

Genome-wide identification of OSC gene family and potential function in the synthesis of ursane- and oleanane-type triter-pene in *Momordica charantia*

Yutong Han^{a,b,c,d,f,l}, Ya Yang^{b,c,d,l}, Yan Li^{a,b,c,d,f,l}, Xin Yin^{b,c,d,f}, Zhiyu Chen^{a,b,c,d,f}, Danni Yang^{b,c,d,f}, Yongping Yang^{b,c,d}, Yunqiang Yang^{b,c,d*}, Xuefei Yang^{a,b,e*}

^a Key Laboratory of Economic Plants and Biotechnology, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China

^b Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Science, Kunming, 650204, China

^c Plant Germplasm and Genomics Center, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China

^d Institute of Tibetan Plateau Research at Kunming, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China

^e Southeast Asia Biodiversity Research Institute, Chinese Academy of Sciences, Yezin, NayPyiTaw 05282, Myanmar

^f University of Chinese Academy of Sciences, Beijing 100049, China

^l These authors contributed equally to this work.

* Correspondence: Yunqiang Yang, yangyunqiang@mail.kib.ac.cn; Tel.: 86-871-65230873. Xuefei Yang, xuefei@mail.kib.ac.cn; Tel.: 86-871-65223398

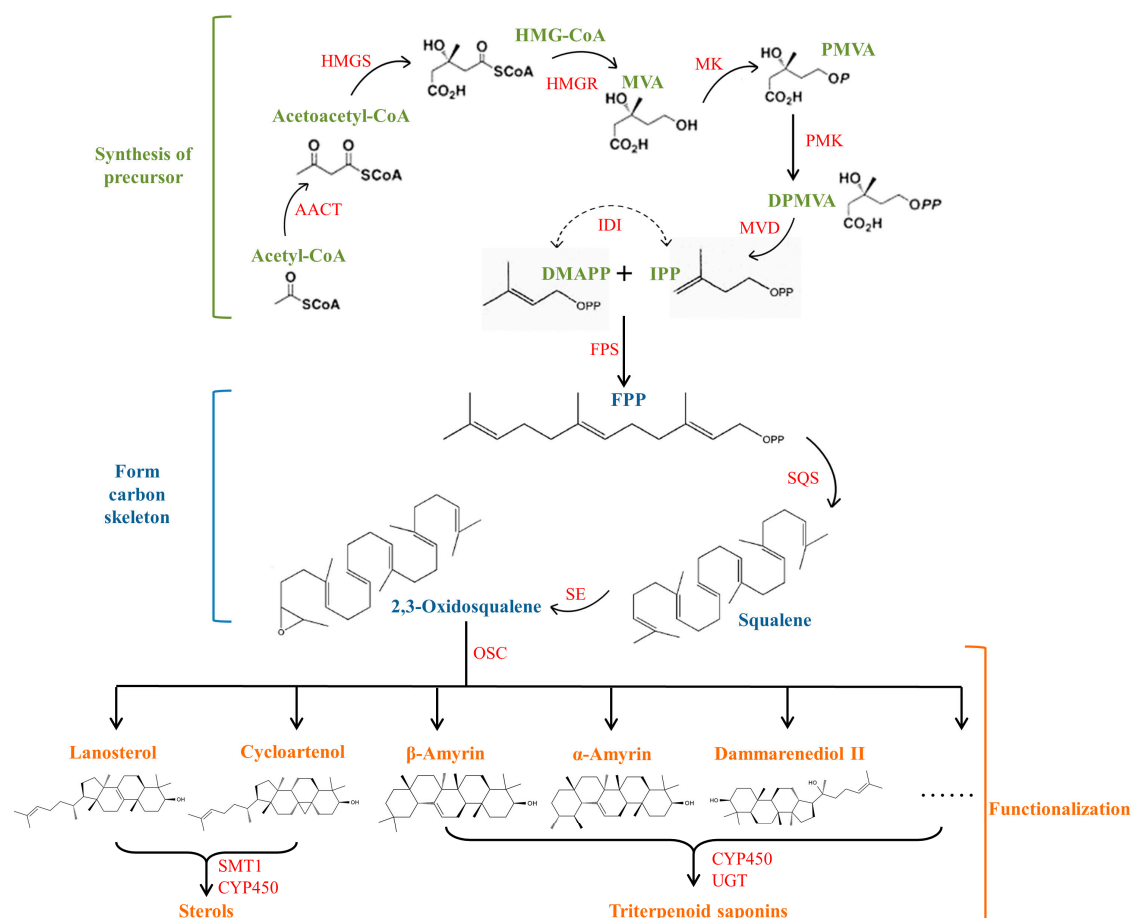
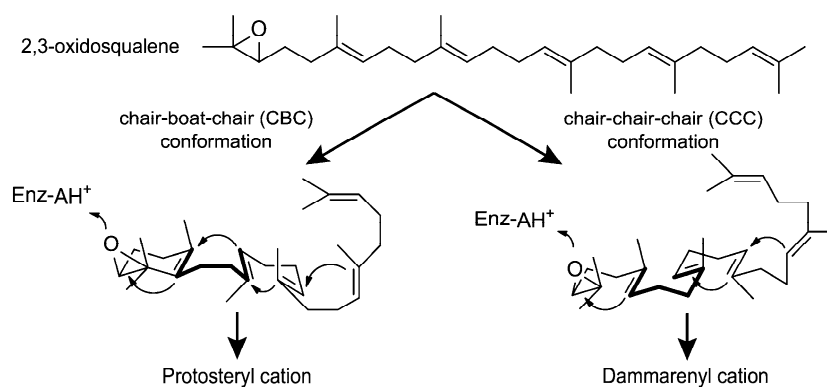
a**b**

Figure S1 Synthesis pathway of triterpenoids. **(a)** Triterpenoids are synthesized by the isoprenoid pathway, which can be generally divided into three stages: precursor synthesis, carbon ring skeleton formation and various complex functional reactions; **(b)** Conformational arrangement of 2,3-oxidosqualene [1].

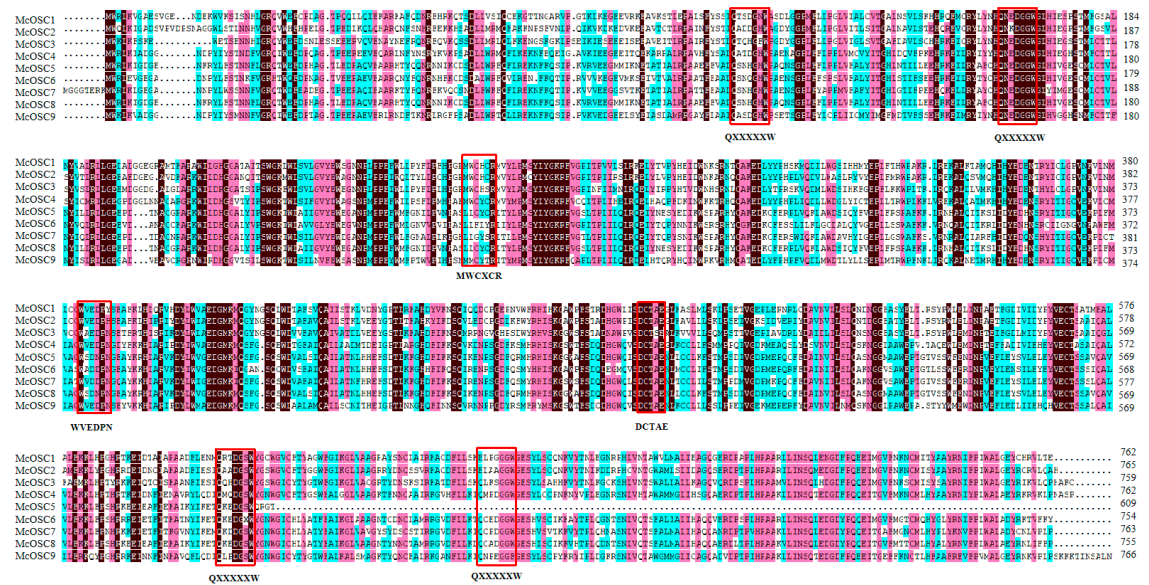


Figure S2 Conserved domain identification of nine McOSC genes. The red boxes in the figure are the conserved sequences of the OSCs gene family DCTAE, QXXXXXXW, WVDPN and MWCYCR.

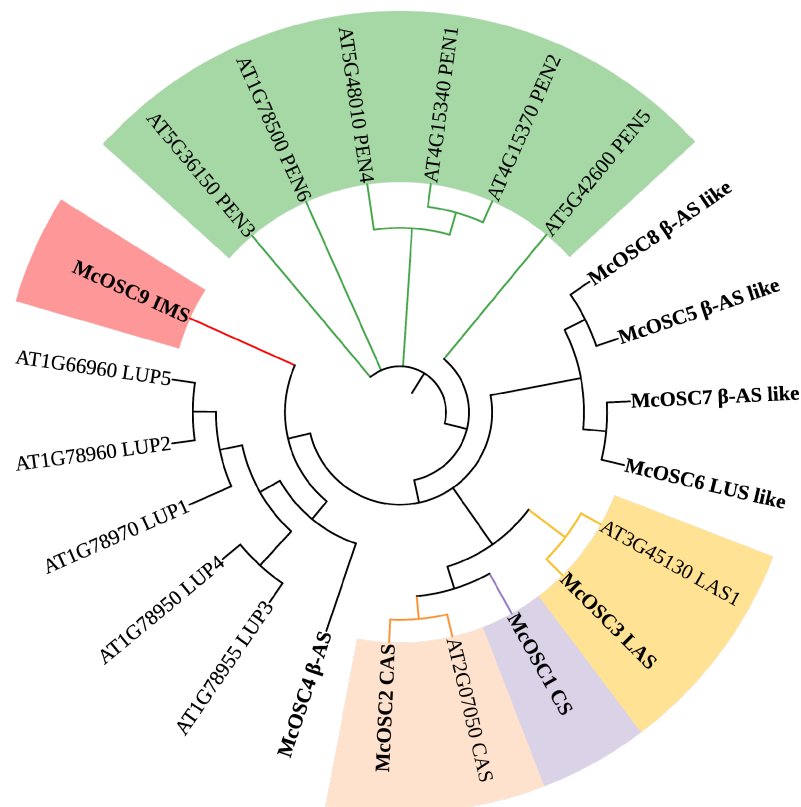


Figure S3 Phylogenetic analysis of nine McOSC genes with AtOSC genes from *Arabidopsis thaliana*. Genes in bold is OSCs from bitter melon. β -AS: β -Amyrin synthase; LUP: Lupeol

synthase; CAS: Cycloartenol synthase; CS: Cucurbitadienol synthase; IMS: isomultiflorenol synthase; PEN: pentacyclic triterpene synthase.

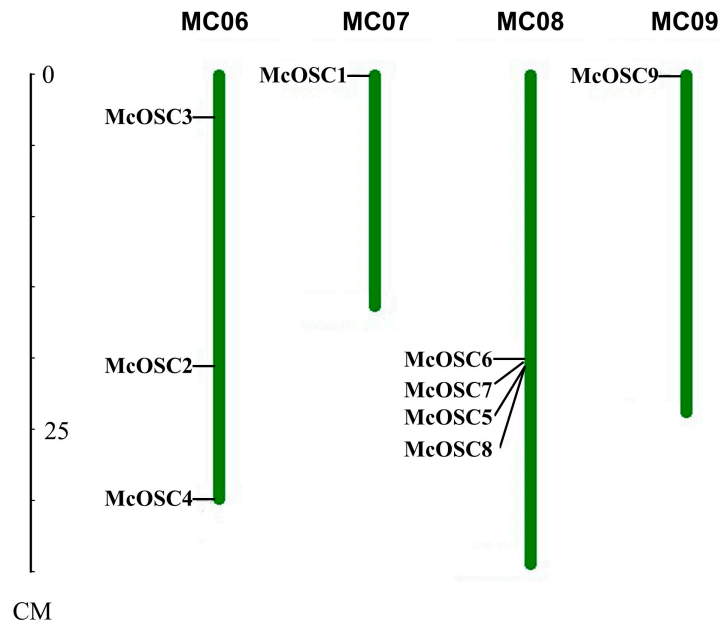


Figure S4 Chromosome location analysis of nine McOSC genes in bitter melon. The chromosome is shown in green, and the McOSC genes are marked on a line at the approximate position.

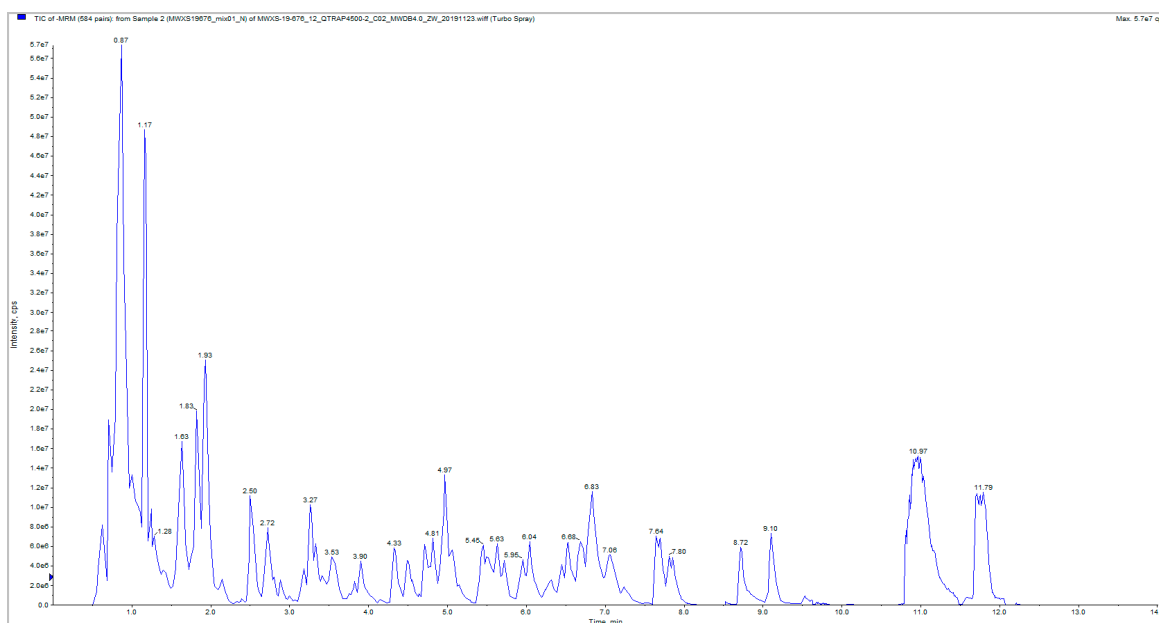
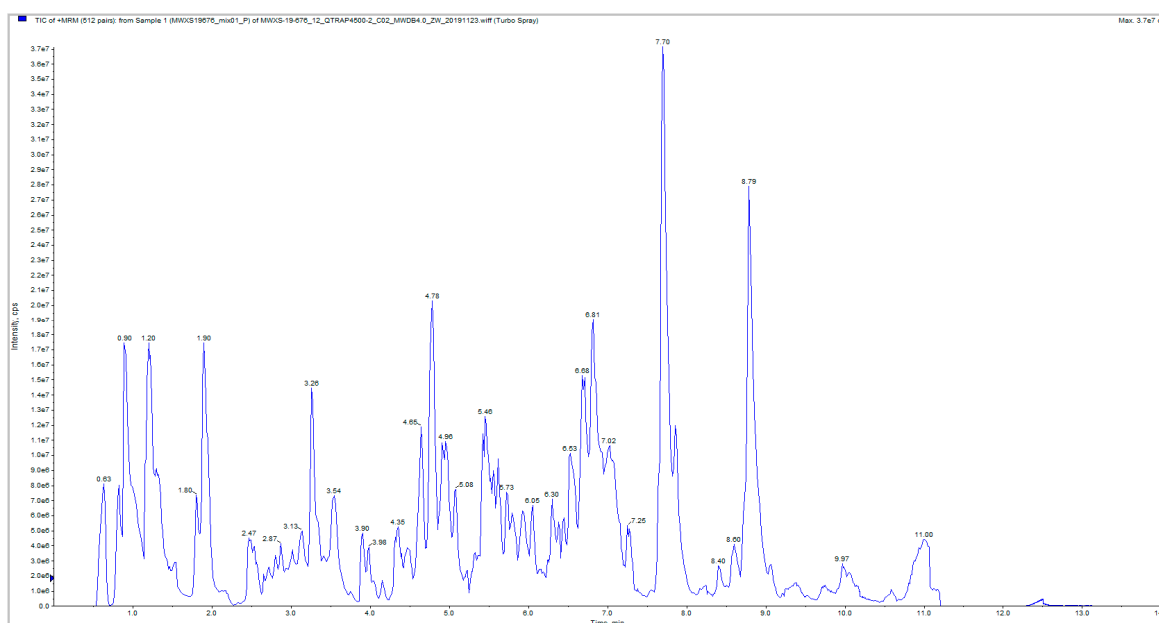
a**b**

Figure S5 Total ion flow diagram by mixed phase mass spectrometry (**a**: positive ion mode, **b**: negative ion mode). The abscissa is the Retention time (Rt) of metabolite detection, and the ordinate is the ion flow intensity of ion detection (intensity unit: CPS, count per second).

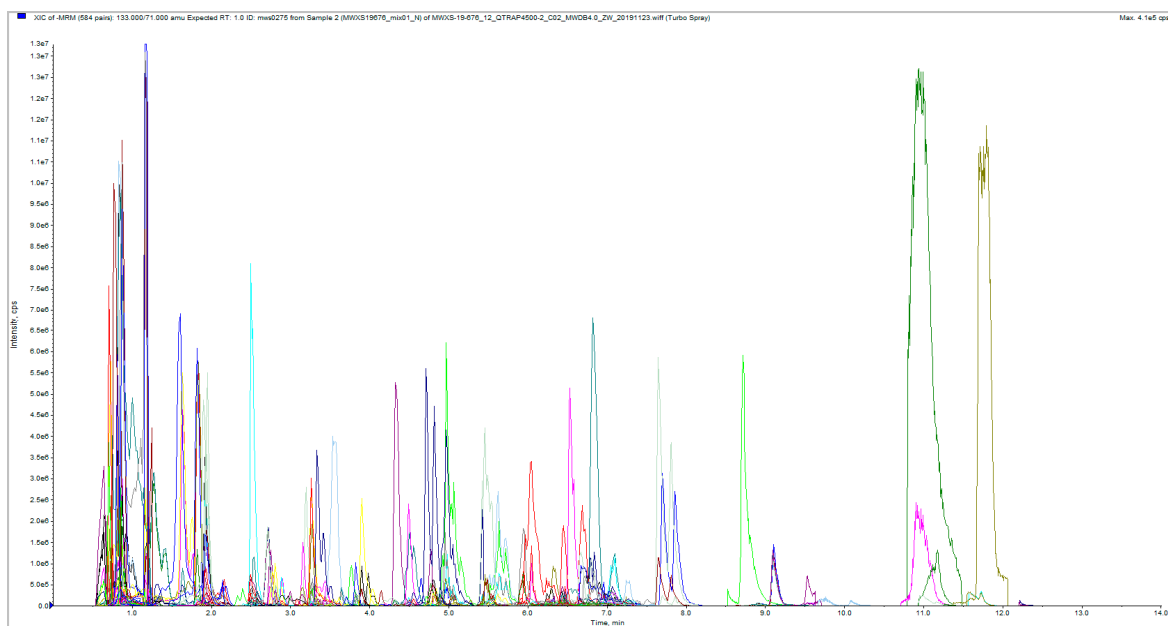
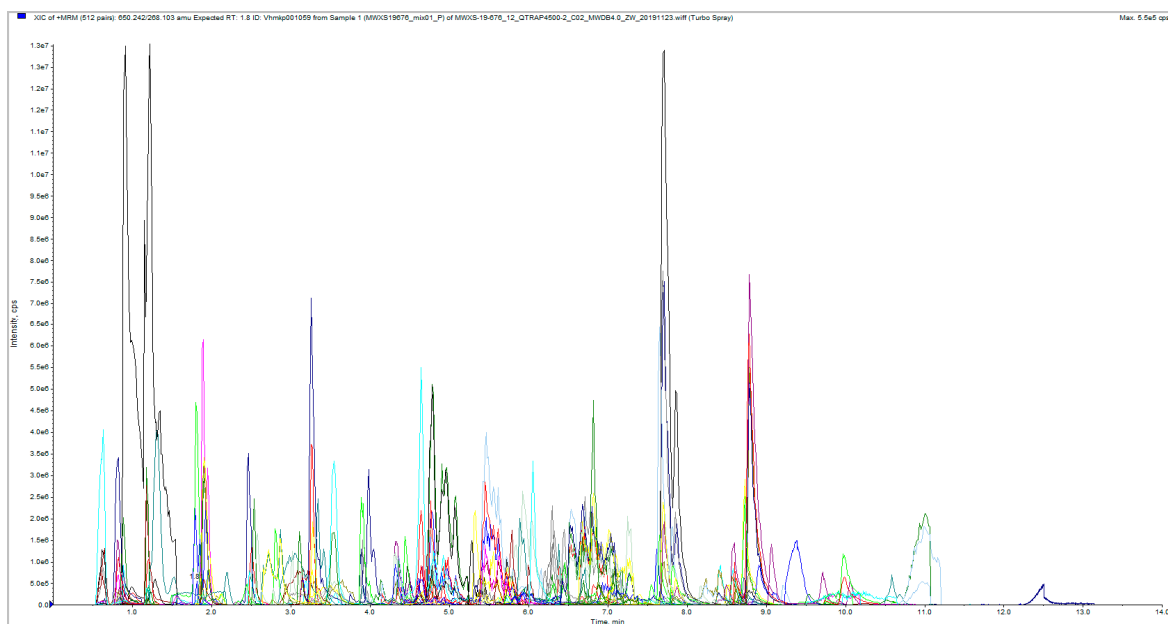
a**b**

Figure S6 Multi-peak map of MRM metabolite detection (**a**: positive ion mode, **b**: negative ion mode). The abscissa is the Retention time (Rt) of metabolite detection, and the ordinate is the ion flow intensity of ion detection (intensity unit: CPS, count per second).

Table S1. Primers used for qPCR analysis

Abbreviation	Primer sequence (5'-3')	Description
<i>McOSC1</i>	AACTGGGCTTCGGATCTTGG	<i>M. charantia</i> <i>McOSC1</i> primer, forward
<i>McOSC1</i>	CAAACATGGTGCTTGGGCTC	<i>M. charantia</i> <i>McOSC1</i> primer, reverse
<i>McOSC2</i>	TGCTGTTCAAGCCATCCTGT	<i>M. charantia</i> <i>McOSC2</i> primer, forward
<i>McOSC2</i>	ATGGGCCATCCATGATCTGC	<i>M. charantia</i> <i>McOSC2</i> primer, reverse
<i>McOSC3</i>	TGCATCCAGGAAGAATGTGGT	<i>M. charantia</i> <i>McOSC3</i> primer, forward
<i>McOSC3</i>	GGAACGTGGGGTGTACAAGT	<i>M. charantia</i> <i>McOSC3</i> primer, reverse
<i>McOSC4</i>	TATGATTGGGCCGGAAGCAA	<i>M. charantia</i> <i>McOSC4</i> primer, forward
<i>McOSC4</i>	GGCGTGAAGTTCCTGTCTCA	<i>M. charantia</i> <i>McOSC4</i> primer, reverse
<i>McOSC5</i>	GGCGTCAAATGGTGAATGG	<i>M. charantia</i> <i>McOSC5</i> primer, forward
<i>McOSC5</i>	TCTTCCTTCCTGTGGCTTGG	<i>M. charantia</i> <i>McOSC5</i> primer, reverse
<i>McOSC6</i>	CGATTGTGGGGCCAATCAC	<i>M. charantia</i> <i>McOSC6</i> primer, forward
<i>McOSC6</i>	AAGAGGGGTTCTCCACGTA	<i>M. charantia</i> <i>McOSC6</i> primer, reverse
<i>McOSC7</i>	ACATTACAGGACATCTTGCACT	<i>M. charantia</i> <i>McOSC7</i> primer, forward
<i>McOSC7</i>	ACAAGCATCGGTATCGGGTT	<i>M. charantia</i> <i>McOSC7</i> primer, reverse
<i>McOSC8</i>	CTCAACAGGCAGAGAGGGAC	<i>M. charantia</i> <i>McOSC8</i> primer, forward
<i>McOSC8</i>	CGATACTCTGCAAGTGCCCA	<i>M. charantia</i> <i>McOSC8</i> primer, reverse
<i>McOSC9</i>	CTCTCGCCATGCAAGCTCTA	<i>M. charantia</i> <i>McOSC9</i> primer, forward
<i>McOSC9</i>	TGCCATCCATGGTCACAGTC	<i>M. charantia</i> <i>McOSC9</i> primer, reverse
<i>β-Tub</i>	CATCTTCCACCTTTACACCCTG	<i>M. charantia</i> <i>β-Tubulin</i> primer, forward
<i>β-Tub</i>	CCATGTATCAATCAAACACTCC	<i>M. charantia</i> <i>β-Tubulin</i> primer, reverse

Table S2. Summary information of 9 OSC genes in bitter gourd.

Gene name	Accession	Types of OSCs	Protein ID	Gene ID	MW(Da)	PI	GRAVY	NO. of amino acids
<i>McOSC1</i>	XM_022293250.1	CS	XP_022148942.1	LOC111017486	187766.25	4.91	0.739	2289
<i>McOSC2</i>	XM_022282110.1	CAS	XP_022137802.1	LOC111009148	188994.65	4.94	0.657	2298
<i>McOSC3</i>	XM_022279428.1	LAS	XP_022135120.1	LOC111007175	186811.53	4.93	0.749	2280
<i>McOSC4</i>	XM_022299947.1	β -AS	XP_022155639.1	LOC111022720	187912.45	4.91	0.753	2289
<i>McOSC5</i>	XM_022294609.1	β -AS-like	XP_022150301.1	LOC111018501	150619.18	4.98	0.747	1830
<i>McOSC6</i>	XM_022303149.1	LUP-like	XP_022158841.1	LOC111025305	186001.19	4.94	0.736	2265
<i>McOSC7</i>	XM_022303150.1	β -AS-like	XP_022158842.1	LOC111025306	189132.05	4.93	0.743	2292
<i>McOSC8</i>	XM_022294610.1	β -AS-like	XP_022150302.1	LOC111018501	187365.93	4.94	0.734	2268
<i>McOSC9</i>	XM_022301789.1	IMS	XP_022157481.1	LOC111024177	190518.75	4.90	0.803	2301

Notes: MW means relative molecular mass; PI means theoretical isoelectric point; GRAVY means grand average of hydropathicity. The types of OSCs in the table are: β -AS: β -Amyrin synthase; LUP: Lupeol synthase; CAS: Cycloartenol synthase; LAS, lanosterol synthase; CS: Cucurbitadienol synthase; IMS: isomultiflorenol synthase.

Table S3. Ten terpene metabolites detected in transgenic hairy roots

Compounds	Formula	Class	CAS	Fold Change (<i>GFP</i> vs. <i>McOSC7-OE</i>)
Cucurbitacin D	C30H44O7	Triterpene	3877-86-9	1.43053
Camaldulenic acid	C30H46O4	Triterpene	71850-15-2	2.89626
Enoxolone	C30H46O4	Triterpene	471-53-4	1.82195
Quinovic acid	C30H46O5	Triterpene	465-74-7	4.30638
11-Carbonyl-20 β -hydroxycucurbitadienol	C30H48O3	Triterpene	-	0.78697
24,30-Dihydroxy-12(13)-enolupinol	C30H48O3	Triterpene	-	1.16759
2-Hydroxyoleanolic acid	C30H48O4	Triterpene	26707-60-8	1.47516
3,24-Dihydroxy-17,21-semiacetal-12(13) oleanolic fruit	C30H48O4	Triterpene	-	1.47508
Maslinic acid	C30H48O4	Triterpene	4373-41-5	1.45899
Deacetoxycucurbitacin B 3-O-glucoside	C36H54O11	Triterpene Saponin	-	1.15847

Notes: Fold Change represents the fold change of the metabolites of transgenic hairy roots compared with the control.