PNAS
www.pnas.org

Supplementary Information for

A conserved mechanism affecting hydride shifting and deprotonation in the synthesis of hopane triterpenes as compositions of wax in oat

Miaomiao Liang, Fan Zhang, Jiaxin Xu, Xiaoning Wang, Ruibo Wu, Zheyong Xue

Corresponding author: Zheyong Xue or Ruibo Wu.
Email: xuezhy@126.com or wurb3@mail.sysu.edu.cn

This PDF file includes:

Supplementary text
Figures S1 to S11
Tables S1 to S2
Legends for Datasets S1 to S5
SI References

Other supplementary materials for this manuscript include the following:

Datasets S1 to S5

Supplementary Information Text

Supplementary experimental methods

Triterpene extraction and gas chromatography-mass spectrometry (GC-MS) analysis

Ten milligrams of dry materials were weighed and them saponified in 1 mL of saponification reagent (10% [w/v] KOH in 90% [v/v] ethanol) containing 10 μg/mL coprostanol (internal standard) (Sigma-Aldrich C7578, Merck, Darmstadt, Germany) with constant shaking at 75°C for 1 h. Then removing the cover and heating the metabolite-extraction samples for an additional hour to allow the ethanol to evaporate. Ethyl acetate (0.5 mL) was added to the resulting product and vortexed. Then adding 0.5 mL water and vortexed again. Solutions were then centrifuged to aid partitioning of ethyl acetate from water. Aliquots of the ethyl acetate solutions (50 μL) were dried with N_2 and derivatized using 50 μL of trimethylsilyl imidazole/pyridine reagent (Sigma-Aldrich 92718, China) at 70°C for 30 min. Reaction solutions were analysed using an GC-QQQ-MS instrument (Agilent technologies, California, USA) set 70 eV with a Zebron ZB-5 HT capillary column (dimensions = 0.25 mm, lengths = 30 m) (Torrance, CA, USA). Reactions were placed in an incubator at a temperature was initially set at 170°C, raised from 170 to 290°C (at a rate of 6°C min^{-1}), maintained for 4 min, and then elevated from 290 to 340°C (at a rate of 10°C min^{-1}). The quantification of the relative amounts of compounds was carried out by comparisons with the amount of the internal standard.

Triterpene purification and structural identification

We carried out batch infiltration of ~200 *N. benthamiana* plants with each of two *A. tumefaciens* mixtures, one carrying pEAQ-HT-DEST1*tHMGR,* the other harboring pEAQ-HT-DEST1*OSCs*. The vacuum dry materials were weighted and saponified in a corresponding volume (100 mL/g) of saponification reagent and incubated at 75°C for 1 h. The products were extracted three times with an equal volume of hexane. The hexane extract was loaded onto a silica gel column (45 g; 40-63μm; CV 60mL; Biotage, Uppsala, Sweden), and eluted with a solution consisting of one of the following ratio of hexane (A): ethyl acetate (B) Elution gradient: 15 volume, 0%~50% B; 3 volume, 50%~100% B; 2 volume, 100% B; the flow rate was 60 mL/min. Fractions were collected in 50 mL tubes and analysed by thin layer chromatography (TLC) and GC-MS.

Fractions containing hopenol B were further purified by reversed-phase high performance liquid chromatography (HPLC) with a semi-preparative column (Thermo Hypersil GOLD C18 column, 250 mm × 10 mm, 5 μm). The mobile phases were water (A) and methanol (B). The

separation was conducted using the elution gradient from 60%~100%B at 0-8 min, maintaining at 100% B for 14 min, and backing to 60% B at 22-25 min. The flow rate was 5 mL/min, the column temperature is 30°C, the injection volume is 80 μL, and the detection wavelength was 210 nm. Hopenol B (3.8 mg, t_R=15.0 min) was purifed. Using the same methodology, hop-17(21)-en-3*β*-ol (2.4 mg, t_R=14.0 min) and isomotiol (3.4 mg, t_R=15.0 min) was purified. The purified compounds were dissolved with CDCl3 for NMR analysis including ^{1}H-NMR, ^{13}C-NMR, HMBC, HSQC, ^{1}H-^{1}H COSY, and ROESY. Chemical shifts were recorded in parts per million (ppm) and referenced to the residual solvent peak or to an internal tetramethylsilane standard.

Extraction of wax compositions of *A.strigosa* tissues

The panicle, sheath, whole root, leaf tissues were collected from two-month-old oats, and tissue samples were lyophilized. For wax surface extraction, the dry materials (1g) were immersed in chloroform (100 mL) for 1 min each. Then remove the plant materials by filtration immediately, the solvent was evaporated, leaving the wax behind as a solid residue. For whole tissue extraction, the dry materials were grinded in liquid nitrogen with a mortar and pestle. The powder (1g) was immersed in chloroform (100 mL) for 1 min. Then remove the plant materials by filtration immediately, the solvent was evaporated, leaving the whole tissue extraction behind as a solid residue. For unsaponification method, add 0.5 mL of chromatographic ethyl acetate to the solid residue and dissolve with ultrasound. Add 0.5 mL water and vortex again. Centriguge to aid partition. Aliquots of ethyl acetate solution (50 μL) were dried under N_2 and derivatized using 50 μL of trimethylsilyl imidazole/pyridine reagent (Sigma-Aldrich 92718, China) at 70°C for 30 min. For saponification method, solid residue was saponified in 1 mL of saponification reagent (10% (w/v) KOH in 90% (v/v) ethanol) with constant shaking at 75°C for 1 h. Then open the lid and heat for an additional hour to allow ethanol to evaporate. Ethyl acetate (0.5 mL) was added to the resulting product and vortex. Add 0.5 mL water and vortex again. Centriguge to aid partition. Aliquots of ethyl acetate solution (50 μL) were dried under N_2 and derivatized using 50 μL of trimethylsilyl imidazole/pyridine reagent (Sigma-Aldrich 92718, China) at 70°C for 30 min.

Mutagenesis experiments

Mutagenesis was performed using the QuikChange site-directed mutagenesis method (Stratagene, Agilents, California, USA). The expression plasmids for AsHS1, AsHS2, and AcHS1 were used as templates for PCR-based site-directed mutagenesis. The primers used for site-directed mutagenesis are listed in Supplementary Table 1 with the substitutions underlined. Reaction for PCR was composed of 0.1 mM of deoxynucleoside triphosphates (dNTPs), 40 ng of plasmid template, 1×Phusion HF buffer, 0.25 μM of each primer, 2 U of Phusion DNA polymerase (New England Biolabs, Ipswich, MA, USA) and double-distilled H2O to a final volume of 20 μL. The reactions were denatured at 98°C for 30 s, and then run

for 20 cycles of denaturation at 98°C for 10 s, annealing at 60°C for 30 s, and polymerization at 72°C for 6 min; a final extension was carried out at 72°C for 30 min. PCR products were incubated with DpnI at 37°C for 2 h to digest the parental supercoiled DNA. The mutations were confirmed by DNA sequencing. Functional characterization of the mutants was conducted as that for the AsHS1 and AsHS2 mentioned above.

Quantitative reverse transcription-polymerase chain reaction (qRT-PCR)

The *N.benthamiana* leaves were harvested after 24 h of infiltration and frozen in liquid nitrogen before RNA extraction using sing the TRI-REAGENT (Invitrogen, California, USA). First-strand cDNA was synthesized from 1 μg of each DNase-treated RNA using SuperScript™ III (Invitrogen, California, USA) in accordance with the manufacturer's instructions. Primers for the amplification of fragments of *AsHS1*, *AsHS2*, *AcHS1* and their mutants (*NtACTIN* as a reference gene) are listed in Table 1. qPCR was carried out on Roche LightCycler 96 (USA) using Talent qPCR PreMix (Tiangen Biotech, China). The relative expression levels for each gene were normalized to the ACTIN gene X69885, with three biological replicates per targeted gene.

Supplementary Computational Methods

System setup

The crystal structure of human OSCs (HsLAS) (PDB entry:1w6k/1w6j) was used for OSCs modeling by AlphaFold2(1). First for benchmark, we found that the AlphaFold2 predicted hLAS structure is highly consistent with crystal structure in secondary structure (0.16 Å RMSD and TM-score of 0.999), indicating the high reliability of predicted OSC models. Then several OSC models including AsHS1, AsHS2, AcHS1 and their corresponding mutants were built. According to the reaction pathways, the key intermediate hopyl C22 cation was used as our starting point for system setup and subsequent computational studies. The intermediate was first docked into the active site, then the diverse intermediate conformations will be further sampled and optimized by longtime scale classical MD simulations on enzyme-ligand complex in a solvent box, to obtain reliable initial structure for subsequent QM/MM simulations.

Classical MD Simulations

The Amber ff99SB force field(2) was employed for the protein and the TIP3P model was used for water molecules(3). The force field parameters of the ligand were generated from the general AMBER force field (GAFF)(4), and the partial atomic charge of substrates was defined by the restrained electrostatic potential (RESP)(5) charge from the HF/6-31G*

calculation with the Gaussian 09 package(6). The initial coordinates and topology files were generated by the *tleap* program in AMBER12(7). The MD simulations were carried out using the AMBER12 molecular simulation package, and the periodic boundary condition with cubic models were employed. The routine minimization, first by constraining all solute atoms then protein backbone and finally no constraint, were carried out to preliminarily relax the solvent and protein-ligand complex. After the optimization, each system was heated from 0 to 300 K gradually under the NVT ensemble for 100 ps, followed by another 100 ps NPT ensemble MD simulations at 300 K and the target pressure of 1.0 atm. Afterward, 50 ns NVT production MD simulations with a target temperature of 300 K were performed to produce trajectories. During the MD simulations, the SHAKE algorithm(8) was applied to constrain the high-frequency stretching vibration of all hydrogen-containing bonds, and a cutoff of 12 Å was set for both vdW (LJ-12 potential) and electrostatic interactions (PME strategy). Finally, snapshots of each system from the stable trajectories were chose to build the initial structures for the subsequent QM/MM simulations.

QM/MM MD Simulations

The periodic boundary condition was also considered in the following QM/MM MD simulations. The F257, C366 and the Y410/A410 together with hopyl C22 cation were included in the QM region of wildtype and mutant AsHS1/AsHS2 models (the charge of QM region is +1). The F260, L413 together with hopyl C22 cation were included in the QM region of AcHS1 model (the charge of QM region is also +1). All of these QM atoms were described with the M06-2X/6-31G(d) basis set which is widely used in studying cyclization reaction9 (9-12), and the model contains about 500 basis functions in total. The QM/MM boundary was treated by the improved pseudo-bond approach (13-15). The same Amber force field in the aforementioned classical MD simulations was used for the remaining atoms. The 12 Å cutoff was employed for both van der Waals interaction by 126 Lennard-Jones potential function and electrostatic interactions by dual-focal ai-QM/MM-PME approach (16). The QM/MM systems were minimized again for several iterations and more than 20 ps QM/MM MD simulations were performed. The resulting conformations of QM/MM MD were used to map out the minimum energy path (MEP) with the reaction coordinate driving (RCD) method (17). RCD method heavily depends on the existence of a good reaction coordinate, thus different RCs were considered for each reaction step and selected based on the obtained MEP. For every point with fixed reaction coordinate, the structure will be minimized with QM/MM microiterative. In order to get smooth enough energy profiles, at least three times of forward and backward reaction path scan were performed. All of these QM/MM calculations were performed with the interfaced QChem (18)-AMBER12 programs (16).

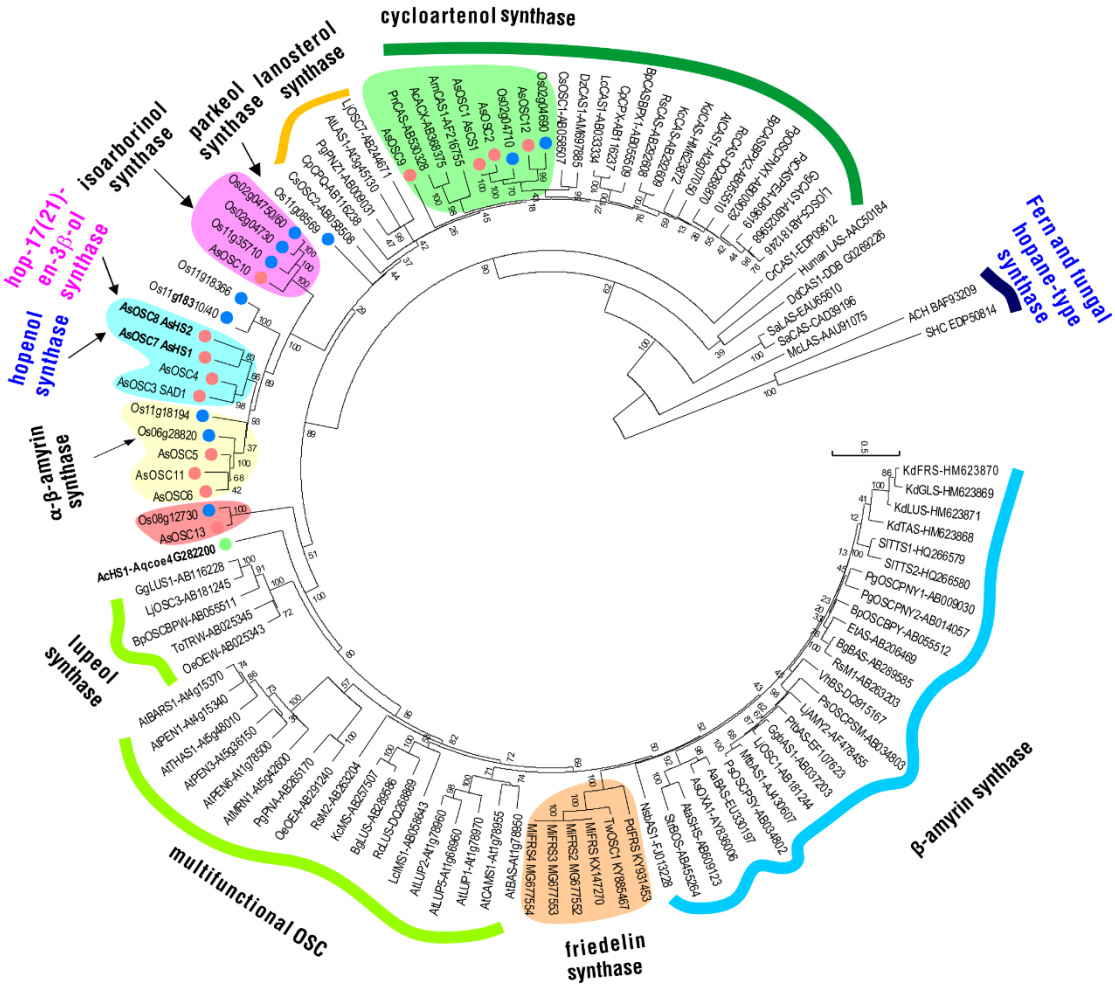

Fig. S1. Phylogenetic tree of oxidosqualene cyclases (OSCs) from *Avena strigosa* and other plant species. *Avena strigosa OSCs* were manually annotated from *A.strigosa* genome data (19). Functionally characterized plant OSCs and other OSCs from lower organisms used as outgroup sequences were derived from review paper (20). Amino acid sequences of OSCs were aligned using MUSCLE with default parameters as implemented in the program MEGA 7 (21). A Neighbor-joining tree was constructed with the evolutionary distances obtained by the JTT matrix-based method (22). Cycloartenol synthases are shown in green, lanosterol synthases are shown in yellow, dicot *β*-amyrin synthases are shown in blue, and lupeol and multifunctional OSCs are shown in light blue. OSCs from *A.strigosa* are marked by red dots, *Aquilegia coerulea* OSC is marked with green dot and OSCs from rice are marked by blue dots.

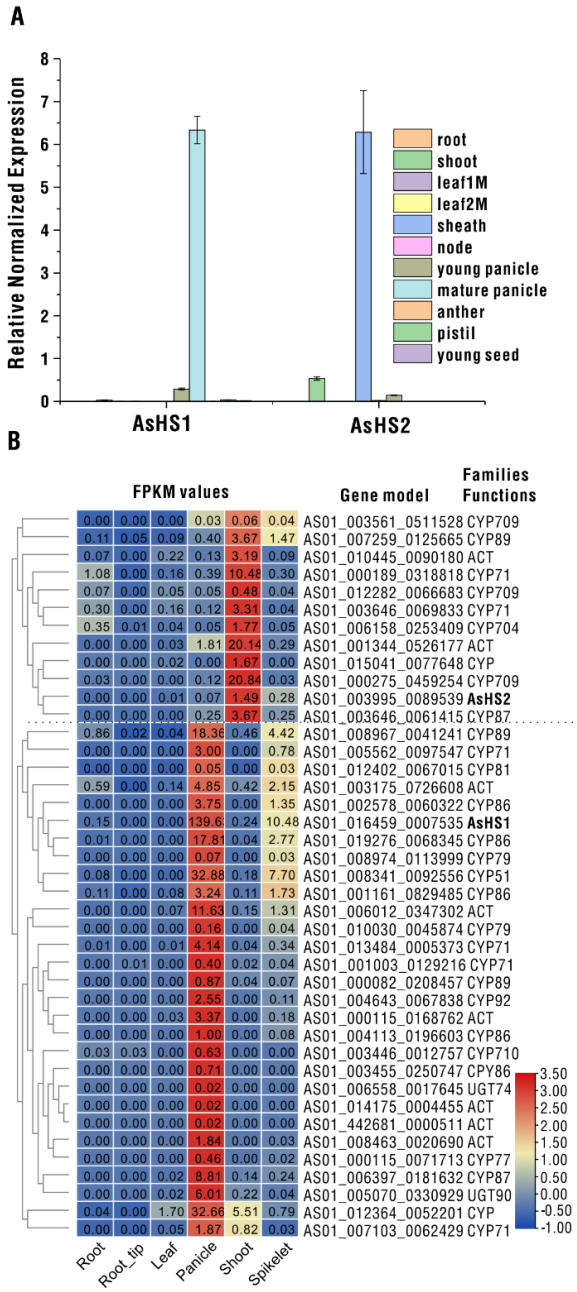

Fig. S2. Expression pattern and coexpression analysis of AsHS1 and AsHS2 in diverse tissues of oat. (A) qRT-PCR analysis to quantify mRNA levels in root, shoot, leaf1M (leaf from one month old plant), leaf2M (leaf from two months old plant), sheath, node, young panicle, mature panicle (heading stage), anther, pistil and immature seed. (**B**) Hot map shows co-expression analysis of AsHS1 and AsHS2 with cytochrome P450s, acyltransferases, and UDP-glucosyltransferases with high pearson correlation coefficient (>0.9) derived from RNA-seq data for six different tissues of *A. strigosa* accession S75. Total 414 P450s, 152 ACTs, 225 UGTs were annotated from transcriptome data. The correlation coefficients between each and AsHS1 or AsHS2 were calculated by Excel Correl function, and the genes with correlation coefficient greater than 0.9 (29 P450s, 8 ACTs, 2 UGTs) were selected (Dataset S5).

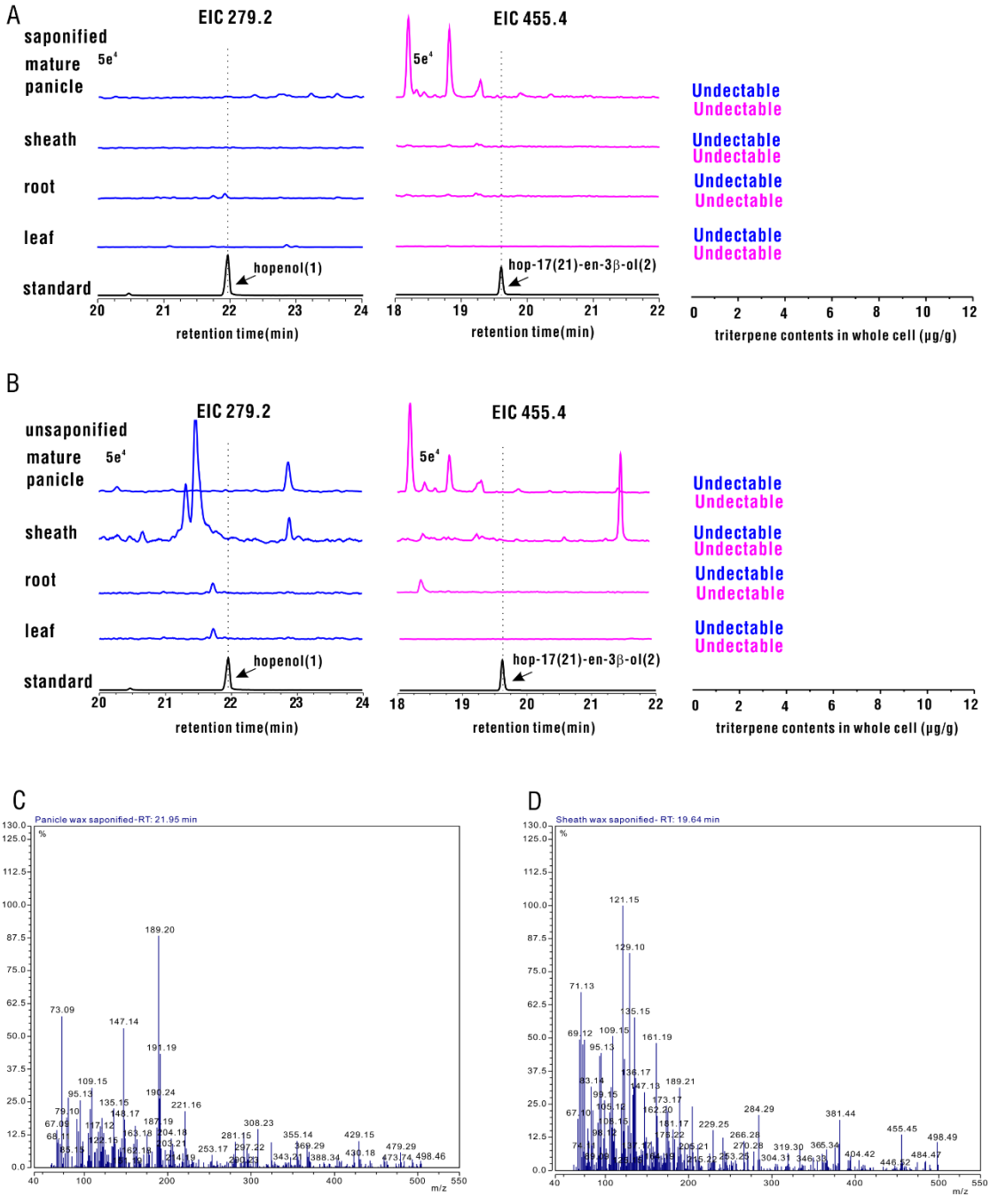

Fig. S3. Identification of targeted triterpene constituents in oat tissues. (A&B) Extracted ion chromatogram and absolute amounts of hopenol B and hop-17(21)-en-3*β*-ol were quantified by GC-MS analysis of extracts from saponified or unsaponified whole cell of four tissues using standard curve method. EIC279.2 and EIC455.4: extracted ion chromatogram at m/z 279.2 and m/z 455.4. All value from three biological replicates. (**C&D**) Electron impact (EI) mass spectral fragmentation of hopenol B (**C**) and hop-17(21)-en-3*β*-ol (**D**) extracted from surface wax of panicle and sheathes, respectively.

Fig. S4. Sequences alignments of AsHS1 and AsHS2 showing diversity of sites.

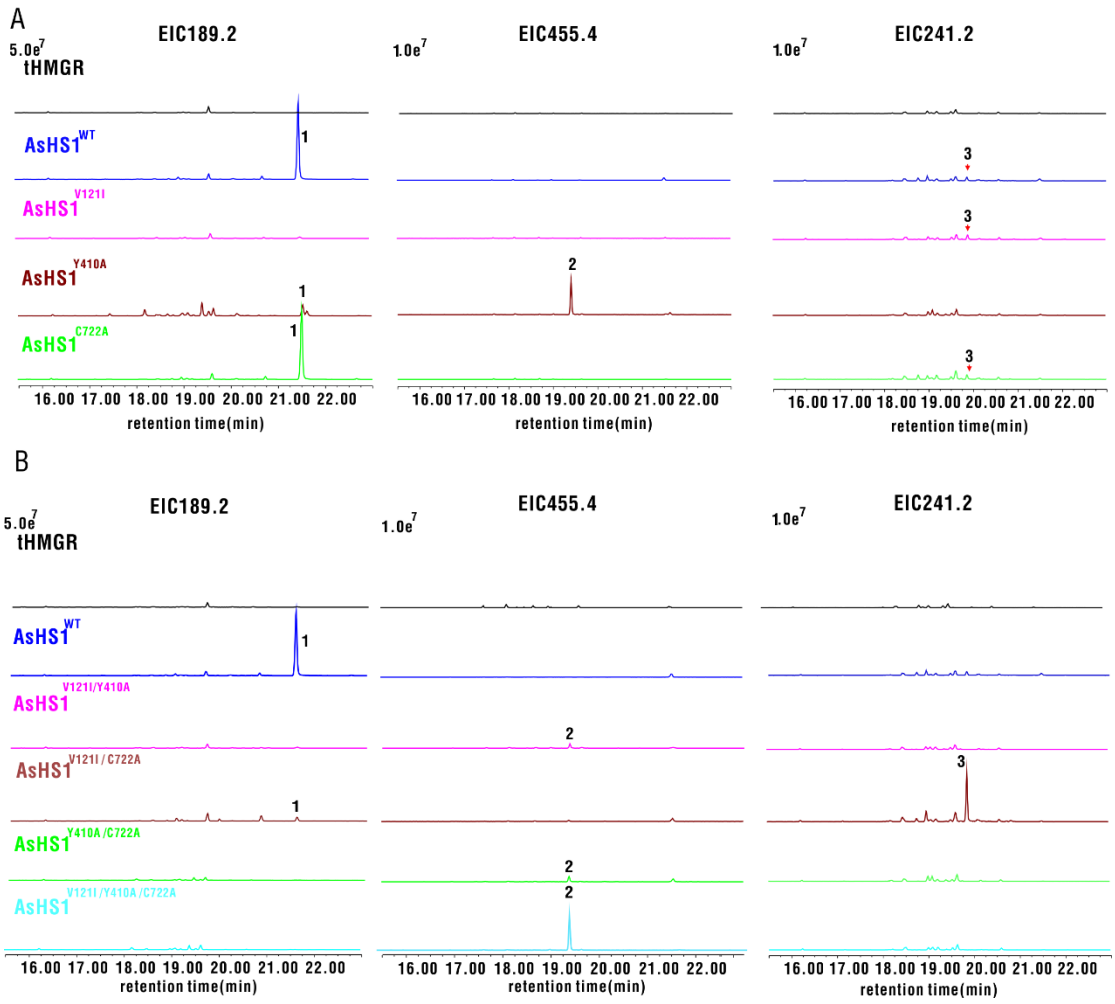

Fig. S5. Extracted ion chromatograms (EIC) of *N. benthamiana* leaf extracts expressing the wild-type and mutants of AsHS1. EICs obtained at m/z 189.2 (EIC189.2) and m/z 455.4 (EIC455.4) and m/z 241.2 (EIC241.2) are shown. tHMGR, a truncated 3-hydroxy-3-methyl glutaryl coenzyme A reductase from oat. 1: hopenol B, 2: hop-17(21)-en-3*β*-ol, 3: isomotiol.

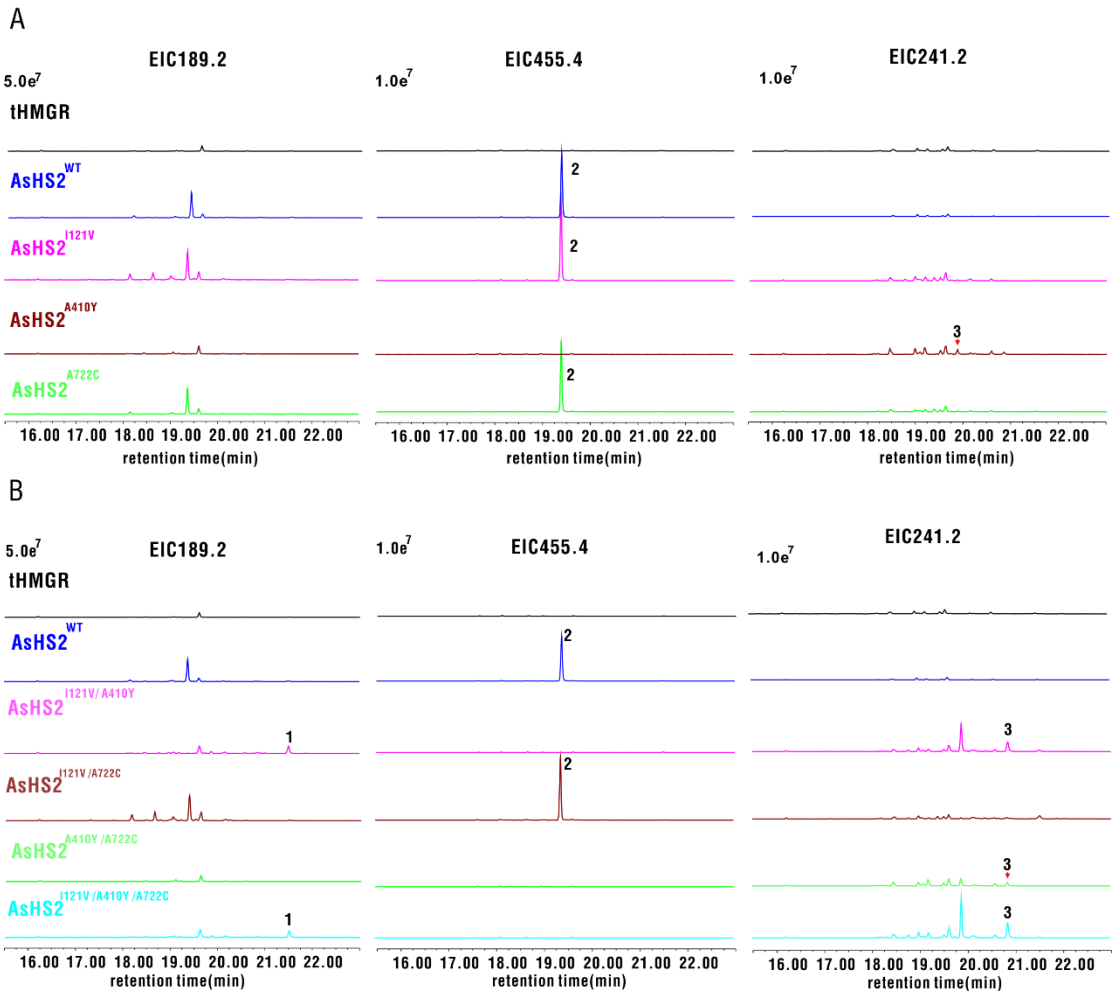

Fig. S6. Extracted ion chromatograms (EIC) of *N. benthamiana* leaf extracts expressing the wild-type and mutants of AsHS2. EICs obtained at m/z 189.2 (EIC189.2) and m/z 455.4 (EIC455.4) and m/z 241.2 (EIC241.2) are shown. tHMGR, a truncated 3-hydroxy-3-methyl glutaryl coenzyme A reductase from oat. 1: hopenol B, 2: hop-17(21)-en-3*β*-ol, 3: isomotiol.

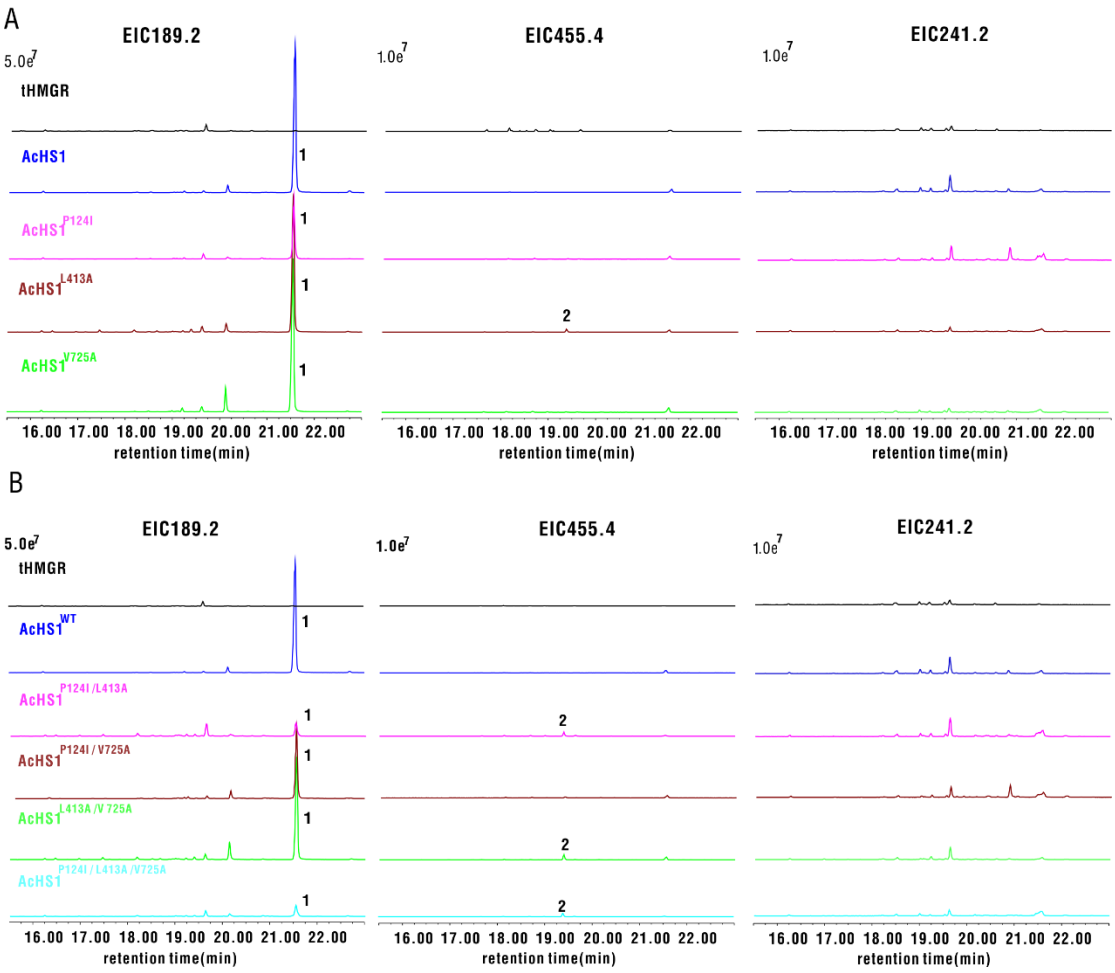

Fig. S7. Extracted ion chromatograms (EIC) of *N. benthamiana* leaf extracts expressing the wild-type and mutants of *Aquilegia* hopenol B synthase. EICs obtained at m/z 189.2 (EIC189.2) and m/z 455.4 (EIC455.4) and m/z 241.2 (EIC241.2) are shown. tHMGR, a truncated 3-hydroxy-3-methyl glutaryl coenzyme A reductase from oat. 1: hopenol B, 2: hop-17(21)-en-3*β*-ol, 3: isomotiol.

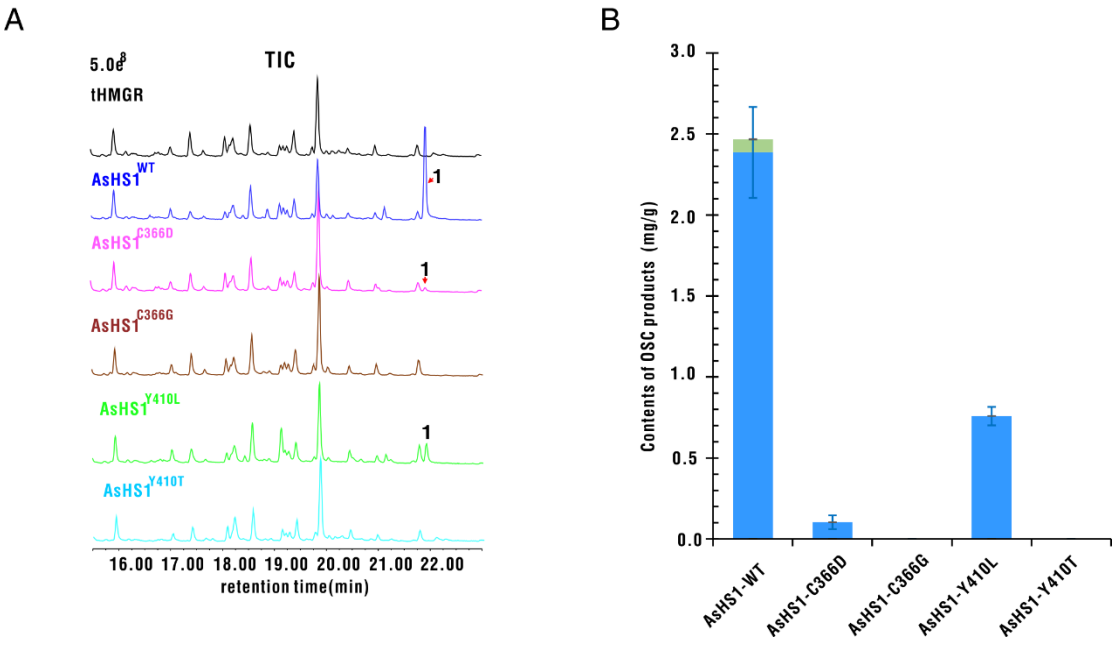

Fig. S8. Total ion chromatograms (TIC) (A) and the quantified triterpene contents (B) of *N. benthamiana* leaves extracts expressing the wild-type AsHS1 and AsHS1^{C366G}, AsHS1^{C366D}, AsHS1^{Y410L}, AsHS1^{Y410T}. The quantities of hopenol B (blue) and isomotiol (green) were determined by integrating peak areas from gas chromatography-mass spectrometry (GC-MS) analysis using coprostanol as an internal standard.The activities of wild-type and mutants are presented as means ± SE, n = 3 (contents of products mg/g). tHMGR, a truncated 3-hydroxy-3-methyl glutaryl coenzyme A reductase from oat (43). 1: hopenol B.

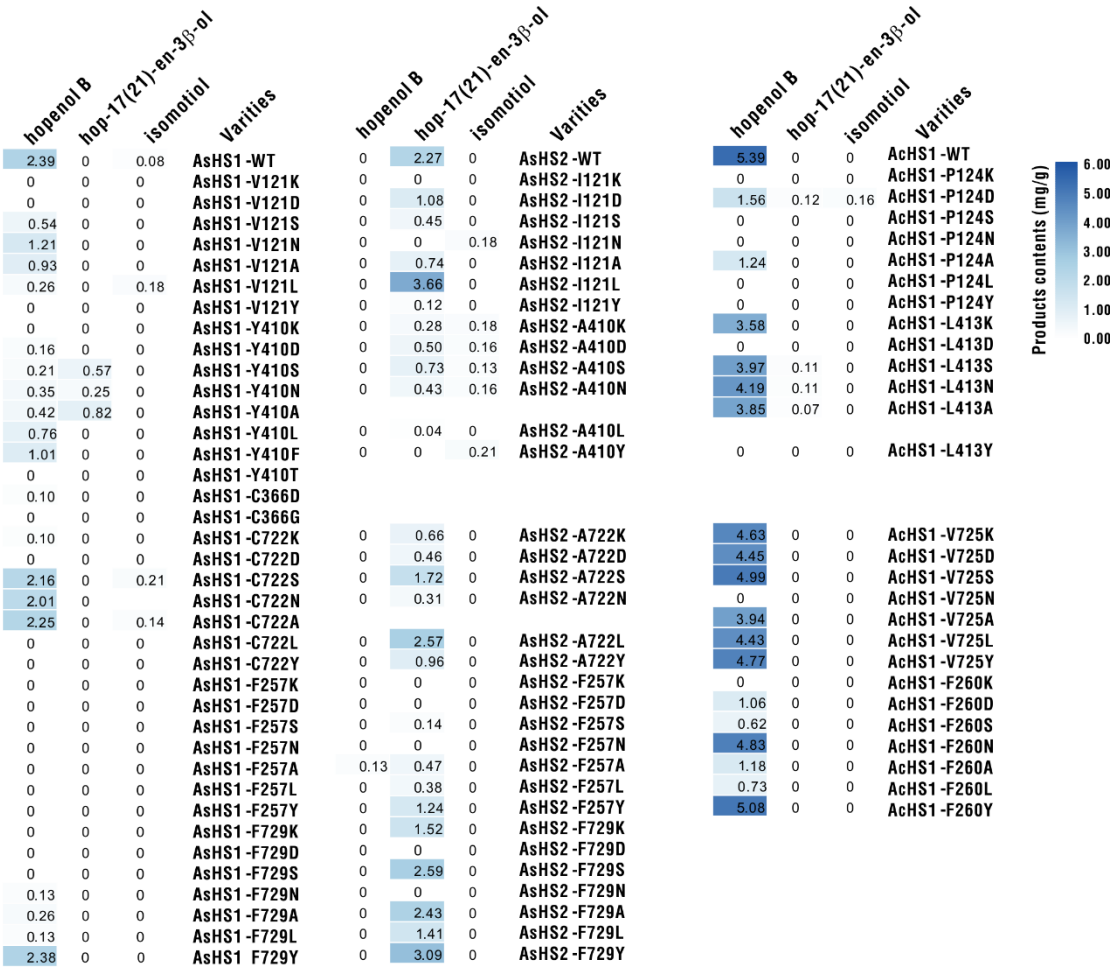

Fig. S9. Product profiles of *N. benthamiana* leaf extracts expressing the mutants of AsHS1, AsHS2 and AcHS1. Value represented by mean, n=3 biological replicates (contents of products mg/g).

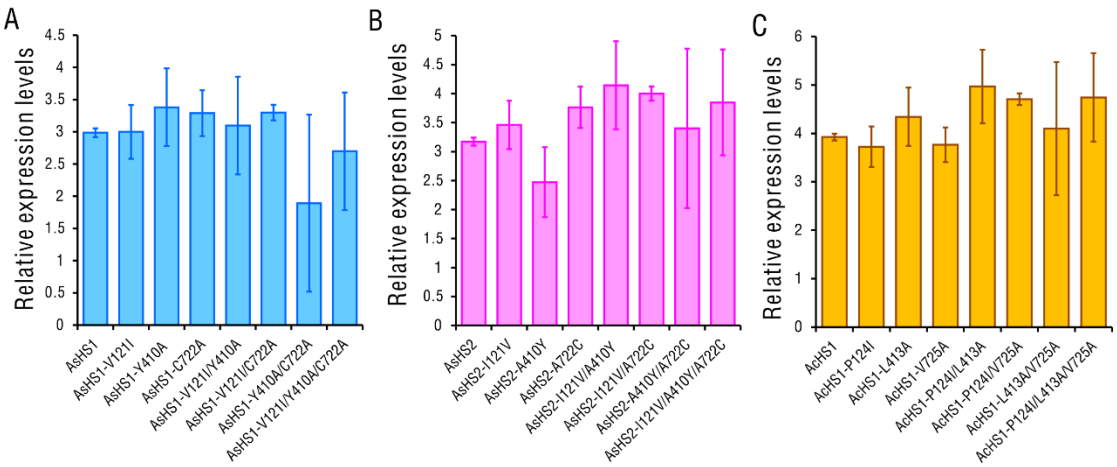

Fig. S10. **qRT-PCR analysis to quantify *OSC* mRNA levels in leaves of *N.benthamiana* expressing wild-type and mutant of OSCs.** (**A**) AsHS1 and its mutants; (**B**) AsHS2 and its mutants; (**C**) AcHS1 and its mutants. The relative expression levels of wild-type and all mutants are presented as means ± SE, n = 3 biologic replicates.

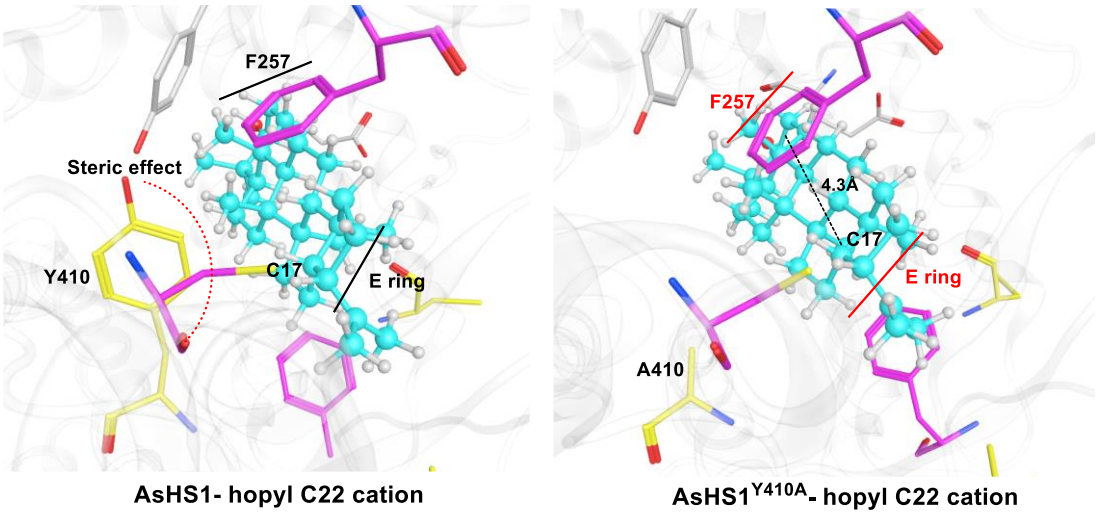

Fig. S11. The side view of representative structures of hopyl C22 cation (A state) in wild type AsHS1 and AsHS1Y410A mutant. The interactions between the intermediate and F257 is different due to the steric hinder effect of Y410 side chain. All these enzyme-intermediate structures are obtained from QM/MM modelling.

Table S1. List of primer sequences

Primers	Sequences
	For *AsHS1, AsHS2* clone
AsHS1F	5'-ATGTGGAGGCTGAAGGTGAGCGAG-3'
AsHS1R	5'-TCAGCTTTTGCTTGAAATAAGTCTGCGCC-3'
AsHS2F	5'-ATGTGGCGGCTGAAGGTGAGC-3'
AsHS2R	5'-TTACTTTCTGCTCGCAGCAAGTTTGCG-3'
	For AsHS1, AsHS2 and AcHS1 substitution
AsHS1V121I-F	5'-ACAGCGGGATTTTGTTCA**TAATGCCTATACTTATATTCTCTTTATATGTTAC-**3'
AsHS1V121I-R	5'-TGAACAAAATCCCGCTGT**AATCACCAGGCCAGTGGC**-3'
AsHS1Y410A-F	5'-GCAAAGATAGCCGATGG**TTGTCAAGGCTGGGAGATGC-**3'
AsHS1Y410A-R	5'-CCATCGGCTATCTTTGC**CTTCATGCCATCTTCAGCAAGC-**3'
AsHS1C722A-F	5'-GCGGGAGCGGCTA**ACTGCGCCTTTTACTTTAACTACC-**3'
AsHS1C722A-R	5'-TAGCCGCTCCCGC**GTGTTCCTGCTGAGGAAACTCG-**3'
AsHS2I121V-F	5'-CAGCGGAGTATTGTT**CCTTATGCCAATGTTTATATTCTCATTATATG-**3'
AsHS2I121V-R	5'-AACAATACTCCGCTG**TAATCACCGGGCCAATGGC**-3'
AsHS2A410Y-F	5'-GGCAAAGATATATGATGGCTG**CACTGGCTGGGATACATCATTT-**3'
AsHS2A410Y-R	5'-CAGCCATCATATATCTTTGCC**TTCATGCCATCTTCAGCGAGC-**3'
AsHS2A722C-F	5'-TACACTGGATGTGCTAACTC**GGCTTTTTTCTTTAACTAC-**3'
AsHS2A722C-R	5'-GAGTTAGCACATCCAGTGTA**TTCCTGCTGAGGAAACTCTCC-**3'
AcHS1P124I-F	5'-AATTCAGGAATCTTATTTCTGC**TTCCTTACTTGGTCATCA-**3'
AcHS1P124I-R	5'-GCAGAAATAAGATTCCTGAATT**CTCCGCAGGCCAATGCCC-**3'
AcHS1L413A-F	5'-AATGCAGAACGCTGGTAGTC**AAAACTGGGATACCTCGTTTG-**3'
AcHS1L413A-R	5'-GACTACCAGCGTTCTGCATT**TTCATTCCATCTTCTGCAATCC-**3'
AcHS1V725A-F	5'-ATATAACTGGAGCTTCTTTGAAG**AATTTAATGTTGCATT-**3'
AcHS1V725A-R	5'-CTTCAAAGAAGCTCCAGTTATAT**CCTGTTGGGGGA-**3'
	For qRT-PCR
AsHS1F	5'-CTTGTGGCTTGCTGAAGATG**-3'**
AsHS1R	5'-CGATGAGATCCGTGGAGTAAA**-3'**
AsHS2F	5'-CTTCTCCTACGCTTCCAGTATG**-3'**
AsHS2R	5'-CTTCAGTGACTTCGGTACTCTTT**-3'**
AcHS1F	5'-TCTGCTTCCTTACTTGGTCATC**-3'**
AcHS1R	5'-TGTGGCCCTCTATGTGAAATC**-3'**
NtACTIN-F	5'-CCTGAGGTCCTTTTCCAACCA**-3'**
	For AsHS1 substitution
AsHS1V121D-R	5'-TTATGAACAAATCCCCGCTGTAATCACCAGGCCAGTGGCCATCATGTGCC-3'
AsHS1V121S-F	5'-GATTACAGCGGGTCTTTGTTCATAATGCCTATACTTATATTCTCTTTAT-3'
AsHS1V121S-R	5'-CATTATGAACAAAGACCCGCTGTAATCACCAGGCCAGTGGCCATCATG-3'
AsHS1V121N-F	5'-GATTACAGCGGGAATTTGTTCATAATGCCTATACTTATATTCTCTTTATA-3'
AsHS1V121N-R	5'-GCATTATGAACAAATTCCCGCTGTAATCACCAGGCCAGTGGCCATCATGT-3'
AsHS1V121A-F	5'-GATTACAGCGGGGCTTTGTTCATAATGCCTATACTTATATTCTCTTTAT-3'
AsHS1V121A-R	5'-GCATTATGAACAAAGCCCCGCTGTAATCACCAGGCCAGTGGCCATCA-3'
AsHS1V121L-F	5'-TGATTACAGCGGGCTTTTGTTCATAATGCCTATACTTATATTCTCTTTATATG-3'
AsHS1V121L-R	5'-GCATTATGAACAAAAGCCCGCTGTAATCACCAGGCCAGTGGCCATCAT-3'
AsHS1V121Y-F	5'-GTGATTACAGCGGGTATTTGTTCATAATGCCTATACTTATATTCTCTTTATAT-3'
AsHS1V121Y-R	5'-CATTATGAACAAATACCCGCTGTAATCACCAGGCCAGTGGCCATCATG-3'
AsHS1Y410K-F	5'-GAAGGCAAAGATAAAGGATGGTTGTCAAGGCTGGGAGATGCCGTTTATAGTT-3'
AsHS1Y410K-R	5'-CTTGACAACCATCCTTTATCTTTGCCTTCATGCCATCTTCAGCAAGCCA-3'
AsHS1Y410D-F	5'-TGAAGGCAAAGATAGATGATGGTTGTCAAGGCTGGGAGATGCCGTTTATAG-3'
AsHS1Y410D-R	5'-CTTGACAACCATCATCTATCTTTGCCTTCATGCCATCTTCAGCAAGCCA-3'
AsHS1Y410S-F	5'-GAAGGCAAAGATATCTGATGGTTGTCAAGGCTGGGAGATGCCGTTTATAGTT-3'
AsHS1Y410S-R	5'-CTTGACAACCATCAGATATCTTTGCCTTCATGCCATCTTCAGCAAGC-3'
AsHS1Y410N-F	5'-AAGGCAAAGATAAATGATGGTTGTCAAGGCTGGGAGATGCCGTTTA-3'
AsHS1Y410N-R	5'-TTGACAACCATCATTTATCTTTGCCTTCATGCCATCTTCAGCAAGCC-3'
AsHS1Y410A-F	5'-AGGCAAAGATAGCTGATGGTTGTCAAGGCTGGGAGATGCCGTTTATA-3'
AsHS1Y410A-R	5'-CTTGACAACCATCAGCTATCTTTGCCTTCATGCCATCTTCAGCAAGC-3'

AsHS1Y410L-F	5'-AAGGCAAAGATACTTGATGGTTGTCAAGGCTGGGAGATGCCGTTTATAGTTC-3'
AsHS1Y410L-R	5'-CTTGACAACCATCAAGTATCTTTGCCTTCATGCCATCTTCAGCAAGCCACA-3'
AsHS1Y410F-F	5'-AAGGCAAAGATATTTGATGGTTGTCAAGGCTGGGAGATGCCGTTTATAGT-3'
AsHS1Y410F-R	5'-TTGACAACCATCAAATATCTTTGCCTTCATGCCATCTTCAGCAAGCCAC-3'
AsHS1C722K-F	5'-GGAACACGCGGGAAAAGCTAACTGCGCCTTTTACTTTAACTACCCAAATTAC-3'
AsHS1C722K-R	5'-AGGCGCAGTTAGCTTTTCCCGCGTGTTCCTGCTGAGGAAACTCGCCTGTGCC-3'
AsHS1C722D-F	5'-CACGCGGGAGATGCTAACTGCGCCTTTTACTTTAACTACCCAAATT-3'
AsHS1C722D-R	5'-GCAGTTAGCATCTCCCGCGTGTTCCTGCTGAGGAAACTC-3'
AsHS1C722S-F	5'-CACGCGGGATCTGCTAACTGCGCCTTTTACTTTAACTACCCA-3'
AsHS1C722S-R	5'-GCAGTTAGCAGATCCCGCGTGTTCCTGCTGAGGAAACTCGCCTGT-3'
AsHS1C722N-F	5'-ACGCGGGAAATGCTAACTGCGCCTTTTACTTTAACTACCCAAAT-3'
AsHS1C722N-R	5'-GCAGTTAGCATTTCCCGCGTGTTCCTGCTGAGGAAACTCGCCTGT-3'
AsHS1C722A-F	5'-CACGCGGGAGCTGCTAACTGCGCCTTTTACTTTAACTACCCAAAT-3'
AsHS1C722A-R	5'-GCGCAGTTAGCAGCTCCCGCGTGTTCCTGCTGAGGAAACTCGCCT-3'
AsHS1C722L-F	5'-CACGCGGGACTTGCTAACTGCGCCTTTTACTTTAACTACCCAAAT-3'
AsHS1C722L-R	5'-GCAGTTAGCAAGTCCCGCGTGTTCCTGCTGAGGAAACTCGCC-3'
AsHS1C722Y-F	5'-CACGCGGGATATGCTAACTGCGCCTTTTACTTTAACTACCCAAATT-3'
AsHS1C722Y-R	5'-GCAGTTAGCATATCCCGCGTGTTCCTGCTGAGGAAACTCGCCTGTG-3'
AsHS1F257K-F	5'-TCTGGTGCAAGTGCCGATTGGTGTATATGGCAATGGCATA-3'
AsHS1F257K-R	5'-AATCGGCACTTGCACCAGAATCGCCCTGGATGAATAGGAAGA-3'
AsHS1F257D-F	5'-ATTCTGGTGCGATTGCCGATTGGTGTATATGGCAATGGCATA-3'
AsHS1F257D-R	5'-CAATCGGCAATCGCACCAGAATCGCCCTGGATGAATAGGAAG-3'
AsHS1F257S-F	5'-TTCTGGTGCTCTTGCCGATTGGTGTATATGGCAATGGCATA-3'
AsHS1F257S-R	5'-CAATCGGCAAGAGCACCAGAATCGCCCTGGATGAATAGGAAGA-3'
AsHS1F257N-F	5'-TTCTGGTGCAATTGCCGATTGGTGTATATGGCAATGGCATA-3'
AsHS1F257N-R	5'-CAATCGGCAATTGCACCAGAATCGCCCTGGATGAATAGGAAGAA-3'
AsHS1F257A-F	5'-CTGGTGCGCTTGCCGATTGGTGTATATGGCAATGGCAT-3'
AsHS1F257A-R	5'-CAATCGGCAAGCGCACCAGAATCGCCCTGGATGAATAGGAAGA-3'
AsHS1F257L-F	5'-TCTGGTGCCTTTGCCGATTGGTGTATATGGCAATGGCAT-3'
AsHS1F257L-R	5'-AATCGGCAAAGGCACCAGAATCGCCCTGGATGAATAGGAAGAA-3'
AsHS1F257Y-F	5'-TTCTGGTGCTATTGCCGATTGGTGTATATGGCAATGGCATA-3'
AsHS1F257Y-R	5'-AATCGGCAATAGCACCAGAATCGCCCTGGATGAATAGGAAGA-3'
AsHS1F729K-F	5'-CCTTTTACAAAAACTACCCAAATTACAGGAATATTTTCCC-3'
AsHS1F729K-R	5'-GTAGTTTTTGTAAAAGGCGCAGTTAGCACATCCCGCGTGTTCC-3'
AsHS1F729D-F	5'-CCTTTTACGATAACTACCCAAATTACAGGAATATTTTCCC-3'
AsHS1F729D-R	5'-GGTAGTTATCGTAAAAGGCGCAGTTAGCACATCCCG-3'
AsHS1F729S-F	5'-CTTTTACTCTAACTACCCAAATTACAGGAATATTTTCCCCATTT-3'
AsHS1F729S-R	5'-GGTAGTTAGAGTAAAAGGCGCAGTTAGCACATCCCGCGTGTTCCT-3'
AsHS1F729N-F	5'-CCTTTTACAATAACTACCCAAATTACAGGAATATTTTCC-3'
AsHS1F729N-R	5'-GGTAGTTATTGTAAAAGGCGCAGTTAGCACATCCCGCGTGTTCCT-3'
AsHS1F729A-F	5'-CTTTTACGCTAACTACCCAAATTACAGGAATATTTTCCCCA-3'
AsHS1F729A-R	5'-GGTAGTTAGCGTAAAAGGCGCAGTTAGCACATCCCGCGTGTTCC-3'
AsHS1F729L-F	5'-CTTTTACCTTAACTACCCAAATTACAGGAATATTTTCCCC-3'
AsHS1F729L-R	5'-GGTAGTTAAGGTAAAAGGCGCAGTTAGCACATCCCGCGTGTTC-3'
AsHS1F729Y-F	5'-CTTTTACTATAACTACCCAAATTACAGGAATATTTTCCCCAT-3'
AsHS1F729Y-R	5'-GGTAGTTATAGTAAAAGGCGCAGTTAGCACATCCCGCGTGTT-3'
	For AsHS2 substitution
AsHS2I121K-F	5'-CAGCGGAAAGTTGTTCCTTATGCCAATGTTTATATTCTCA-3'
AsHS2I121K-R	5'-GGAACAACTTTCCGCTGTAATCACCGGGCCAATGGCCATCATGT-3'

AsHS2I121D-F	5'-CAGCGGAGATTTGTTCCTTATGCCAATGTTTATATTCTCATT-3'
AsHS2I121D-R	5'-GAACAAATCTCCGCTGTAATCACCGGGCCAATGGCCATCAT-3'
AsHS2I121S-F	5'-CAGCGGATCTTTGTTCCTTATGCCAATGTTTATATTCTCATTATA-3'
AsHS2I121S-R	5'-GGAACAAAGATCCGCTGTAATCACCGGGCCAATGGCCATCATG-3'
AsHS2I121N-F	5'-CAGCGGAAATTTGTTCCTTATGCCAATGTTTATATTCTCA-3'
AsHS2I121N-R	5'-GGAACAAATTTCCGCTGTAATCACCGGGCCAATGGCCATCATGT-3'
AsHS2I121A-F	5'-CAGCGGAGCTTTGTTCCTTATGCCAATGTTTATATTCTCAT-3'
AsHS2I121A-R	5'-GGAACAAAGCTCCGCTGTAATCACCGGGCCAATGGCCATCATG-3'
AsHS2I121L-F	5'-ACAGCGGACTTTTGTTCCTTATGCCAATGTTTATATTCTCAT-3'
AsHS2I121L-R	5'-AGGAACAAAAGTCCGCTGTAATCACCGGGCCAATGGCCATCATGT-3'
AsHS2I121Y-F	5'-ACAGCGGATATTTGTTCCTTATGCCAATGTTTATATTCTCATTA-3'
AsHS2I121Y-R	5'-AGGAACAAATATCCGCTGTAATCACCGGGCCAATGGCCATCATGTG-3'
AsHS2A410K-F	5'-CAAAGATAAAGGATGGCTGCACTGGCTGGGATACATCATTTA-3'
AsHS2A410K-R	5'-GCAGCCATCCTTTATCTTTGCCTTCATGCCATCTTCAGCGAG-3'
AsHS2A410D-F	5'-AAAGATAGATGATGGCTGCACTGGCTGGGATACAT-3'
AsHS2A410D-R	5'-CAGCCATCATCTATCTTTGCCTTCATGCCATCTTCAGC-3'
AsHS2A410S-F	5'-CAAAGATATCTGATGGCTGCACTGGCTGGGATACATCATTT-3'
AsHS2A410S-R	5'-GCAGCCATCAGATATCTTTGCCTTCATGCCATCTTCAGCGAG-3'
AsHS2A410N-F	5'-AAAGATAAATGATGGCTGCACTGGCTGGGATACATCA-3'
AsHS2A410N-R	5'-CAGCCATCATTTATCTTTGCCTTCATGCCATCTTCAGCGAG-3'
AsHS2A410L-F	5'-AAAGATACTTGATGGCTGCACTGGCTGGGATACATCA-3'
AsHS2A410L-R	5'-CAGCCATCAAGTATCTTTGCCTTCATGCCATCTTCAGCGAG-3'
AsHS2A410Y-F	5'-GCAAAGATATATGATGGCTGCACTGGCTGGGATACATCATTTATAATTCA-3'
AsHS2A410Y-R	5'-GCAGCCATCATATATCTTTGCCTTCATGCCATCTTCAGCGAG-3'
AsHS2A722K-F	5'-ACACTGGAAAGGCTAACTCGGCTTTTTTCTTTAACTACTCCA-3'
AsHS2A722K-R	5'-GAGTTAGCCTTTCCAGTGTATTCCTGCTGAGGAAACTCTCCT-3'
AsHS2A722D-F	5'-CACTGGAGATGCTAACTCGGCTTTTTTCTTTAACTACTCCAAC-3'
AsHS2A722D-R	5'-AGTTAGCATCTCCAGTGTATTCCTGCTGAGGAAACTCTCCTGT-3'
AsHS2A722S-F	5'-ACACTGGATCTGCTAACTCGGCTTTTTTCTTTAACTACTCCAAC-3'
AsHS2A722S-R	5'-AGTTAGCAGATCCAGTGTATTCCTGCTGAGGAAACTCTCCTG-3'
AsHS2A722N-F	5'-CACTGGAAATGCTAACTCGGCTTTTTTCTTTAACTACTCCAAC-3'
AsHS2A722N-R	5'-AGTTAGCATTTCCAGTGTATTCCTGCTGAGGAAACTCTCCTGT-3'
AsHS2A722L-F	5'-CACTGGACTTGCTAACTCGGCTTTTTTCTTTAACTACTCCA-3'
AsHS2A722L-R	5'-AGTTAGCAAGTCCAGTGTATTCCTGCTGAGGAAACTCTCCTG-3'
AsHS2A722Y-F	5'-TACACTGGATATGCTAACTCGGCTTTTTTCTTTAACTACTCCAACT-3'
AsHS2A722Y-R	5'-GAGTTAGCATATCCAGTGTATTCCTGCTGAGGAAACTCTCCTGTGTCTAG-3'
AsHS2F257K-F	5'-TATTGGTGCAAGTGCCGACTAGTGTATATGTCAATGGCATA-3'
AsHS2F257K-R	5'-AGTCGGCACTTGCACCAATATCGTCCTGGGTGTATTGGAAGAAA-3'
AsHS2F257D-F	5'-TTGGTGCGATTGCCGACTAGTGTATATGTCAATGGCATA-3'
AsHS2F257D-R	5'-GTCGGCAATCGCACCAATATCGTCCTGGGTGTATTGGAAGAA-3'
AsHS2F257S-F	5'-TTGGTGCTCTTGCCGACTAGTGTATATGTCAATGGCATA-3'
AsHS2F257S-R	5'-GTCGGCAAGAGCACCAATATCGTCCTGGGTGTATTGGAAGAAAATA-3'
AsHS2F257N-F	5'-ATTGGTGCAATTGCCGACTAGTGTATATGTCAATGGCATA-3'
AsHS2F257N-R	5'-GTCGGCAATTGCACCAATATCGTCCTGGGTGTATTGGAAGAAAA-3'
AsHS2F257A-F	5'-ATTGGTGCGCTTGCCGACTAGTGTATATGTCAATGGCATATCTT-3'
AsHS2F257A-R	5'-GTCGGCAAGCGCACCAATATCGTCCTGGGTGTATTGGAAGAAAATA-3'
AsHS2F257L-F	5'-TTGGTGCCTTTGCCGACTAGTGTATATGTCAATGGCATATCT-3'
AsHS2F257L-R	5'-GTCGGCAAAGGCACCAATATCGTCCTGGGTGTATTGGAAGAAAATAA-3'
AsHS2F257Y-F	5'-TTGGTGCTATTGCCGACTAGTGTATATGTCAATGGCATATCTTTA-3'

AsHS2F257Y-R	5'-GTCGGCAATAGCACCAATATCGTCCTGGGTGTATTGGAAGAAAATAA-3'
AsHS2F729K-F	5'-CTTTTTTCAAGAACTACTCCAACTATCGCAACATCTACCCCATTATGGCTCT-3'
AsHS2F729K-R	5'-GAGTAGTTCTTGAAAAAAGCCGAGTTAGCAGCTCCAGTGTATTCCTGC-3'
AsHS2F729D-F	5'-CTTTTTTCGATAACTACTCCAACTATCGCAACATCTACCCCATTATGG-3'
AsHS2F729D-R	5'-GAGTAGTTATCGAAAAAAGCCGAGTTAGCAGCTCCAGTGTATTCCTG-3'
AsHS2F729S-F	5'-CTTTTTTCTCTAACTACTCCAACTATCGCAACATCTACCCCATTATGG-3'
AsHS2F729S-R	5'-AGTAGTTAGAGAAAAAAGCCGAGTTAGCAGCTCCAGTGTATTCC-3'
AsHS2F729N-F	5'-CTTTTTTCAATAACTACTCCAACTATCGCAACATCTACCCCATTATGG-3'
AsHS2F729N-R	5'-AGTAGTTATTGAAAAAAGCCGAGTTAGCAGCTCCAGTGTATT-3'
AsHS2F729A-F	5'-CTTTTTTCGCTAACTACTCCAACTATCGCAACATCTACCCCATTATGGCTC-3'
AsHS2F729A-R	5'-GAGTAGTTAGCGAAAAAAGCCGAGTTAGCAGCTCCAGTGTATTCCTGCTG-3'
AsHS2F729L-F	5'-CTTTTTTCCTTAACTACTCCAACTATCGCAACATCTACCCCATTATGGCTCT-3'
AsHS2F729L-R	5'-AGTAGTTAAGGAAAAAAGCCGAGTTAGCAGCTCCAGTGTATTCCT-3'
AsHS2F729Y-F	5'-CTTTTTTCTATAACTACTCCAACTATCGCAACATCTACCCCATTATGGCTCT-3'
AsHS2F729Y-R	5'-AGTAGTTATAGAAAAAAGCCGAGTTAGCAGCTCCAGTGTATTCC-3'
	For AcHS1 substitution
AcHS1P124K-F	5'-ATTCAGGAAAGTTATTTCTGCTTCCTTACTTGGTCATCACATTGTACAT-3'
AcHS1P124K-R	5'-AGAAATAACTTTCCTGAATTCTCCGCAGGCCAATGCCCATCGTGGCT-3'
AcHS1P124D-F	5'-TTCAGGAGATTTATTTCTGCTTCCTTACTTGGTCATCACATTG-3'
AcHS1P124D-R	5'-AAATAAATCTCCTGAATTCTCCGCAGGCCAATGCCCATCGTG-3'
AcHS1P124S-F	5'-ATTCAGGATCTTTATTTCTGCTTCCTTACTTGGTCATCACATTGTACAT-3'
AcHS1P124S-R	5'-GAAATAAAGATCCTGAATTCTCCGCAGGCCAATGCCCATCGTGGCTTTGC-3'
AcHS1P124N-F	5'-ATTCAGGAAATTTATTTCTGCTTCCTTACTTGGTCATCACATTGTACAT-3'
AcHS1P124N-R	5'-GAAATAAATTTCCTGAATTCTCCGCAGGCCAATGCCCATCGTGGC-3'
AcHS1P124A-F	5'-ATTCAGGAGCTTTATTTCTGCTTCCTTACTTGGTCATCACATTGTACATTA-3'
AcHS1P124A-R	5'-AGAAATAAAGCTCCTGAATTCTCCGCAGGCCAATGCCCATCGTGGCTTT-3'
AcHS1P124L-F	5'-ATTCAGGACTTTTATTTCTGCTTCCTTACTTGGTCATCACATTGTACA-3'
AcHS1P124L-R	5'-GAAATAAAAGTCCTGAATTCTCCGCAGGCCAATGCCCATC-3'
AcHS1P124Y-F	5'-ATTCAGGATATTTATTTCTGCTTCCTTACTTGGTCATCACATTG-3'
AcHS1P124Y-R	5'-GAAATAAATATCCTGAATTCTCCGCAGGCCAATGCCCATCGTGGC-3'
AcHS1L413K-F	5'-TGCAGAACAAGGGTAGTCAAAACTGGGATACCTCGTTTGCTCTTC-3'
AcHS1L413K-R	5'-GACTACCCTTGTTCTGCATTTTCATTCCATCTTCTGCAATCC-3'
AcHS1L413D-F	5'-TGCAGAACGATGGTAGTCAAAACTGGGATACCTCGTTTGCTCT-3'
AcHS1L413D-R	5'-TGACTACCATCGTTCTGCATTTTCATTCCATCTTCTGCAATCCA-3'
AcHS1L413S-F	5'-GCAGAACTCTGGTAGTCAAAACTGGGATACCTCGTTTGCTC-3'
AcHS1L413S-R	5'-GACTACCAGAGTTCTGCATTTTCATTCCATCTTCTGCAATCCA-3'
AcHS1L413N-F	5'-TGCAGAACAATGGTAGTCAAAACTGGGATACCTCGTTTGCTCTT-3'
AcHS1L413N-R	5'-TGACTACCATTGTTCTGCATTTTCATTCCATCTTCTGCAATCC-3'
AcHS1L413A-F	5'-GCAGAACGCTGGTAGTCAAAACTGGGATACCTCGTTTGCTCTTC-3'
AcHS1L413A-R	5'-TGACTACCAGCGTTCTGCATTTTCATTCCATCTTCTGCAATCCATA-3'
AcHS1L413Y-F	5'-TGCAGAACTATGGTAGTCAAAACTGGGATACCTCGTTTGCTCTTC-3'
AcHS1L413Y-R	5'-TTGACTACCATAGTTCTGCATTTTCATTCCATCTTCTGCAATCCA-3'
AcHS1V725K-F	5'-TAACTGGAAAGTCTTTGAAGAATTTAATGTTGCATTACGCAGCATA-3'
AcHS1V725K-R	5'-TTCAAAGACTTTCCAGTTATATCCTGTTGGGGGAAACTCCCATTGT-3'
AcHS1V725D-F	5'-TAACTGGAGATTCTTTGAAGAATTTAATGTTGCATTACGCAGCATAC-3'
AcHS1V725D-R	5'-TTCAAAGAATCTCCAGTTATATCCTGTTGGGGGAAACTCCCATTGTCC-3'
AcHS1V725S-F	5'-TAACTGGATCTTCTTTGAAGAATTTAATGTTGCATTACGCAGCATAC-3'
AcHS1V725S-R	5'-TTCAAAGAAGATCCAGTTATATCCTGTTGGGGGAAACTCCCATTG-3'
AcHS1V725N-F	5'-TAACTGGAAATTCTTTGAAGAATTTAATGTTGCATTACGCAGCATAC-3'

AcHS1V725N-R	5'-TCAAAGAATTTCCAGTTATATCCTGTTGGGGGAAACTCCCATTG-3'
AcHS1V725A-F	5'-TAACTGGAGCTTCTTTGAAGAATTTAATGTTGCATTACGCAGCATACA-3'
AcHS1V725A-R	5'-TCAAAGAAGCTCCAGTTATATCCTGTTGGGGGAAACTCCCATTGT-3'
AcHS1V725L-F	5'-TAACTGGACTTTCTTTGAAGAATTTAATGTTGCATTACGCAGCATACA-3'
AcHS1V725L-R	5'-TCAAAGAAAGTCCAGTTATATCCTGTTGGGGGAAACTCCCATTGT-3'
AcHS1V725Y-F	5'-TAACTGGATATTCTTTGAAGAATTTAATGTTGCATTACGCAGCATACAGA-3'
AcHS1V725Y-R	5'-TTCAAAGAATATCCAGTTATATCCTGTTGGGGGAAACTCCCATTGTCC-3'
AcHS1F260K-F	5'-ATGGTGTAAGGCTCGCTTGACATATTTGCCTATGTCATATTTATATGG-3'
AcHS1F260K-R	5'-AGCGAGCCTTACACCATAGATTTCCTGGATACATAGGAAAATATG-3'
AcHS1F260D-F	5'-TATGGTGTGATGCTCGCTTGACATATTTGCCTATGTCATATTTATATGG-3'
AcHS1F260D-R	5'-AGCGAGCATCACACCATAGATTTCCTGGATACATAGGAAAATATG-3'
AcHS1F260S-F	5'-TATGGTGTTCTGCTCGCTTGACATATTTGCCTATGTCATATTTATATGG-3'
AcHS1F260S-R	5'-AGCGAGCAGAACACCATAGATTTCCTGGATACATAGGAAAATATGA-3'
AcHS1F260N-F	5'-ATGGTGTAATGCTCGCTTGACATATTTGCCTATGTCATATTTATATGG-3'
AcHS1F260N-R	5'-AGCGAGCATTACACCATAGATTTCCTGGATACATAGGAAAATAT-3'
AcHS1F260A-F	5'-ATGGTGTGCTGCTCGCTTGACATATTTGCCTATGTCATATTTATATG-3'
AcHS1F260A-R	5'-AGCGAGCAGCACACCATAGATTTCCTGGATACATAGGAAAATATGAT-3'
AcHS1F260L-F	5'-TATGGTGTCTTGCTCGCTTGACATATTTGCCTATGTCATATTTATATGG-3'
AcHS1F260L-R	5'-AGCGAGCAAGACACCATAGATTTCCTGGATACATAGGAAAATATG-3'
AcHS1F260Y-F	5'-ATGGTGTTATGCTCGCTTGACATATTTGCCTATGTCATATTTATATGGG-3'
AcHS1F260Y-R	5'-AGCGAGCATAACACCATAGATTTCCTGGATACATAGGAAAATATGA-3'

Table S2. Annotation of the 13 oxidosqualene cyclase (OSC) genes in *Avena strigosa*

Gene name	Scaffold name (V0.8)	Stand	Gene coverage in genome (bp)	Coding sequence length (bp)	Predict protein length (aa)	No. Exons	Ro	Rt	Le	Pa	Sh	Sp
							Low				High	
AsOSC1/AsCS1	AS01_000854_0062888	+	7,607	2280	759	18	34.11	18.63	31.95	87.84	42.17	108.38
AsOSC2	AS01_000854_0080506	+	12,074	2283	760	18	0.06	0.02	0.41	13.37	0.15	7.75
AsOSC3/Sad1	AS01_003827_0085775	-	7,340	2274	757	18	103.11	259.97	0	0.13	0.1	0.13
AsOSC4	AS01_010666_0134954	-	14,394	2271	756	18	5.21	0	0.01	0.01	0	0
AsOSC5	AS01_002190_0178042	-	28,176	2256	751	18	1	0.59	131.46	27.65	67.75	10.2
AsOSC6	AS01_003668_0201207	-	56,897	2256	751	18	0.28	0.07	36.96	5.78	19.2	1.75
AsOSC7/AsHS1	AS01_016459_0007535	+	35,790	2268	755	18	0.15	0	0	139.63	0.24	10.48
AsOSC8/AsHS2	AS01_003995_0089539	+	24,563	2268	755	18	0	0	0.01	0.07	1.49	0.28
AsOSC9	AS01_002578_0337646	+	10,504	2292	763	18	0	0	0	0.3	0	0.01
AsOSC10	AS01_000627_0460766	-	42,532 (partial)	2028	675	16	0.03	0	0.34	0.03	0.08	0.04
AsOSC11	AS01_002190_0218503	-	85,577 (partial)			-	0	0	0	0	0	0
AsOSC12	AS01_001450_0008578	+	35,852	2301	766	18	0	0	27.92	0.82	0.87	0.05
AsOSC13	AS01_001226_0028600	+	6,559	2271	756	18	0.47	0.14	0.66	79.66	0.9	3.3

Note：The expression patterns of OSC genes (FPKM) are derived from RNA-seq data for six different tissues of *A. strigosa* accession S75 (whole roots: Ro, root tips: Rt, leaves: Le, panicles: Pa, shoots: Sh and spikelets: Sp) were generated previously (23).

Supplementary Dataset

Dataset S1(separate file). Structural identification of hopenol B

Electron impact (EI) mass spectral fragmentation, ^{1}H-NMR, ^{13}C-NMR and 2D NMR spectrum (in the following order: ^{1}H-^{1}H COSY, HSQC, HMBC, ROESY) of hopenol B isolated from *N.benthamiana* leaves expressing AsHS1

Dataset S2 (separate file). Structural identification of hop-17(21)-en-3*β*-ol.

Electron impact (EI) mass spectral fragmentation, ^{1}H-NMR, ^{13}C-NMR and 2D NMR spectrum (in the following order: ^{1}H-^{1}H COSY, HSQC, HMBC, ROESY) of hop-17(21)-en-3*β*-ol isolated from *N.benthamiana* leaves expressing AsHS2

Dataset S3 (separate file). Structural identification of isomotiol.

Electron impact (EI) mass spectral fragmentation, ^{1}H-NMR, ^{13}C-NMR and 2D NMR spectrum (in the following order: ^{1}H-^{1}H COSY, HSQC, HMBC, ROESY) of isomotiol isolated from *N.benthamiana* leaves expressing the mutant AsHS2$^{I121V/A410V}$.

Dataset S4 (separate file). Sequence alignments of all OSCs in Fig.S1

Dataset S5 (separate file). Sequences of candidate metabolic genes for coexpression analysis in Fig. S2.

SI References

1. K. Tunyasuvunakool *et al.*, Highly accurate protein structure prediction for the human proteome. *Nature.* **596**,590-596 (2021).
2. Y. Duan *et al.*, A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comput Chem.* **24**,1999-2012 (2003).
3. W. Jorgensen *et al.*, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926-935 (1983).
4. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general amber force field. *J Comput Chem.* **25**,1157-1174 (2004).
5. C Bayly et al., A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **97**, 10269-10280 (1993).
6. M. Frisch, Gaussian 09, Revision A.02. Gaussian Inc. Wallingford, CT, USA, (2009).
7. DA. Case, TA. Darden, TE. Cheatham, C. Simmerling，PA Kollman, AMBER 12, University of California, San Francisco (2012).
8. JP. Ryckaert，G. Ciccotti，HJC. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327-341 (1977).
9. Y. Zhao, D. Truhlar, The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **120**, 215-241 (2008).
10. Y. Zhao, D. Truhlar, Exploring the limit of accuracy of the global hybrid meta density functional for main-group thermochemistry, kinetics, and noncovalent interactions. *J. Chem. Theory Comput.* **4**, 1849-1868 (2008).
11. N. Chen, J. Zhou, J. Li, J. Xu, R. Wu, Concerted cyclization of lanosterol C-ring and D-ring under human oxidosqualene cyclase catalysis: an ab initio QM/MM MD study. *J Chem Theory Comput.* **10**,1109-1120 (2014).
12. N. Chen, S. Wang, L. Smentek, B. A. Hess, Jr.R. Wu, Biosynthetic mechanism of lanosterol: cyclization. *Angew Chem Int Ed Engl.* **54**,8693-8696 (2015).
13. Y. Chu, Q. Xu, Q. H. Guo, Understanding energetic origins of product specificity of SET8 from QM/MM free energy simulations: what causes the stop of methyl addition during histone lysine

methylation? *J. Chem. Theory Comput.* **6**, 1380-1389 (2010).
14. Y. Chu, J.Yao, H. Guo, QM/MM MD and free energy simulations of G9a-Like Protein (GLP) and its mutants: understanding the factors that determine the product specificity. *Plos One*. **7**, e37674 (2012).
15. P. Hu, S. Wang, Y. Zhang, How do SET-domain protein Lysine methyltransferases achieve the methylation state specificity? Revisited by ab initio QM/MM molecular dynamics simulations. *J. Am. Chem. Soc.* **130**, 3806-3813 (2008).
16. Y. Zhou, S. Wang, Y. Li, Y. Zhang, Born-oppenheimer ab initio QM/MM molecular dynamics simulations of enzyme reactions. *Methods Enzymol*. **577**, 105-118 (2016).
17. Y. Zhang, H. Liu, W. Yang, Free energy calculation on enzyme reactions with an efficient iterative procedure to determine minimum energy paths on a combined ab initio QM/MM potential energy surface. *J. Chem. Phys.* **112**, 3483-3492 (2000).
18. Y. Shao *et al.*, Advances in methods and algorithms in a modern quantum chemistry program package. *Phys Chem Chem Phys* **8**,31723191 (2006)
19. A. Waterhouse *et al.*, SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296-W303 (2018).
20. R. Thimmappa, K. Geisler, T. Louveau, P. O'Maille, A. Osbourn, Triterpene biosynthesis in plants. *Annu Rev Plant Biol.* **65**, 225-257 (2014).
21. D. L. Filiault *et al.*, The Aquilegia genome provides insight into adaptive radiation and reveals an extraordinarily polymorphic chromosome with a unique history. *Elife.* **7** (2018).
22. K. Tunyasuvunakool *et al*., Highly accurate protein structure prediction for the human proteome. *Nature* **596**, 590-599 (2021).
23. A. Orme *et al.*, A noncanonical vacuolar sugar transferase required for biosynthesis of antimicrobial defense compounds in oat. *Proc Natl Acad Sci U S A.* **116**, 27105-27114 (2019).