

Supplementary Data

Supplementary Tables

Table S1. Length of the five Formosan gum terpene synthase mRNAs. ORF: open reading frame; UTR: untranslated region. ORFs and UTRs are stated as nucleotides (nt).

	mRNA			Protein		
	ORF (nt)	5'UTR (nt)	3'UTR (nt)	Residue number	Theoretical <i>pI</i>	Theoretical Mw (Da)
<i>LfTPS01</i>	1680	149	134	559	5.06	64533.60
<i>LfTPS02</i>	1668	57	399	555	5.19	64487.70
<i>LfTPS03</i>	1686	73	200	561	5.10	64543.97
<i>LfTPS04</i>	1722	27	233	573	5.50	66670.35
<i>LfTPS05</i>	1725	60	181	574	5.09	65554.04

Table S2. Protein sequence identities and similarities (identities/similarities) between five *LfTPS*s and reference terpene synthases from different species.

Identity/similarity (%)	<i>LfTPS01</i>	<i>LfTPS02</i>	<i>LfTPS03</i>	<i>LfTPS04</i>	<i>LfTPS05</i>
<i>LfTPS01</i>	-	66/82	53/72	69/83	32/54
<i>LfTPS02</i>	-	-	51/71	61/78	31/54
<i>LfTPS03</i>	-	-	-	50/69	35/55
<i>LfTPS04</i>	-	-	-	-	32/53
<i>LfTPS05</i>	-	-	-	-	-

<i>VvGer</i>	67/83	65/81	52/71	60/80	31/51
<i>GhCad</i>	51/70	53/74	45/67	48/70	26/47
<i>ZmFar</i>	36/55	37/56	36/54	36/54	24/40
<i>AtOci</i>	26/48	26/49	31/51	26/51	28/49
<i>SaLin</i>	30/51	29/51	31/52	28/51	30/52
<i>AtGer</i>	12/25	12/25	11/26	12/26	13/25
<i>PtKau</i>	17/31	16/30	15/29	15/30	16/30
<i>AgHum</i>	26/46	25/47	26/46	25/45	22/41
<i>SmCar</i>	6/18	5/17	7/18	5/17	4/15

Table S3. Terpene synthases from different species used in phylogenetic analysis.

Species	Name	Accession No.	Products
<i>Arabidopsis thaliana</i>	AtGer	At1g61120	(E,E)-Geranyllinalool
	AtLin	At1g61680	S-(+)-Linalool
	AtOciMyr	At2g24210	(E)- β -Ocimene, Myrcene
	AtOci	At3g25810	(E)- β -Ocimene
<i>Abies grandis</i>	AgBis	3SDV_A	α -Bisabolene
	AgHum	AC05728	γ -Humulene
	AgMyr	AAB71084	Myrcene
	AgBis	3SDV_A	α -Bisabolene
<i>Vitis vinifera</i>	VvBer	ADR74195	(E)- α -Bergamotene
	VvCar	ADR74194	(E)- β -Caryophyllene
	VvGer	AAS66357	(-)-Germacrene D
	VvNer	ADR74219	(E)-Nerolidol, (E,E)-Geranyl linalool
<i>Populus trichocarpa</i>	PtKau	XP_002311286	<i>ent</i> -Kaurene
	PtFar	AAO61226	α -Farnesene
	PtPin	AAO61225	(-)- α -Pinene
<i>Physcomitrella patens</i>	PpKau	BAF61135	<i>ent</i> -Kaurene
<i>Sorghum bicolor</i>	SbZin	C5YHH7	β -Sesquiphellandrene

	SbSes	C5YHI2	Zingiberene
<i>Selaginella</i>	SmCar	J9QS25	(E)-2- <i>epi</i> - β -Caryophyllene
<i>moellendorffii</i>	SmNer	D8RNZ9	(3S,6E)-Nerolidol
<i>Taxus brevifolia</i>	TbTax	Q41594	Taxadiene
<i>Picea abies</i>	PaLon	Q675L0	Longifolene
	PaLin	AAS47693	Linalool
<i>Picea sitchensis</i>	PsCar	ADZ45511	(+)-3-Carene
<i>Pinus sylvestris</i>	PsFar	ADH29869	β -Farnesene
	PsCar	ABV44452	Caryophyllene, Humulene
<i>Gossypium hirsutum</i>	GhCad	AAC12784	(+)- δ -Cadinene
<i>Solanum lycopersicum</i>	SlCar	ADD96698	Caryophyllene, α -Humulene
<i>Zea mays</i>	ZmFar	Q2NM15	(E)- β -Farnesene
<i>Santalum album</i>	SaLin	B5A434	(+)- α -Terpineol, (-)-Limonene, (E)-Geraniol, Linalool
<i>Citrus limon</i>	ClLim	AAM53946	(+)-Limonene

Table S4 ^{13}C -nuclear magnetic resonance spectroscopy (^{13}C -NMR) (CDCl_3) signals of the compounds.

No.	Compound	Reference Compound	^{13}C -NMR (CDCl_3) signals ^a (ppm)
1	longicyclene	-	-
2	α -copaene	α -copaene	143.26 (143.73), 115.69 (116.30), 53.63 (54.50), 44.92 (45.06), 44.57 (44.63), 39.55 (39.58), 37.34 (37.33), 36.54 (36.51), 32.32 (32.33), 30.23 (30.20), 23.02 (23.00), 21.96 (21.93), 19.90 (20.00), 19.64 (19.72), 19.27 (19.30)
3	trans- β -caryophyllene	trans- β -caryophyllene	154.59 (154.47), 134.88 (134.89), 125.16 (124.96), 112.18 (112.12), 53.63 (53.96), 48.35 (48.32), 40.71 (40.76), 40.38 (40.40), 35.18 (35.10), 32.95 (32.96), 30.15 (30.17), 29.53 (29.50), 28.60 (28.62), 16.33 (16.36)
4	aromadendrene	aromadendrene	154.25 (154.17), 105.95 (106.05), 53.63 (54.06), 43.99 (44.00), 39.40 (39.44), 35.55 (35.58), 34.43 (34.46),

			29.75 (29.88), 29.42 (29.43), 28.83 (28.83), 27.85 (27.82), 25.10 (25.13), 20.03 (20.09), 17.52 (17.41), 15.80 (15.89)
5	α -humulene	α -humulene	140.55 (141.11), 138.15 (138.83), 132.91 (132.96), 128.40 (128.06), 126.54 (126.54), 125.40 (125.38), 42.38 (42.42), 40.59 (40.61), 40.07 (40.01), 37.41 (37.38), 27.28 (27.31), 23.65 (23.62), 17.93 (17.98), 15.17 (15.12)
6	germacrene D	germacrene D	148.84 (148.94), 136.29 (136.37), 133.30 (133.66), 133.19 (133.22), 129.75 (129.87), 109.27 (109.37), 53.16 (53.27), 40.91 (41.09), 34.91 (34.93), 33.01 (33.06), 29.53 (29.68), 26.84 (26.88), 21.00 (21.02), 19.64 (19.60), 15.96 (16.01)
7	α -selinene	β -selinene	151.28 (150.72), 150.34 (150.42), 108.81 (108.77), 105.95 (105.90), 50.15 (50.03), 46.67 (46.15), 42.14 (42.24), 41.33 (41.38), 37.21 (37.21), 36.17 (36.08), 29.75 (29.84), 27.09 (27.13), 23.76 (23.82), 21.04 (21.04), 16.49 (16.49)
8	α -muurolene	γ -muurolene	154.25 (154.32), 133.87 (133.86), 124.62 (124.57), 105.95 (106.51), 44.80 (44.80), 43.63 (43.54), 39.81 (39.75), 31.58 (31.62), 30.90 (30.86), 26.76 (26.71), 25.88 (25.87), 25.38 (25.41), 23.94 (23.89), 21.77 (21.72), 15.47 (15.49)
9	δ -cadinene	δ -cadinene	133.87 (133.93), 130.26 (130.35), 125.16 (125.05), 124.05 (124.05), 45.73 (45.75), 39.91 (39.88), 32.71 (32.69), 32.32 (32.32), 27.09 (27.16), 26.92 (27.02), 23.65 (23.67), 21.96 (21.89), 21.61 (21.61), 18.66 (18.57),

10	β -elemol	β -elemol	15.96 (15.90) 150.34 (150.58), 147.85 (147.97), 112.49 (112.42), 110.05 (110.04), 71.89 (71.85), 53.16 (52.94), 49.37 (49.45), 40.25 (40.18), 39.91 (39.94), 28.74 (28.77), 27.41 (27.43), 27.28 (27.20), 25.10 (25.14), 22.78 (22.76), 16.81 (16.75)
11	nerolidol	-	-
12	guaiol	-	-
13	unknown	-	-
14	γ -eudesmol	β -eudesmol	134.99 (134.97), 125.87 (125.76), 74.84 (74.85), 44.35 (44.24), 39.40 (39.41), 38.08 (38.12), 34.43 (34.37), 32.71 (32.70), 29.53 (29.59), 27.85 (27.80), 26.03 (26.00), 25.33 (25.32), 22.46 (22.48), 19.45 (19.53), 18.85 (18.90)
15	T-cadinol	-	-
16	α -cadinol	α -cadinol	134.88 (134.90), 122.38 (122.32), 72.81 (72.88), 50.15 (50.04), 46.83 (46.75), 42.14 (42.23), 39.91 (39.88), 30.90 (30.95), 26.03 (26.00), 23.28 (23.31), 22.64 (22.69), 21.96 (21.99), 21.48 (21.52), 20.69 (20.76), 15.17 (15.15)

a, The signal which most close to the reference is listed. Values in parentheses are ^{13}C -NMR spectra of reference compounds according to Kubeczka and Formáček 2002.

Supplementary Figures

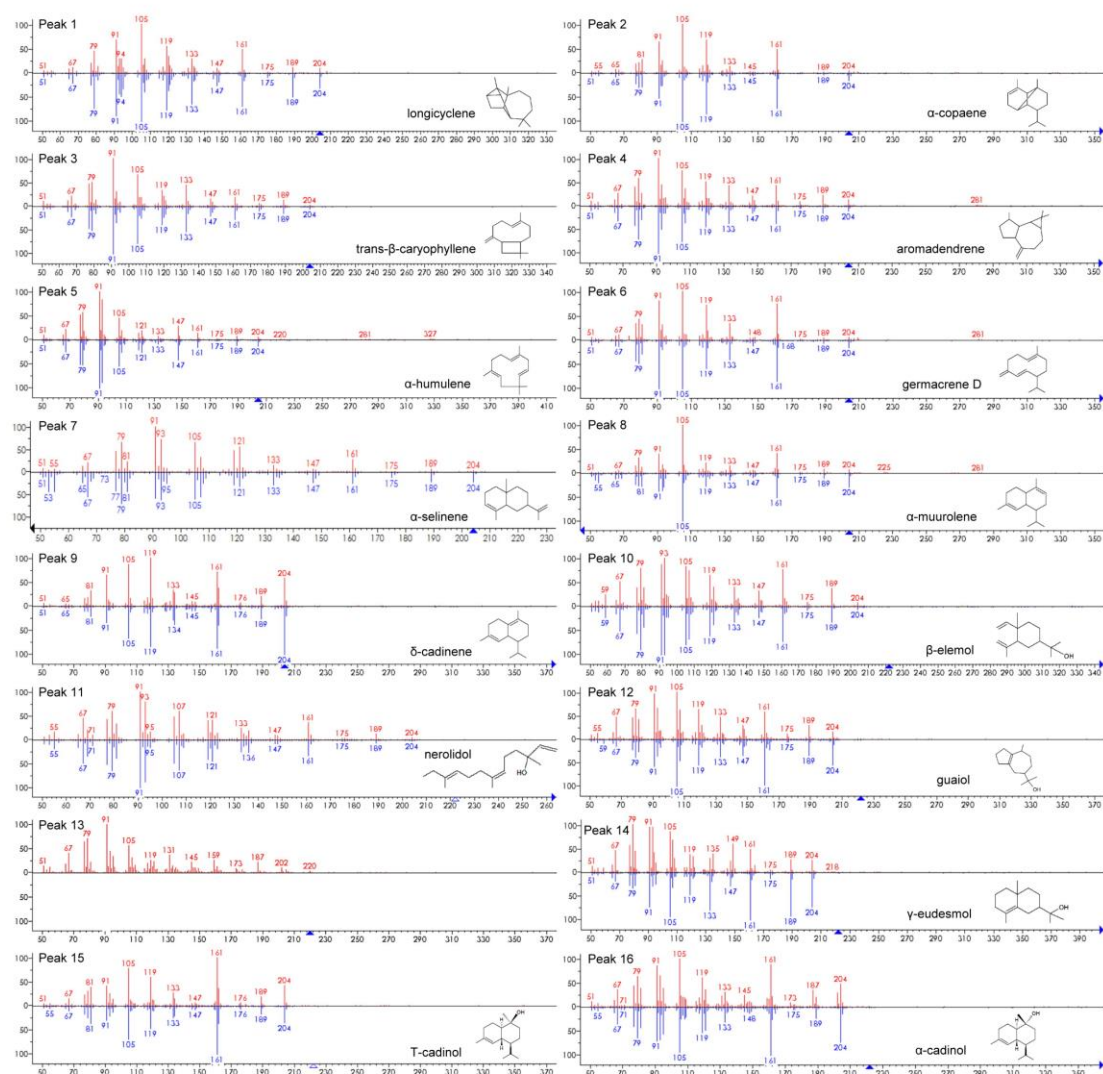


Fig. S1 Mass spectrophotometer (MS) plot of the 16 sesquiterpene compound listed in Table 1. The upper part in each plot represent MS in enzyme reaction and lower part represent the best hit reference MS in Wiley/NBS Mass Spectral library.

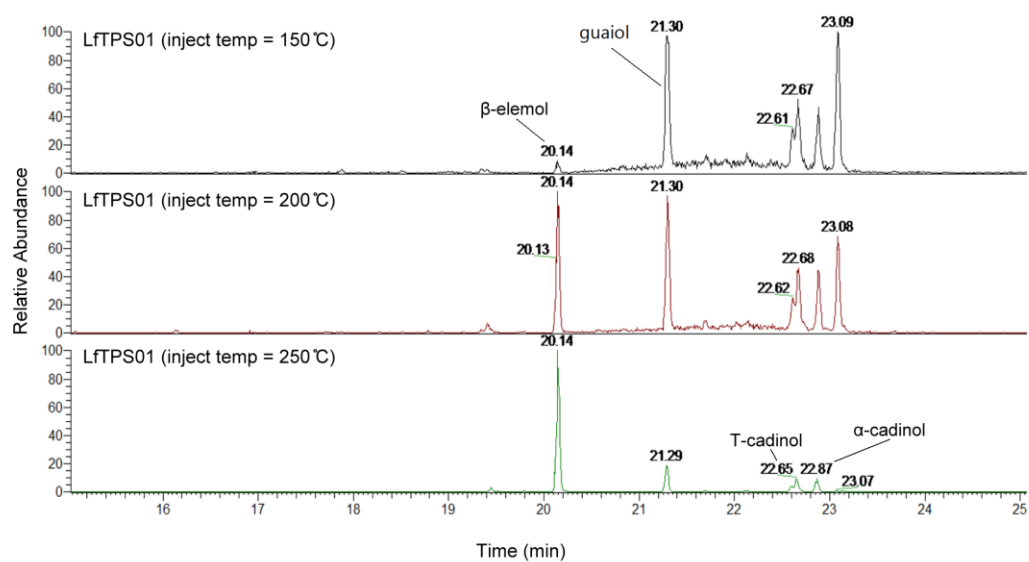


Fig. S2 Gas chromatography (GC) of LfTPS01 enzyme reaction product with inject temperature set to 150°C, 200°C and 250°C. The most intense ion detected in the spectrum was set to 204.

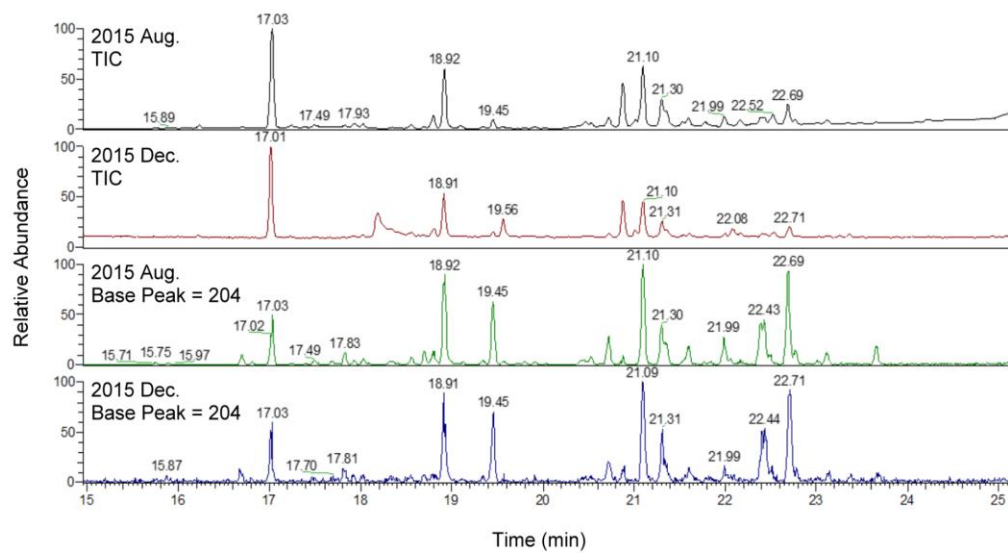


Fig. S3 Gas chromatography (GC) of Formosan gum essential oil. TIC: total ion current chromatogram; Base peak = 204: the most intense ion detected in the spectrum was set to 204.