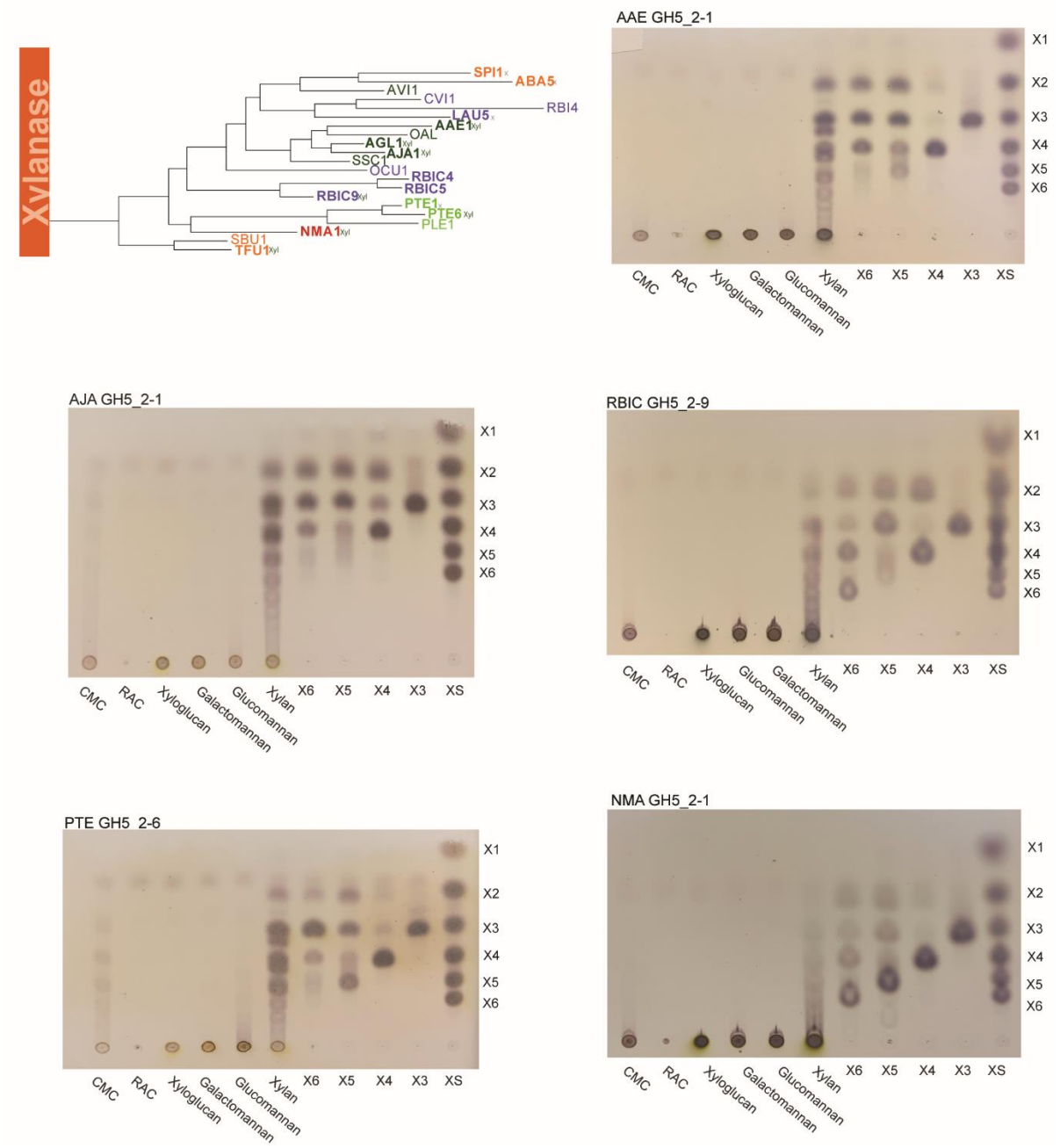

Figure S1. Thin-layer chromatography of cerambycid derived GH5_2s belonging to clade I ("xylanase clade"). GH5_2 proteins were expressed in Sf9 cells, and assayed against a range of polysaccharides and xylooligosaccharides. Products were then developed on thin-layer chromatography (TLC) plates. For more details, please refer to Fig. 1. The polysaccharide tested were carboxymethyl cellulose (CMC), regenerated amorphous cellulose (RAC), xyloglucan, galactomannan, Glucomannan, xylan and xylooligosaccharides (xylotriose (X3) to xylohexaose (X6)). Xylose (X1) to xylohexaose (X6) were used as standards.

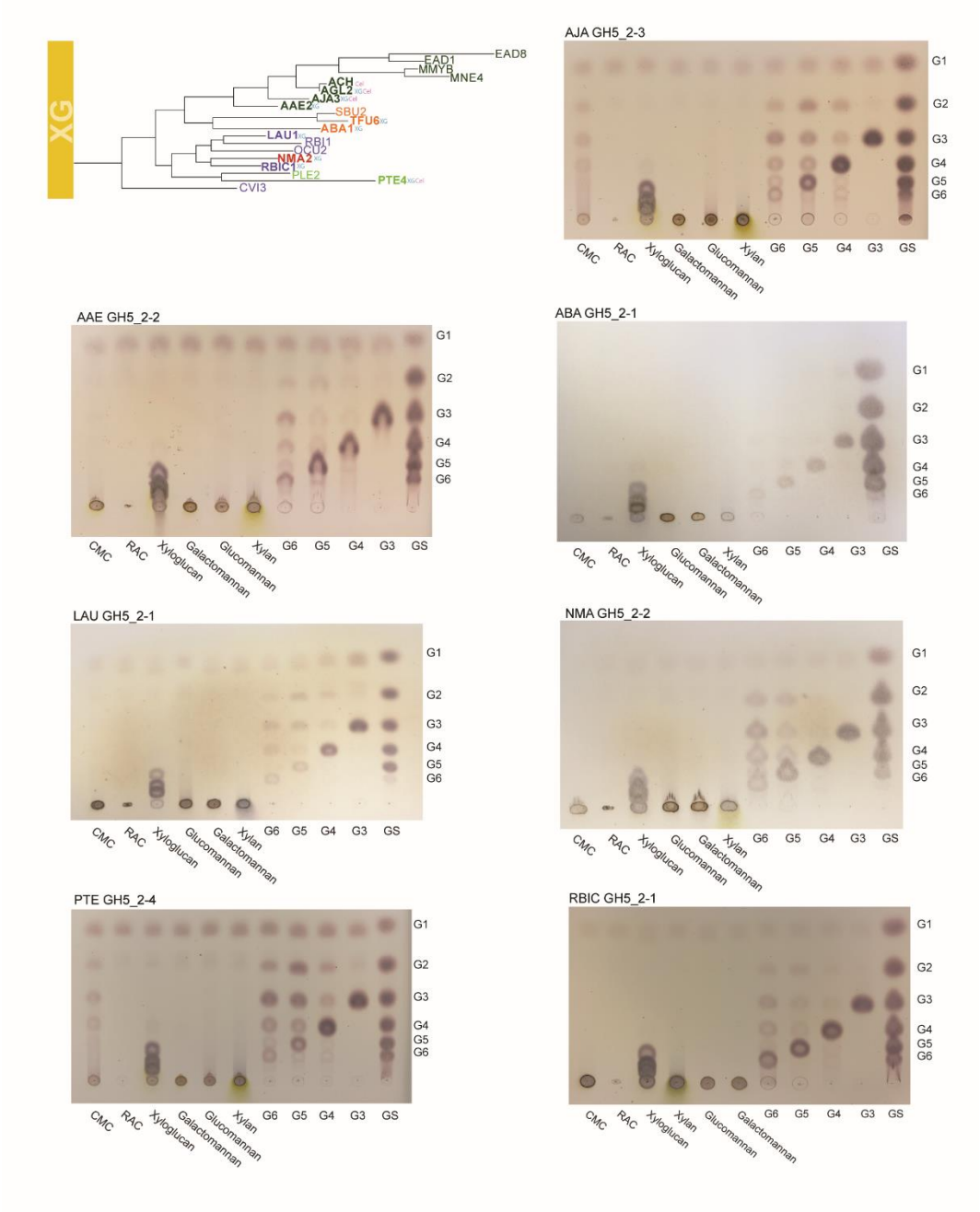

Figure S2. Thin-layer chromatography of cerambycid derived GH5_2s belonging to clade II ("xyloglucanase clade"). GH5_2 proteins were expressed in Sf9 cells, and assayed against a range of polysaccharides and cellooligosaccharides. Products were then developed on thin-layer chromatography (TLC) plates. For more details, please refer to Fig. 1. The polysaccharide tested were carboxymethyl cellulose (CMC), regenerated amorphous cellulose (RAC), xyloglucan, galactomannan, Glucomannan, xylan and cellooligosaccharides (cellotriose (G3) to cellohexaose (G6)). Glucose (G1) to cellohexaose (G6) were used as standards.

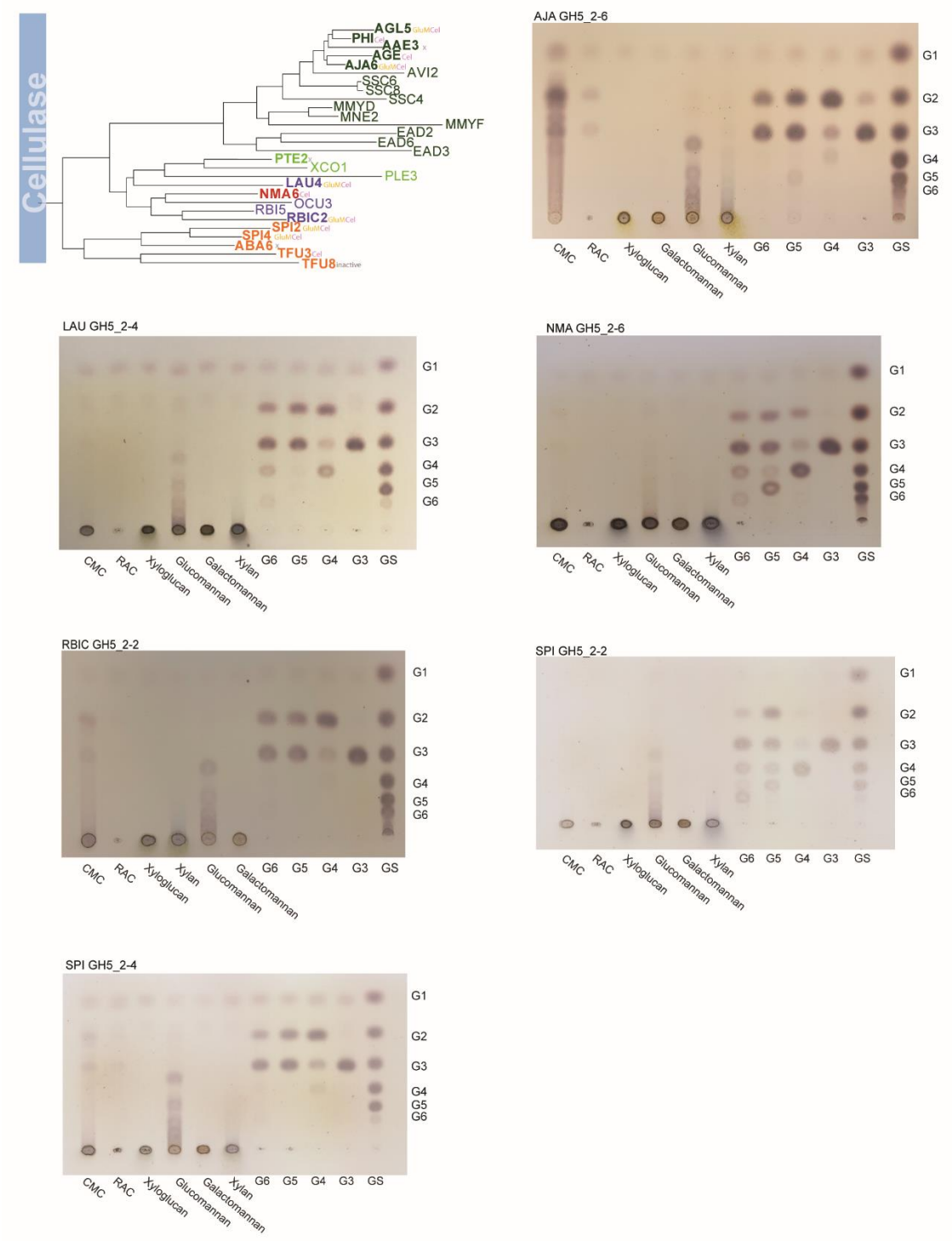

Figure S3. Thin-layer chromatography of cerambycid derived GH5_2s belonging to clade III ("cellulase clade"). GH5_2 proteins were expressed in Sf9 cells, and assayed against a range of polysaccharides and cellooligosaccharides. Products were then developed on thin-layer chromatography (TLC) plates. For more details, please refer to Fig. 1. The polysaccharide tested were carboxymethyl cellulose (CMC), regenerated amorphous cellulose (RAC), xyloglucan, galactomannan, Glucomannan, xylan and cellooligosaccharides (cellotriose (G3) to cellohexaose (G6)). Glucose (G1) to cellohexaose (G6) were used as standards.

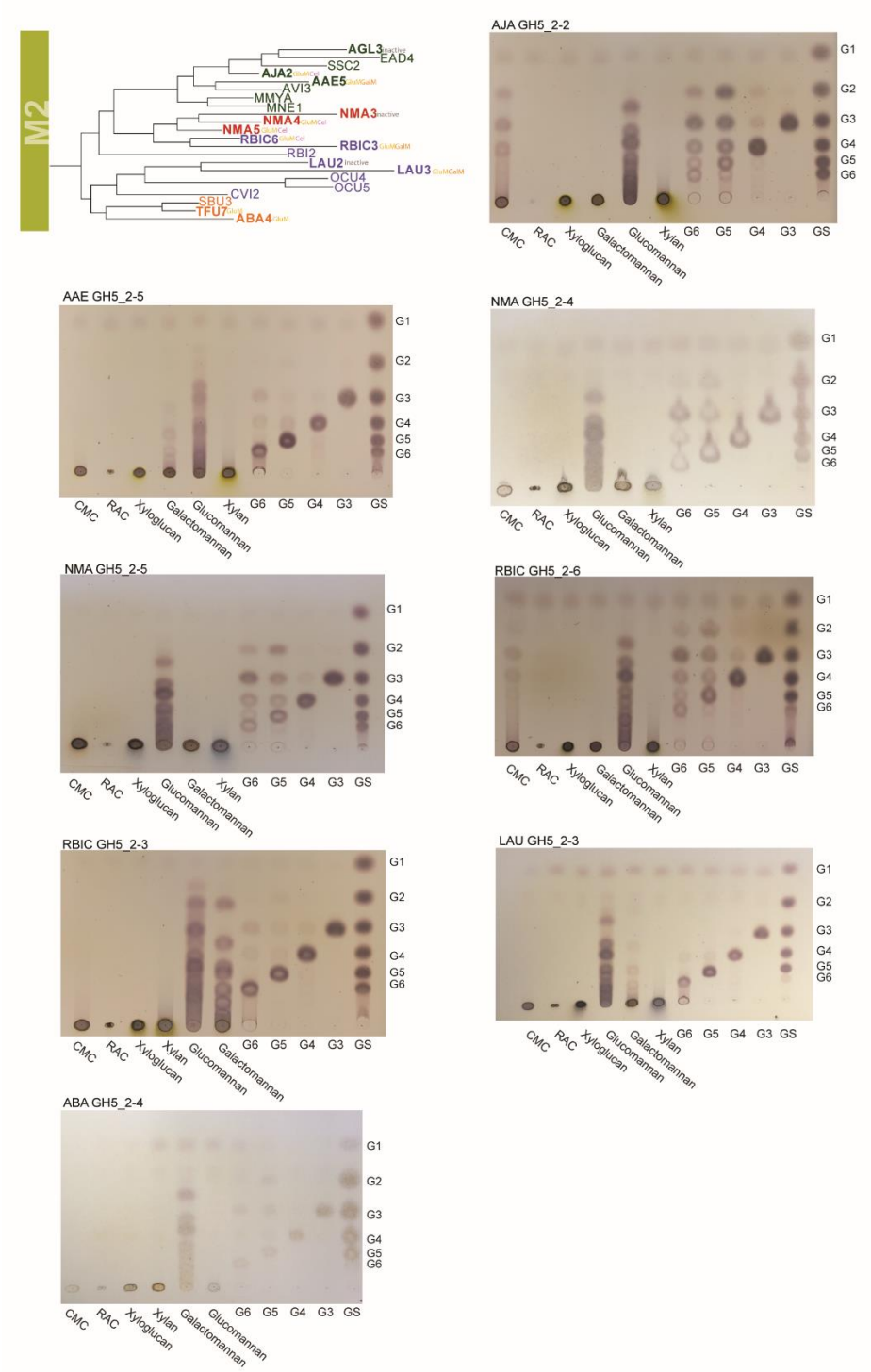

Figure S4. Thin-layer chromatography of cerambycid derived GH5_2s belonging to clade IV ("mannanase 2 clade"). GH5_2 proteins were expressed in Sf9 cells, and assayed against a range of polysaccharides and cellooligosaccharides. Products were then developed on thin-layer chromatography (TLC) plates. For more details, please refer to Fig. 1. The polysaccharide tested were carboxymethyl cellulose (CMC), regenerated amorphous cellulose (RAC), xyloglucan, galactomannan, Glucomannan, xylan and cellooligosaccharides (cellotriose (G3) to cellohexaose (G6)). Glucose (G1) to cellohexaose (G6) were used as standards.

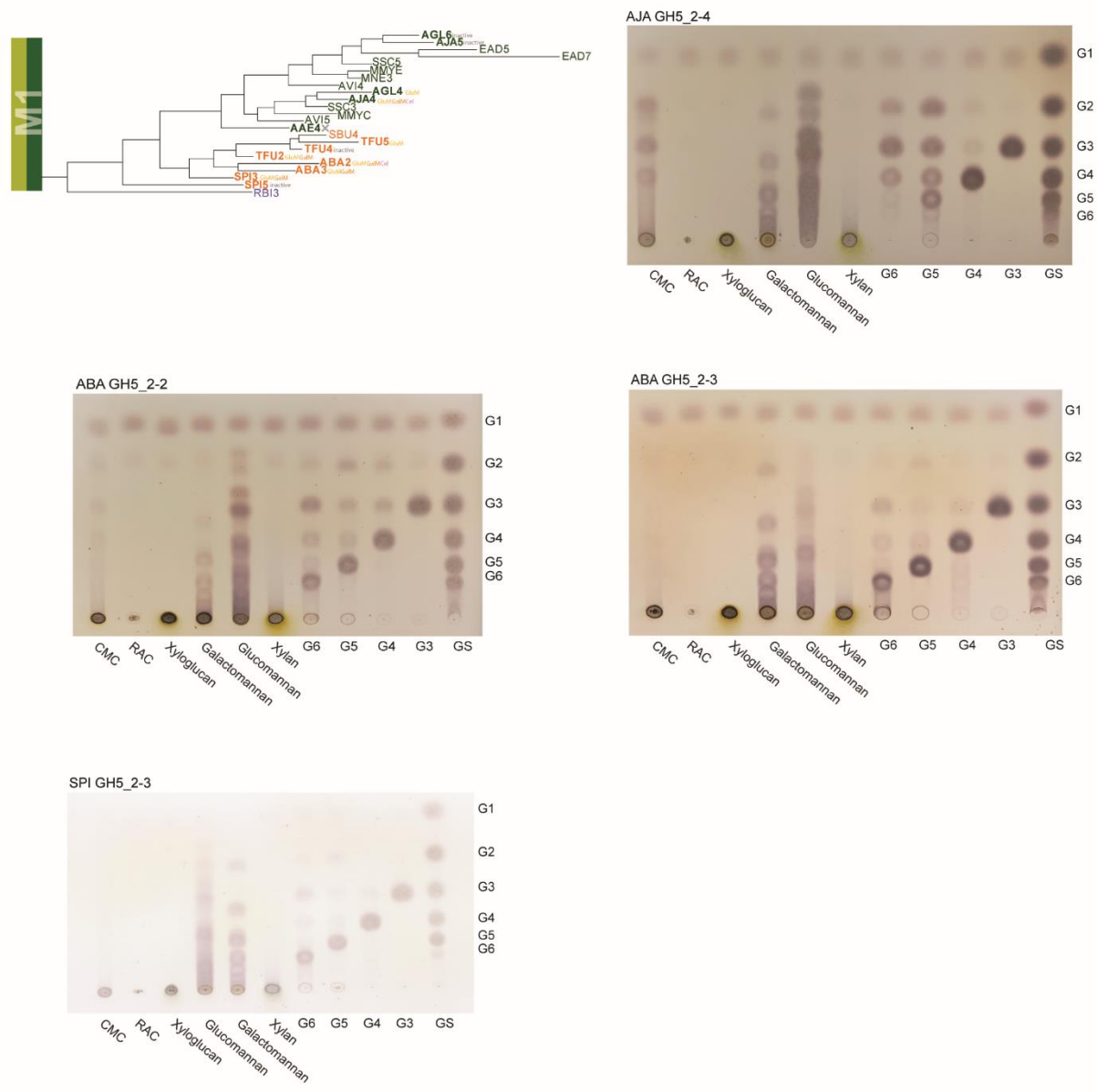

Figure S5. Thin-layer chromatography of cerambycid derived GH5_2s belonging to clade IV ("mannanase 1 clade"). GH5_2 proteins were expressed in Sf9 cells, and assayed against a range of polysaccharides and cellooligosaccharides. Products were then developed on thin-layer chromatography (TLC) plates. For more details, please refer to Fig. 1. The polysaccharide tested were carboxymethyl cellulose (CMC), regenerated amorphous cellulose (RAC), xyloglucan, galactomannan, Glucomannan, xylan and cellooligosaccharides (cellotriose (G3) to cellohexaose (G6)). Glucose (G1) to cellohexaose (G6) were used as standards.

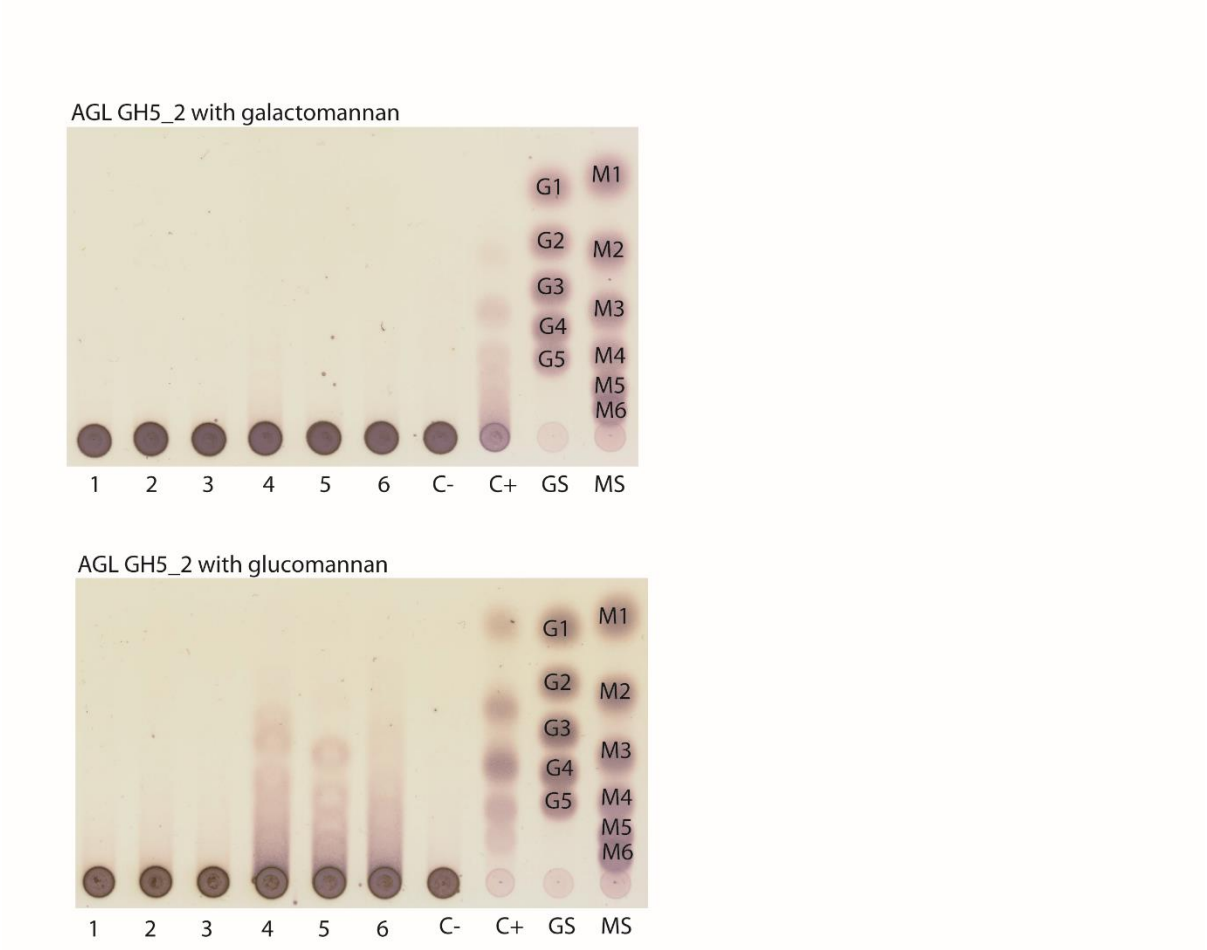

Figure S6. Thin-layer chromatography of *Anoplophora glabripennis* derived GH5_2s assayed against two types of mannan polysaccharides. *Anoplophora glabripennis* GH5_2 enzymes were not tested against mannans in the genome paper (McKenna, et al. 2016). GH5_2-1 to -6 were expressed in Sf9 cells and tested against galactomannan and glucomannan as substrates. Activity on glucomannan was detected for GH5_2-4, -5 and -6. No enzyme was active against galactomannan. Glucose (G1) to cellohexaose (G6), as well as mannose (M1) to mannohexaose (M6) were used as standards.

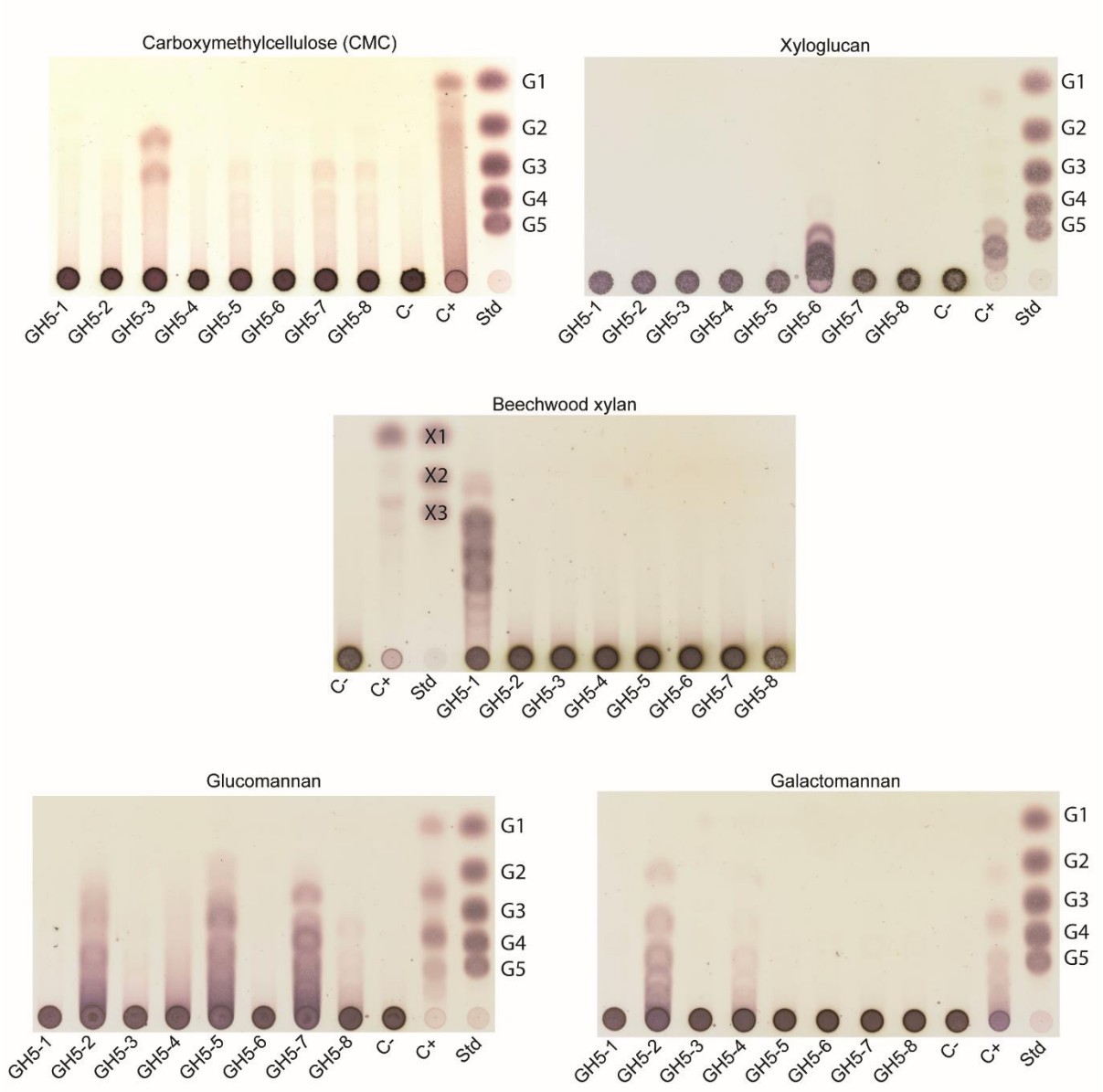

Figure S7. Thin-layer chromatography of *Tetropium fuscum* derived GH5_2s assayed against PCW polysaccharides. GH5_2 proteins were expressed in Sf9 cells, and assayed against a range of polysaccharides. Products were then developed on thin-layer chromatography (TLC) plates. For more details, please refer to Fig. 1. The polysaccharides tested were carboxymethyl cellulose (CMC), xyloglucan, galactomannan, Glucomannan, xylan. Glucose (G1) to cellopentaose (G5), as well as xylose (X1) to xylotriose (X3) were used as standards.

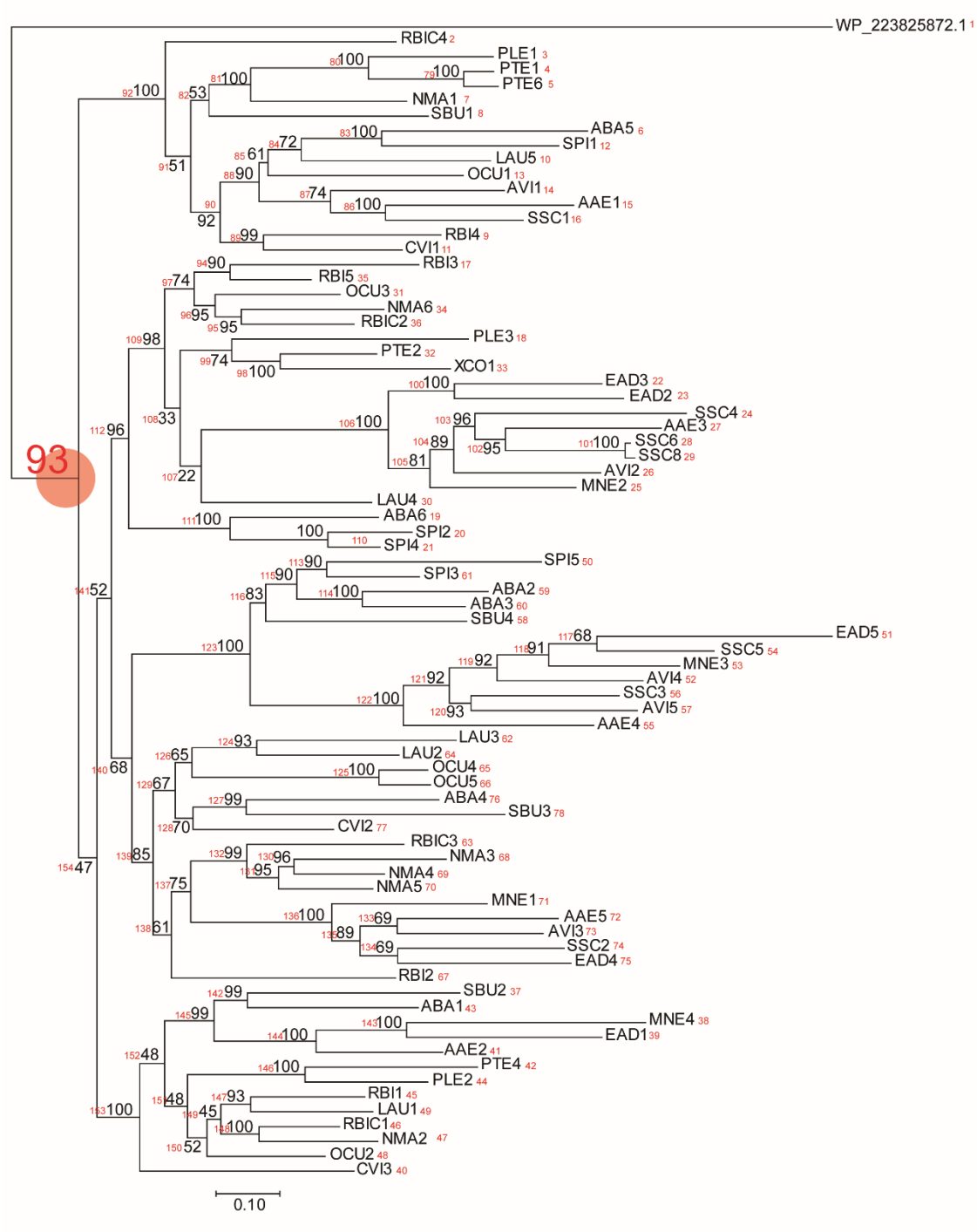

Figure S8. Phylogenetic tree generated for reconstruction of ancestral sequence of GH5_2. We performed IQ-TREE to infer maximum likelihood tree to estimate the putative sequences of ancestral GH5_2 in cerambycid beetles. 77 genes encoding cerambycid derived GH5_2 and one bacterial GH5_2 (WP_223825872.1) were aligned using MAFFT (v7.471). Ultra bootstrap value were labelled with black on the node, and node number from ancestral reconstruction analysis in MEGA are labelled with Red on the node. The red circle indicates the node for reconstruction of ancestral GH5_2 sequence.

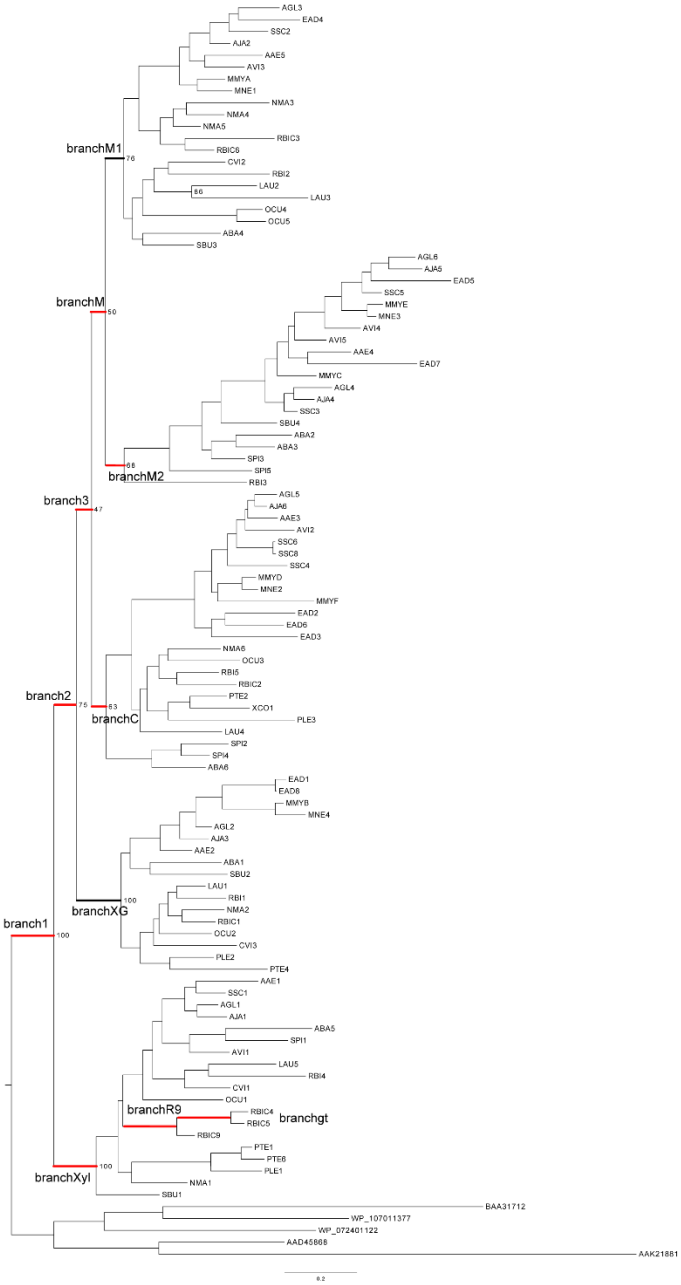

Figure S9. Phylogenetic tree generated for the branch site model selection. We tested 11 individual branches for signal of positive selection: branch 1; branch2; branch3; branch Xyl; branch XG; branch C; branch M; branch M1; branch M2; branch R9; and branch gt. Branchesfor which a signal of positive selection was detected from the branch site model test and BEB analyses, are labelled in red. Further information can be found in the Table S3.

Literature cited in the legends

McKenna DD, Scully ED, Pauchet Y, Hoover K, Kirsch R, Geib SM, Mitchell RF, Waterhouse RM, Ahn SJ, Arsala D, et al. 2016. Genome of the Asian longhorned beetle (*Anoplophora glabripennis*), a globally significant invasive species, reveals key functional and evolutionary innovations at the beetle-plant interface. Genome Biol 17:227.