Supplemental Materials and Methods

LC-MS conditions for quantification of endogenous aroma glycosides in *C. sinensis*.

LC–MS analysis of ten aroma glycosidic precursors was performed on LCMS-2010A system (Shimadzu; www.shimadzu.com) equipped with Capcell Pak UG120 C_{18} reversed phase column (2.0 mm i.d. x 150 mm, 5 µm, Shiseidou; hplc.shiseido.co.jp/e) and electrospray operating conditions were used the followings: dry gas 1.5 l/min, capillary voltage 1.5 kV, dry gas temperature 250°C. The endogenous aroma glycosides were quantified with the SIM negative mode. For benzyl-glc, benzyl-pri, 2PE-glc, 2PE-pri, (*Z*)-3-hexenyl-glc and (*Z*)-3-hexenyl-pri, the glycosidic fractions of were subjected to the LC-MS using gradient elution with aqueous formic acid (0.1%,v/v) as solvent A and acetonitrile as solvent B at a flow rate of 0.2 ml/min at 40°C. The gradient condition started with isocratic conditions of 16% of solvent B for 5 min and then increased up to 29% of solvent B for 24.5 min. The *m/z* 293 ion for benzyl-glc (*tR* 10.5 min), the *m/z* 425 ion for benzyl-pri (*tR* 14.9 min) the *m/z* 307 ion for 2PE-glc (*tR* 18.3 min), the *m/z* 439 ion for 2PE-pri (*tR* 22.4 min) The *m/z* 285 ion for (*Z*)-3-hexenyl-glc (*tR* 21.6 min), the *m/z* 417 ion for (Z)-3-hexenyl-pri (*tR* 26.3 min) and the *m/z* 456 ion for *p*NP-pri (*tR* 8.2 min) as a internal standard. For geranyl-glc, geranyl-pri, linalyl-glc and linalyl-pri, the glycosidic fractions of were subjected to the LC-MS using gradient elution with aqueous formic acid (0.1%,v/v) as solvent A and acetonitrile as solvent B at a flow rate of 0.2 ml/min at 40°C. The gradient condition started with 7% of solvent B then increased up to 30% of solvent B for 4 min and keep 30% of solvent B for 13 min. The *m/z* 339 ion for geranyl-glc (*tR* 14.5 min) and linalyl-glc (*tR* 13.0 min), and the *m/z* 471 ion for geranyl-pri (*tR* 11.9 min) and linalyl-glc (*tR* 11.0 min)

Enzyme purification of a xylosyltransferase specific for monoglucoside-bound

volatiles

The enzyme solution was applied to a HiTrap DEAE FF (5 ml, GE Healthcare) equilibrated with buffer B and eluted by a linear gradient of NaCl (0 to 340 mM) of buffer B. The active fraction was equilibrated with buffer C using an Amicon Ultra-15 centrifugal filter device (Mr 30,000 cutoff; Millipore). Then, the enzyme fraction was applied to a HiTrap Q FF (5 ml, GE Healthcare) and eluted with a linear gradient of NaCl (0 to 1 M) of buffer C. The active fractions was equilibrated with buffer D and applied to a Macro-prep Ceramic Hydroxyapatite Type III (5ml, Bio-Rad), and eluted with a linear gradient of NaCl (0 to 1 M) in buffer D. The active fraction was equilibrated with buffer E, applied to a HiTrap Blue HP (1 ml, GE Healthcare) and eluted with a linear gradient of NaCl (0 to 2 M) of buffer E. The active fraction was equilibrated with buffer C and applied to a Mono Q 5/50 GL (1 ml, GE Healthcare) and eluted with a linear gradient of NaCl (0 to 1 M) of buffer C. Peptide sequences were analyzed by Nano-LC-TOF MS.

Nano-LC-MS/MS analysis of partial peptide sequences of purified CsGT2

Peptide assignments were performed using an LC-ESI-LIT-q-TOF mass spectrometer equipped with NanoFrontier eLD System (Hitachi High-Technologies, www.hitachi-hitec.com) and a nano-flow HPLC, NanoFrontier nLC (Hitachi High-Technologies). The LIT-TOF and CID modes were used for MS detection and peptide fragmentation, respectively. The trypsin-treated sample (10 μl) was injected, and the peptides were trapped on a C18 column, Monolith Trap (50 μm × 150 mm, Hitachi High-Technologies). Peptide separation was achieved using a packed nano-capillary column (capillary-Ex nano mono cap, 0.05 × 150 mm, GL Science, Japan) at a flow rate of 200 nl/min. The peptides were eluted using an acetonitrile gradient (A: 2% acetonitrile containing 0.1% formic acid; B: 98% acetonitrile containing 0.1% formic acid; 0 min with A = 98%, B = 2%, followed by 60 min with A = 60%, B = 40%). All peptide mass data were analyzed using

Peaks software (Bioinformatics Solutions Inc., www.bioinfor.com) and the MASCOT database (Matrix Science, www.matrixscience.com).

LC-MS analysis of CsGTs enzymatic products

Enzymatic reaction mixtures by UGT enzymes were analyzed by LC-MS using Shimadzu LCMS-2010A (www.shimadzu.com) system with negative electrospray ionization mode; dry temperature, 250°C; a flow rate, 0.2 ml/min and column oven temperature, 40 ^{o}C. HPLC separation was performed according to the following conditions: Analysis of CsGT1 enzymatic products was performed using a CAPCELL PAK C_{18} UG120 column (2.0 mm I.D.×150 mm; 5 μm) (Shiseido, hplc.shiseido.co.jp/e/) and elution was started with 15% solvent B, following a linear gradient flow up to 90% in 15 min (solvent A: H_2O containing 0.05% (v/v) formic acid, B: acetonitrile) Analysis of CsGT2 enzymatic products was performed using a CAPCELL PAK C_{18} UG120 column (2.0 mm I.D.×75 mm; 5 μm) (Shiseido) and elution was started with 15% solvent B (0-2 min) following a linear gradient flow up to 60% in 8 min (solvent A: H_2O containing 0.05% (v/v) formic acid, B: acetonitrile)

Subcloning of CsGT1 and CsGT2 to pET15b vector for *E.coli* expression

The amplified cDNA fragments of CsGT1 and CsGT2 genes were subcloned into pENTR/D-TOPO vector (Life Technologies; www.lifetechnologies.com), and then digested with *Xho*I and *Bgl*II for CsGT1, and *Nde*I and *BamH*I for CsGT2 and the resulting DNA fragments were ligated into a pET-15b vector (Merck Millipore; www.merckmillipore.com) previously digested with *Xho*I and *BamH*I for CsGT1, and *Nde*I and *BamH*I for CsGT2, respectively. The resulting plasmids were named pET-15b-CsGT1 and pET-15b-CsGT2, which encoded an N-terminal in-frame fusion of CsGT1 and CsGT2, respectively, with a Hisx6 tag. The inserted fragments were sequenced to confirm the absence of PCR errors and transformed into *E. coli* BL21(DE3) (TOYOBO; www.toyobo-global.com).

Subcloning of CsGT2 (I141S) to pET15b vector

According to the previously published method of Noguchi *et al*. (2007), *in vitro* mutagenesis of the CsGT2 gene was performed using recombinant PCR with the pENTR-Directional-TOPO vector (Life Technologies) containing the wild-type CsGT2 cDNAs as the templates and the specific mutagenic oligonucleotide primer set (CsGT2-134-I141S-FW and CsGT2-134-I141S-Rv) (Supplemental Table S2) to obtain the CsGT2-I141S mutant. The amplified fragment containing the I141S mutation was digested with *Nde*I and *Bgl*II, and the resulting DNA fragments were ligated with pET-15b at *Nde*I and *BamH*I site as described above. The introduced mutation was verified by DNA sequencing of both strands. The resulting CsGT2-I141S fragment was inserted in the expression vector pET-15b as described above.

***E.coli* expression and purification of recombinant CsGT1 and CsGT2**

The transformant cells were precultured at 37°C for 16 hr in a Luria-Bertani broth containing 50 μg/mL ampicillin. Twenty milliliters of the pre-culture was then inoculated into 200 mL, then further incubated at 22°C for 20 hr with Overnight Express Autoinduction System 1 (Merck Millipore) for expression of recombinant proteins. All subsequent operations were conducted at 0 to 4°C. The recombinant *E. coli* cells were harvested by centrifugation (7000*g*, 15 min), washed with distilled water, and resuspended in buffer F containing 20 mM imidazole. The cells were disrupted at 4°C by five cycles of ultrasonication. The cell debris was removed by centrifugation (7,000*g*, 15 min). Polyethyleneimine was slowly added to the supernatant solution to a final concentration of 0.12% (v/v). The mixture was allowed to stand at 4°C for 30 min, followed by centrifugation (7,000*g*, 15 min). The supernatant was applied to a HisTrap HP column (1 mL; GE Healthcare) equilibrated with buffer F containing 20 mM imidazole. The column was washed with buffer F containing 20 mM imidazole, and

the enzyme was eluted with buffer F containing 200 mM imidazole. The active column-bound fractions were concentrated and desalted using Vivaspin 30,000 MWCO (GE Healthcare), followed by substitution with buffer G. The expressed recombinant proteins in the gels were separated by SDS-PAGE and visualized by using PAGE Blue 83 (Cosmo Bio; www.cosmobio.co.jp/index_e) and His・Tag monoclonal antibody (mouse, Merck Millipore) and ECL Anti-mouse IgG Horseradish Peroxidase-linked Whole antibody (sheep, GE Healthcare) respectively. Chemical luminescence derived from the recombinant proteins was detected by using Chemi-Lumi One Ultra (Nacalai tesque; www.nacalai.co.jp/english/) in Molecular Imager ChemiDoc XRS+ (BioRad; www.bio-rad.com/)

Preparation of cryosections of tea leaf for MS imaging

For MS imaging analysis, cryosections of the young tea leaf were prepared as described elsewhere. Briefly, tea leaves were embedded in SCEM embedding medium (Leica, www.leica-microsystems.com) in a Tissue-Tek Cryomold (Sakura Finetek, www.sakura-finetek.com/) and desiccated in *vacuo* for 5 min at ambient temperature. The leaves were frozen by placing the mold containing tea leaves embedded in the medium on a block of dry ice for 15 min. Longitudinal sections of the leaves in 5 μm thickness were cut using a cryotome (Cryostar NX70, Thermo Scientific; www.thermosci.jp) with blade and block temperatures of −20 and −25 °C, respectively, and the sections were transferred onto a glass slide coated with indium tin oxide (Bruker Daltonics; www.bruker.jp). The sections were dried in 50 ml tubes containing dried silica gel. The cryosections were uniformly sprayed with a solution of 50 mg/ml dihydrobenzoic acid (DHBA) in 70% methanol containing 0.1% trifluoroacetic acid using a 0.2-mm caliber airbrush (Procon Boy FWA Platinum; Mr. Hobby, Japan), air-dried for 10 min and mounted on a MALDI target plate.

MALDI-MS imaging analysis

Longitudinal sections of seeds were selected for MS imaging analysis, performed by MALDI-TOF-MS/MS (Ultraflextream III, Bruker Daltonics). Parameters of the laser shots were as follows: laser diameter, 15 μm; raster, 15 μm; laser power, 3_mid; laser power boost, 50%. Data from 100 laser shots were obtained at each position using the FlexImaging software program (Bruker Daltonics). The spectra were acquired in the positive mode. For MS/MS analysis, the parent ion of (*Z*)-3-hexenyl-pri (*m/z* 417), (*Z*)-3-hexenyl- glu (*m/z* 284) and geranyl-glc (*m/z* 340) were monitored for a peak tolerance of 0.5 kDa. A cutoff value of 50% was used for the production of the images. The instrument was calibrated in the positive ion mode in the low-mass range using a mixture of selected chemical standards (DHBA and 4-hydroxycinnamic acid) prior to data acquisition. Authentic (*Z*)-3-hexenyl-pri was used as a reference compound. The images of the parent ion were overlaid onto bright-field images of the corresponding leaf sections with the use of FlexImaging software.