Creation of new germplasm resources, development of SSR molecular markers and screen of monoterpene synthase in thyme

Meiyu Sun
IB CAS: Institute of Botany Chinese Academy of Sciences
Li Zhu
IB CAS: Institute of Botany Chinese Academy of Sciences
Yanan Zhang
IB CAS: Institute of Botany Chinese Academy of Sciences
Ningning Liu
IB CAS: Institute of Botany Chinese Academy of Sciences
Jinzheng Zhang
IB CAS: Institute of Botany Chinese Academy of Sciences
Hui Li
IB CAS: Institute of Botany Chinese Academy of Sciences
Hongtong Bai
IB CAS: Institute of Botany Chinese Academy of Sciences
Lei Shi (shilei_67@126.com)
IB CAS: Institute of Botany Chinese Academy of Sciences https://orcid.org/0000-0001-8918-7470

Research Article

Keywords: Thyme, Essential oil, Hybrid, SSR molecular markers, Monoterpene synthase

Posted Date: June 30th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1643537/v1

Version of Record: A version of this preprint was published at BMC Plant Biology on January 6th, 2023. See the published version at https://doi.org/10.1186/s12870-022-04029-2.

Abstract

Background

Thymus-derived essential oil (EO) and their compositions have numerous applications in medicine, food and cosmetics, while also displaying antibacterial, antifungal, and antiviral properties. In order to obtain EOs different terpene compositions of thyme, we want to use hybridization to create new germplasm resources of thyme.

Results

The phenotypic of three Chinese wild thymes and seven European thymes, including EO yield (extraction rate), glandular trichomes density, EO compositions content, plant type and fertility had been evaluated and analysed. Two F_1 populations were constructed with EOs terpene compositions content and yield as the main breeding goal, and refer to other traits to select suitable parental design hybrid combinations. At the same time, simple sequence repeat (SSR) primers were developed based on *T. quinquecostatus* whole-genome sequencing (date unpublished) to authenticate of F_1 hybrid progenies. The primers of 300 pairs were selected from the designed primers, and polymerase chain reaction (PCR) amplification was carried out in the parents of the two populations of in *T. longicaulis* (Tl) × *T. vulgaris* 'Fragrantissimus' (Tvf) and *T. vulgaris* 'Elsbeth' (Tve) × *T. quinquecostatus* (Tq). Based on the chemotype of the parents and F_1 progenies, we screened two γ-terpinene synthase, one α-terpineol synthase, and one geraniol synthase, and performed real-time quantitative PCR (qRT-PCR) of them in the parents and some F_1 progenies.

Conclusions

We use hybridization to create new germplasm resources of thyme, develop SSR molecular markers from *T. quinquecostatus*, and screen of monoterpene synthase in thyme. The above results laid the foundation for the creation of new germplasm resources, construction of genetic linkage map, location of quantitative trait locus (QTL) and insight into the mechanism of monoterpenoids biosynthesis in thymes.

Background

The genus *Thymus*, belonging to the Lamiaceae family, are commonly used for food, cosmetic, and medicinal purposes [1–3]. *Thymus* are important medicinal plants that have been used in traditional medicine for thousands of years in Mediterranean basin [4]. They are highly appreciated thanks for the wide spectrum of pharmacological properties of their essential oils (EOs). Anti-rheumatic, anti-septic, anti-spasmodic, anti-microbial, anti-inflammatory, carminative, diuretic, and expectorant activities have been investigated [5–8]. The main chemical classes of the EOs compositions in *Thymus* species are terpenoids, terpene alcohols, phenolic derivatives, ketones, aldehydes, ethers, and esters [9]. Generally these oils consist of oxygenated monoterpenoids (e.g. thymol, carvacrol, γ-terpinene, *p*-cymene, 1,8-cineole, linalool, α-terpineol, geraniol and borneol), sesquiterpenoids (e.g. β-caryophyllene) and oxygenated sesquiterpenoids (e.g. caryophyllene oxide) [4, 8–9]. Depending on the characteristics of the EOs profiles and main compositions, the occurrence of different chemotypes have been described in several species of *Thymus* plants [2, 9–11], including *T. vulgaris*.

At present, in the breeding research of thyme, there are a lot of researches work done, especially in European countries. Many scientific research institutions and companies have participated in, and many excellent industrial varieties and horticultural varieties have also been cultivated. In the 4th International Symposium on Breeding of Medicinal Aromatic Plants, it was mentioned that in order to optimize the OE terpenoid compositions and yield (extraction rate) performance of *Thymus* (*T. vulgaris*), a breeding program was carried out by crossing male sterile and male fertile clones [12]. Fifty-six new hybrids which were tested by assessing homogeneity, dry weight, EO yield, winter frost tolerance, and seed production potential of the parents were obtained from 2000 to 2002. The most dominant hybrid is named 'Varico 3', its EO yield is 4.9%, and the chemical type is thymol type [12]. These excellent varieties have been put on the market. The clone T-12 with higher phenol content among 10 clone lines of *T. vulgaris* were selected [13]. *Thymus* resources are relatively rich in European (Peter, 2012), and are 15 species, two varieties and one variant in China [14]. Therefore, according to different breeding goals, Chinese wild thymes can be crossed with European thymes to cultivate a series of new varieties that can be used in different fields, thereby promoting the development planting industry of thymes.

With the rapid development of modern molecular technology, DNA-based molecular markers have become an important tool for cultivar identification, fingerprinting [15–16] and genetic diversity analysis [17–19]. To improve the industrial uses of medicinal and aromatic plants, breeders often select plants with a high genetic distance and a high EO content [16]. Variation phenotype could be very valuable for molecular breeding approaches such as the molecular marker-assisted selection (MAS) breeding which have been very helpful in elucidating genetic diversity. In the study of *T. daenensis* and *T. fallax*, the dendrogram of cluster analysis showed the clear separation of these two species from the other *Thymus*, revealing that *T. daenensis* shared some genetic similarity with *T. fallax* [20]. Previously, several researches have reported the use of randomly amplified polymorphic DNA (RAPD) markers to study the genetic diversity, phylogenetic relationship, and its combination with the analysis of EOs compositions in various *Thymus* species [11, 16, 21].

The biosynthesis pathway for phenolic monoterpenoids in thyme is derived from the mevalonate (MVA) and 2-C-methyl-D-erythritol-4-phosphate (MEP) pathways [22]. Isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) are canonically condensed head-to-tail by trans-prenyltransferases to generate geranyl diphosphate (GPP). Terpene synthase (TPS) catalyzes GPP to form the basic skeleton of monoterpenes (C10) and sesquiterpenes (C15), respectively. To date, TPSs have been characterized as multifunctional enzymes which have enough single amino acid substitutions to make an altered metabolic profile [23–25]. In fact, a wide range of TPSs and other terpene-modifying enzymes have been characterized [26]. In the *Thymus* genus, *TPS* genes

from *T. caespititius* [27–29], *T. vulgaris* [30–33], *T. serpyllum* [30], *T. albicans* [34], and *T. citriodorus* (unpublished, direct submission to GenBank) has been identified previously.

In our study, two F_1 populations have constructed by crossing male sterile lines with male fertile lines in thyme. Simple sequence repeat (SSR) primers were developed based on *T. quinquecostatus* whole-genome sequencing (Chinese wild thyme, data unpublished), and the authenticity of the F_1 individuals of two populations *T. longicaulis* (Tl) × *T. vulgaris* 'Fragrantissimus' (Tvf) and *T. vulgaris* 'Elsbeth' (Tve) × *T. quinquecostatus* (Tq) were verified using SSR, and F_1 progenies were all true hybrids. At the same time, two γ-terpinene synthase gene *Tq13G005250.1* and *Tq02G002290.1*, one geraniol synthase gene *Tq04G005190.1*, and one α-terpineol synthase gene *Tq03G001560.1* were verified by qRT-PCR. The results of this study will be very useful for the selection of new thyme varieties, MAS breeding, and will lay the foundation for further gene function verification of *TPSs*.

Results

Evaluation phenotypes of three Chinese wild thymes and seven European thymes, and construction of F_1 hybrid population

Thymus is a herbaceous perennial or sub-shrub with valuable medicinal and aromatic properties. The types of *Thymus* are erect-type and creeping-type, which display remarkable differences in morphology. The plant types of 10 different thyme varieties are shown in Fig. 1A (Table 1). Among them, *T. rotundifolius* (Tr), *T. vulgaris* 'Elsbeth' (Tve), *T. thracicus* (Tt), and *T. vulgaris* 'Fragrantissimus' (Tvf) are erect-type, and *T. serpyllum* 'Aureus' (Ts), *T. guberlinesis* 'lijin' (Tg), *T. longicaulis* (Tl), *T. quinquecostatus* (Tq), *T. quinquecostatus* var. *przewalskii* (Tqp), and *T. mongolicus* (Tm) are creeping-type. Thyme fertility is shown in Fig. 1B. Tr, Tve, Tg, Tt, Ts, and Tl have only stigma and no pollen (stamen), and they are male-sterile. Tq, Tqp, Tm, and Tvf have both stigma and pollen (stamen), and they are male-fertile (Table 1). The increased genetic diversity of the thymes could be partially attributed to the consistent introgression of Chinese wild thymes into the male-sterility and erect-type European thymes during long cultivation periods and partially to the adaptations of thymes to new environments during spreading.

Table 1
Phenotype survey results of three Chinese wild thymes and seven European thymes.

Species	Chemotype	Plant type	Oil yield (%)	Pollen	Seed	Sterility	Flowering phase
Thymus quinquecostatus	carvacrol	creeping	0.50	yes	yes	fertile	2020.4.20–2020.8.30
Thymus quinquecostatus var. *przewalskii*	carvacrol	creeping	0.50	yes	yes	fertile	2020.4.20–2020.8.30
Thymus mongolicus	thymol	creeping	0.40	yes	yes	fertile	2020.5.12–2020.8.30
Thymus rotundifolius	thymol	erect	0.95	no	no	sterile	2020.6.20–2020.7.30
Thymus vulgaris 'Elsbeth'	thymol	erect	0.75	no	no	sterile	2020.6.1–2020.6.20
Thymus guberlinesis 'lijin'	thymol	creeping	0.50	no	no	sterile	2020.5.6–2020.5.20
Thymus thracicus	thymol	erect	1.20	no	no	sterile	2020.5.8–2020.5.30
Thymus serpyllum 'Aureus'	thymol	creeping	0.52	no	no	sterile	2020.5.22–2020.6.15
Thymus longicaulis	geraniol	creeping	0.70	no	no	sterile	2020.4.20–2020.5.15
Thymus vulgaris 'Fragrantissimus'	α–terpineol	erect	1.25	yes	yes	fertile	2020.4.20–2020.5.10

Glandular trichomes are specialized hairs found on the surface of about 30% of all vascular plants and are responsible for a significant portion of a plant's secondary chemistry [35]. Terpenoids are stored and synthesized in glandular trichomes which is an organ that originates from epidermal cells of flowers, leaves, and stems [36]. Glandular trichomes are widely found in lavender, thyme, rosemary, oregano, basil, and other Lamiaceae, with two existing types: peltate and capitate trichomes. The regulation of these trichome-related genes may underlie the regulation of glandular trichomes density to increase the terpenoid content. The adaxial and abaxial planes of thyme leaves are shown in Fig. 1C, E. The most glandular trichomes per unit area (glandular trichomes density) of adaxial were Tve and Tr, followed by Tvf and Tt. The most glandular trichomes per unit area of abaxial (glandular trichomes density) were Tve and Tr. In total, The most glandular trichomes density of leaves are Tve, Tr, Tvf, Tt, and Ts. The EO yield (extraction rate) are shown in Fig. 1D, F (Supplementary Fig. S1). The highest oil yield are Tt (1.20%) and Tvf (1.25%), followed by Tr (0.95%), Tve (0.75%), and Tl (0.70%), respectively. The EO yield of Tq, Tqp, Tm, Tg, and Ts are between 0.40% and 0.50% (Table 1; Supplementary Fig. S1). There was a certain correlation between EO yield and glandular trichomes density (Fig. 1E, F).

The relative contents of EOs compositions of 10 different thymes are shown in Fig. 2A. In this study, the 20 main compositions (relative content > 0.3%) shared by 10 different thymes were counted and showed in Table 2. The most volatile compositions of Tq essential oil are *p*-cymene and carvacrol, which are account for 23.00% and 20.74%, respectively. The volatile composition with the most Tqp content is carvacrol, which is accounts for 48.37%. The volatile compositions with the most Tm content are thymol (38.57%) and *p*-cymene (16.40%). The volatile compositions with the most Tve content are thymol (35.43%), *p*-cymene

(18.42%), and γ-terpinene (13.96%). The volatile composition with the highest content of Tr, Tt, Tg, and Ts is thymol, and their contents are 36.02%, 41.04%, 26.26%, and 28.96%, respectively. The volatile compositions with the most Tl content are geraniol (28.54%) and geranyl acetate (34.81%). The volatile compositions with the most Tvf content are α-terpineol (30.84%) and α-terpineol acetate (45.46%).

Table 2
Relative contents of volatile terpenoid compositions in 10 thyme essential oils.

No.	Composition	RI Cal	RI Lit	Relative content(%)									
				Tq	Tqp	Tm	Tve	Tr	Tt	Tg	Ts	Tl	Tvf
1	γ-Terpinene	1059	1060	11.45 ± 0.66	10.77 ± 0.71	8.19 ± 0.61	13.96 ± 0.50	13.90 ± 0.51	12.61 ± 0.55	11.18 ± 1.02	10.59 ± 0.32	0.69 ± 0.01	0.48 ± 0.03
2	*p*-Cymene	1026	1025	23.00 ± 1.02	15.34 ± 0.52	16.40 ± 0.72	18.42 ± 1.01	18.08 ± 0.20	17.98 ± 0.97	16.50 ± 1.03	12.03 ± 0.54	-	0.62 ± 0.02
3	Thymol	1292	1291	2.12 ± 0.25	7.59 ± 0.02	38.57 ± 2.01	35.43 ± 3.02	36.02 ± 3.04	41.04 ± 4.02	26.26 ± 2.02	28.96 ± 2.03	0.57 ± 0.05	0.59 ± 0.06
4	Carvacrol	1304	1299	20.74 ± 1.02	48.37 ± 4.05	6.87 ± 0.56	2.27 ± 0.23	2.48 ± 0.22	6.45 ± 0.72	1.95 ± 0.05	3.78 ± 0.05	-	-
5	Geraniol	1259	1255	-	-	-	-	-	0.49 ± 0.02	-	0.69 ± 0.06	28.54 ± 1.05	-
6	Geraniol acetate	1388	1382	-	-	-	-	-	-	-	1.50 ± 0.12	34.81 ± 2.07	-
7	α-Terpineol	1193	1189	0.95 ± 0.02	-	1.11 ± 0.08	-	-	-	0.75 ± 0.06	0.50 ± 0.05	0.85 ± 0.03	30.84 ± 2.02
8	α-Terpineol acetate	1355	1350	-	-	-	-	-	-	-	-	1.17 ± 0.09	45.46 ± 5.05
9	Eucalyptol	1030	1032	5.78 ± 0.32	0.93 ± 0.05	7.02 ± 0.23	0.81 ± 0.06	0.84 ± 0.03	0.64 ± 0.07	1.71 ± 0.02	-	0.76 ± 0.07	0.67 ± 0.12
10	endo-Borneol	1165	1167	7.89 ± 0.66	1.37 ± 0.22	4.96 ± 0.44	0.55 ± 0.02	1.23 ± 0.02	0.56 ± 0.01	9.02 ± 0.54	4.13 ± 0.33	8.17 ± 0.55	1.26 ± 0.06
11	β-Caryophyllen	1421	1419	1.74 ± 0.24	1.56 ± 0.34	1.75 ± 0.03	3.37 ± 0.21	4.01 ± 0.24	3.11 ± 0.16	3.79 ± 0.22	3.07 ± 0.03	4.19 ± 0.02	2.60 ± 0.04
12	Caryophyllen oxide	1585	1581	1.14 ± 0.02	-	0.51 ± 0.01	0.47 ± 0.02	0.74 ± 0.05	-	0.93 ± 0.12	0.34 ± 0.01	0.48 ± 0.01	-
13	α-Thujene	926	929	0.91 ± 0.02	1.97 ± 0.01	1.21 ± 0.02	2.09 ± 0.26	1.27 ± 0.04	1.53 ± 0.05	1.07 ± 0.11	1.31 ± 0.02	-	-
14	Sabinene	972	974	0.81 ± 0.04	-	-	-	-	-	-	-	-	2.18 ± 0.03
15	1-Octen-3-ol	978	980	0.64 ± 0.14	1.12 ± 0.05	2.07 ± 0.33	1.20 ± 0.02	1.30 ± 0.14	0.92 ± 0.01	0.88 ± 0.04	1.39 ± 0.07	0.69 ± 0.03	-
16	β-Myrcene	991	991	0.96 ± 0.02	1.09 ± 0.04	1.15 ± 0.05	1.74 ± 0.13	1.37 ± 0.02	1.86 ± 0.15	0.51 ± 0.01	1.12 ± 0.01	0.38 ± 0.01	1.62 ± 0.25
17	D-Limonene	1028	1032	0.90 ± 0.05	-	0.70 ± 0.12	0.90 ± 0.01	0.79 ± 0.04	0.81 ± 0.06	1.09 ± 0.25	0.77 ± 0.03	-	6.04 ± 0.72

Continued Table 2

No.	Composition	RI Cal	RI Lit	Relative content（%）								
				Tq	Tqp	Tm	Tve	Tr	Tt	Tg	Ts	Tl
18	α-Terpinolen	1087	1088	1.65±0.13	2.10±0.18	1.58±0.04	3.24±0.32	-	2.76±0.27	0.66±0.08	2.83±0.12	-
19	Linalool	1099	1099	1.17±0.05	1.52±0.05	1.33±0.15	1.06±0.01	1.21±0.03	5.33±0.75	0.72±0.42	0.38±0.01	4.12±0.6
20	Camphene	949	952	2.60±0.56	-	0.92±0.02	-	-	-	1.84±0.08	1.04±0.04	2.15±0.1

(Notes: Tq, *T. quinquecostatus*; Tqp, *T. quinquecostatus* var. *przewalskii*; Tm, *T. mongolicus*; Tve, *T. vulgaris* 'Elsbeth'; Tr, *T. rotundifolius*; Tt, *T. thracicus*; Tg, *T. guberlinesis* 'lijin'; Ts, *T. serpyllum* 'Aureus'; Tl, *T. longicaulis*; Tvf, *T. vulgaris* 'Fragrantissimus'. RI Cal, calculated according to C7–C40; RI Lit, obtained by searching the mass spectrum database NIST14.0)

Based on the cluster analysis of the 20 kinds main common compositions in 10 different thymes are shown in Fig. 2B, all thymes are clustered into four categories. The one most abundant composition of Tq and Tqp is carvacrol, and their relative contents are respectively 20.74% and 48.37%, so its chemical type is carvacrol-type; Tm, Tve, Tr, Tt, Tg, and Ts are grouped together, and the most content of them is thymol, so its chemical type is thymol-type; Tl is divided to one group which is geraniol-type, Tvf is divided to other type which is α-terpineol-type, respectively. Principal component analysis (PCA) analysis was carried

out on the main compositions common to these 10 different thymes, as shown in Fig. 2C, D. It can be seen that among these 10 thymes, and they contribute to the volatile compositions of their EOs. The most compositions are thymol, carvacrol, *p*-cymene, and γ-terpinene. It can be seen in Fig. 2D that Tq and Tqp are distributed in the first quadrant, and their corresponding characteristic volatile substances are carvacrol, *p*-cymene, and γ-terpinene; Tvf and Tl are distributed in the third quadrant, which correspond to the characteristic volatile substances of are geraniol and α-terpineol; Tm, Tve, Tr, Tt, Tg, and Ts are distributed in the fourth quadrant, and the corresponding characteristic volatile substance is thymol, which can be used to verify cluster analysis the result.

Use male sterile (without pollen) thyme as the female parent, and male fertile (with pollen) thyme as the male parent, the EO compositions and yield were taken as the main breeding goals, and other traits were referenced. Different cross combinations were designed, and the F_1 hybrid populations of thyme were constructed by means of cross breeding. Finally, two hybrid populations were obtained: *T. longicaulis* × *T. vulgaris* 'Fragrantissimus' (Tl × Tvf, 14 lines) and *T. vulgaris* 'Elsbeth' × *T. quinquecostatus* (Tve× Tq, 11 lines) (Supplementary Table S1).

Ssr Development And Its Application In F1 Hybrid Progenies

Through the assembly and annotation of *T. quinquecostatus* via high-fidelity (HiFi) and chromatin conformation capture (Hi-C) technologies, its genome was revealed at the chromosome level, which contained 13 chromosomes at a total length of 528.66 megabases (Mb) (Data unpublished). The *T. quinquecostatus* genome was highly repetitive with a total of 373.28 Mb of repetitive sequences annotated, accounting for 70.61% of the genome (Data unpublished). Finally, a total of 239,400 tandem repeats were identified, accounting for 48.30 Mb (9.14%) of the genome. A total of 191,847 SSR loci were detected, 183,536 SSR loci used for primer design, accounting for 95.67%, and 8,311 SSR loci not used for primer design, accounting for 4.33%. Among the top ten Contig sequences with the largest distribution of SSR sites, Contig00377 has the largest number of SSR sites (4,774), and Contig00721 has the least number of SSR sites (2,089) (Supplementary Fig. S2A). Among these ten Contig sequences, the most frequent occurrence of SSR per Mb is Contig00808 (435), and the least frequent occurrence of SSR is Contig00630 (264) (Supplementary Fig. S2B). The results showed that there were seven dinucleotide repeat types, 12 trinucleotide repeat types and one tetranucleotide repeat type in the SSR sequence of thyme genome (Supplementary Fig. S2C). CT/AG had the highest proportion of dinucleotide repeats (23.70%), followed by TC/GA (20.50%), TA/TA (14.80%), AT/AT (12.00%), TG/CA (3.80%), GT/AC (3.60%), and GC/GC (0.40%). ATT/AAT accounted for the highest proportion of trinucleotide repeats (2.50%), followed by TTA/TAA (2.00%), TTC/GAA (1.60%), ATA/TAT (1.60%), AGA/TCT (1.50%), CTT/AAG (1.20%), GCC/GGC (0.60%), CCG/CGG (0.50%), GAG/CTC (0.50%), CGC/GCG (0.40%), ATC/GAT (0.40%), and GGA/TCC accounted for the least, which was 0.30%. Only AAAT/ATTT tetraponucleotide repeats accounted for 0.40% (Supplementary Fig. S2 C). The length of SSR sequence in *T. quinquecostatus* genome ranged from 18 bp to 87 bp, and 10 bp SSR locus was the most (61,865), accounting for 32.00% of all SSR loci, while 25 bp SSR locus was the least (1,356), accounting for 0.71% of all SSR loci. The overall trend of changes in SSR motif length was that the number of SSR gradually decreased with the increase of the length (Supplementary Fig. S2D).

The primers of 300 pairs were selected from the designed primers, and PCR amplification was carried out in the parents of the two populations of Tl × Tvf and Tve × Tq (Supplementary Table S2). Through the polymorphism analysis between the parents, the polymorphisms detected in the parents of the primers are 1–2 bands, after many repetitions, the primers with clear and stable bands are screened out (Fig. 3A, D). There are 18 pairs of SSR primers in the hybrid combination Tl × Tvf that show parental co-dominant (Supplementary Table S3). For example, primer TqSSR289 amplifies characteristic band 1 in the female parent T1, and amplifies characteristic band 2 in the male parent Tvf (Fig. 3A). Primer TqSSR292 amplified characteristic band 3 in the female parent T1, and amplified characteristic band 4 in the male parent Tvf (Fig. 3A). There are 23 pairs of SSR primers in the hybrid combination Tve × Tq that are co-dominant with the parents (Supplementary Table S4). For example, the primer TqSSR284 amplifies characteristic band 1 in the female parent Tve, and amplifies characteristic band 2 in the male parent (Fig. 3D), so as to be used for the identification and verification of the hybrid progenies of the population, accounting for 7.60% of the total number of primers. These co-dominant SSR primers were used to identify hybrids in these two populations. In the progeny amplification results, those with parental complementary bands and only paternal-specific bands are true hybrids, and the progeny with only maternal-specific bands are pseudo-hybrids or inbreds. After identification, the 14 F_1 lines of the hybrid combination Tl × Tvf are all true hybrids (Fig. 3B, C), and the 11 F_1 lines of Tve × Tq are all true hybrids (Fig. 3E, F).

Determination of volatile organic compositions (VOC) in the leaves of the population Tl × Tvf and Tve × Tq parents and progenies. In the population Tl × Tvf (Fig. 4A; Supplementary Table S5), the VOC with the most content in the female parent Tl are geraniol (22.75%) and geranyl acetate (41.75%), and the VOC with the most contents in the male parent Tvf are α-terpineol (11.76%) and α-terpineol acetate (61.37%); Among the 14 F_1 lines, the most abundant compositions of the progenies numbered F_1-1, F_1-2, F_1-8, and F_1-9 are thymol (the contents are 12.51%, 8.03%, 9.55%, and 8.39%, respectively), carvacrol (the contents are 13.93%, 9.72%, 12.17%, and 12.92%, respectively), *p*-cymene (the contents are 16.52%, 23.22%, 29.39%, and 23.01%, respectively), and γ-terpene (the contents are 14.16%, 9.50%, 9.02%, and 11.36%, respectively); The most abundant compositions of the progenies numbered F_1-3, F_1-4, F_1-5, F_1-6, F_1-10, F_1-11, F_1-13, and F_1-14 are geraniol (the contents are 18.17%, 22.75%, 23.79%, 28.98%, 28.84%, 26.91%, 25.44%, and 23.57%, respectively) and geranyl acetate (the contents are 15.76%, 29.59%, 21.89%, 23.41%, 15.05%, 23.60%, 22.50%, and 24.56%, respectively), so the compositions of these progenies are biased toward the female parent Tl; The compositions with the most content in the progeny of F_1-7 are geraniol (4.48%), geranyl acetate (27.75%), α-terpineol (3.00%), and α-terpineol acetate (28.03%), a good aggregation of the dominant compositions in two parents; The most abundant compositions in the progeny numbered F_1-12 is α-terpineol (9.07%) and α-terpineol acetate (57.44%), favoring the male parent Tvf. According to the 17 main chemical compositions, clustering of parents and 14 F_1 lines found that F_1-3, F_1-4, F_1-5, F_1-6, F_1-10, F_1-11, F_1-13, and F_1-14 chemical type is geraniol-type which clustered with the female parent Tl; F_1-12 and the male parent Tvf are clustered together, and the chemical type is α-terpineol-type; F_1-1, F_1-2, F_1-8, and F_1-9 are one chemical type, they are thymol and carvacrol polymerization-type; F_1-7 is geraniol and α-terpineol polymerization-type (Fig. 4B).

In the population Tve × Tq (Fig. 4C; Supplementary Table S6), the compositions with the high content of female parent Tve are thymol (35.43%), *p*-cymene (18.42%), and γ-terpinene (13.96%); The compositions with the high content of male parent Tq is carvacrol (20.74%), *p*-cymene (23.23%), and γ-terpinene

(11.45%). Among the 11 F_1 lines, the most abundant compositions of the progenies numbered F_1-3, F_1-4, F_1-5, F_1-8, and F_1-11 are thymol (the contents are 23.57%, 24.71%, 46.19%, 46.38%, and 33.26%, respectively), *p*-cymene (the contents are 37.42%, 27.03%, 23.16%, 18.56%, and 21.14%, respectively) and γ-terpinene (the contents are 17.14%, 9.03%, 8.55%, 12.58%, and 6.85%, respectively). The compositions of these F_1 lines are biased towards the female parent Tve, and the contents of F_1-5 and F_1-8 thymol are higher than the female parent Tve, accounting for 45.45% of the progeny plants; The most compositions contents numbered F_1-1, F_1-2, F_1-6, F_1-7, F_1-9, and F_1-10 are α-terpineol (the contents are 16.84%, 20.34%, 11.25%, 29.45%, 20.72%, and 23.98%, respectively), and α-terpineol acetate (the contents are 75.72%, 71.58%, 80.44%, 60.21%, 72.17%, and 67.11%, respectively), the sum of the proportions of these two compositions is more than 90% in these six F_1 lines, Belongs to the absolute dominant compositions, but these two compositions are not dominant in the parents, these two compositions do not appear in the female parent Tve, and only α-terpineol (0.95%) in the male parent Tq. After cluster analysis, among the 11 progeny lines, F_1-3, F_1-4, F_1-5, F_1-8, and F_1-11 are thymol-type; F_1-1, F_1-2, F_1-6, F_1-7, F_1-9, and F_1-10 are α-terpineol-type (Fig. 4D).

The phenotypes of plant type, fertility, and stem diameter of F_1 lines were obviously different from their parents of Tl × Tvf and Tve × Tq hybrid combinations (Supplementary Fig. S3). The stem about parents and F_1 lines of two populations display remarkable differences in morphology. The types of Tl is creeping-type, Tvf is erect-type, and their F_1 lines are creeping-type or semi-creeping-type (Supplementary Fig. S4). The types of Tve is erect-type, Tq is creeping-type, and their F_1 lines are erect-type or semi-erect-type (Supplementary Fig. S5). The creeping-type parent has a soft stem without a main stem, while the erect-type parent has an obvious main stem with a high degree of lignification, and the stem of its progeny is also between the parents. Some of flowers of F_1 progenies of above two hybrid combinations have stamens and are fertile males, so the inheritance of fertility has the characteristics of paternal inheritance (Supplementary Fig. S3).

Bioinformatics Analysis And Screen Of Terpene Synthase In Thyme

Terpene synthase (TPS) catalyzes GPP to form the basic skeleton of monoterpenes (C10) and sesquiterpenes (C15), respectively. The leaves volatile terpene compositions of Tl × Tvf and Tve × Tq aforementioned populations were determined. The chemical types of Tl × Tvf population were geraniol-type, geraniol/α-terpineol- type, thymol/carvacrol-type, and α-terpineol-type, while the chemical types of Tve × Tq population were carvacrol-type, thymol-type, and α-terpineol-type (Fig. 5B, D). In view of γ-terpinene was the catalytic precursor of thymol and carvacrol, we want to screen of γ-terpinene synthase, geraniol synthase, and α-terpineol synthase. Based on the results of *T. quinquecostatus* whole genome sequencing (Data unpublished), 69 TPS sequences were selected according to gene functional annotation, and 17 sequences were removed by conserved domain analysis of terpene synthase. The phylogenetic tree was constructed using remaining 52 TPS sequences and other species reported TPS (Fig. 6A; Supplementary Table S7). There are 22 TPS sequences belonging to the TPS-b family, all members of this family belong to monoterpene synthase, in which the target genes of γ-terpinene synthase are *Tq02G002290.1* and *Tq13G005250.1*, the target genes of α-terpineol synthase are *Tq03G001560.1*, and the purpose of geraniol synthase gene is *Tq04G005190.1*. The target genes *Tq02G002290.1* and *Tq13G005250.1* of γ-terpinene synthase were compared with the reported genes *TcTPS02.1* and *TcTPS02.2* in *T. caespititius* [28], and the similarities were 96.60% and 89.35%, respectively (Fig. 6B, C); The target gene *Tq03G001560.1* of α-terpineol synthase was compared with the reported genes *TcTPS05.1* and *TcTPS05.2* in *T. caespititius* [28], and the similarity is 95.51% (Fig. 6D); The target gene *Tq04G005190.1* of geraniol synthase is compared with *obGES* (Supplementary Fig. S6), which has been reported in *Ocimum basilicum* [37], and the similarity is 68.90%. Sequence alignment results show that the amino acid sequences of the target genes have the characteristic domain of terpene synthase, *Tq02G002290.1*, *Tq13G005250.1*, and *Tq03G001560.1* sequences have the characteristic domain of terpene synthase, including RRX8W, DDXXD, and NSE/DTE (Fig. 6B–D), and *Tq04G005190.1* sequences has the characteristic domain of terpene synthase, including DDXXD and NSE/DTE (Supplementary Fig. S6).

In population Tve × Tq, the chemical type of the male parent Tq is carvacrol, and the female parent Tve is thymol. Since γ-terpinene is the catalytic precursor of thymol and carvacrol, in order to better verify the γ-terpinene synthase function, *Tq02G002290.1* and *Tq13G005250.1* were verified in Tve, Tq and their progenies F_1-3 and F_1-4 (Fig. 6A). The qRT-PCR results showed that the relative expression of *Tq02G002290.1* and *Tq13G005250.1* were consistent compared with the results of the relative content of γ-terpinene in F_1-3 and F_1-4 lines. These preliminarily inferred that *Tq02G002290.1* and *Tq13G005250.1* maybe catalyze γ-terpinene biosynthesis (Fig. 6A). In population Tl × Tvf, the female parent Tl is geraniol-type and the male parent Tvf is α-terpineol-type. Therefore, in order to better verify the functions of α-terpineol synthase *Tq03G001560.1* and geraniol synthase *Tq04G005190.1*, the two TPS genes were subjected to qRT-PCR in Tl, Tvf, and their progenies F_1-6 and F_1-12 (Fig. 6A). The results showed that the expression levels of *Tq03G001560.1* and *Tq04G005190.1* genes were were consistent with their relative content of α-terpineol and geraniol in F_1-6 and F_1-12 lines. These preliminarily inferred that *Tq03G001560.1* and *Tq04G005190.1* maybe catalyze the α-terpineol and geraniol biosynthesis, respectively (Fig. 6A).

Discussion

Evaluation phenotypes of the thymes and construction of F_1 hybrid population

In this study, the essential oil (EO) yield of erect-type thyme was higher than that of creeping-type thyme, and there was a certain correlation between the density of glandular trichomes and the EO yield (Fig. 1E, F). Some Chinese wild thyme have a long flowering period up to four months, while some European thymes have a short flowering period of one month (Table 1). The tiller capacity of the creeping-type thymes were stronger than that of the erect-type thymes. In addition, the harvest time, extraction method and stem/leaf ratio of materials may also have a certain effect on EO yield. In this study, the composition type and relative content of EO compositions of 10 different thymes were significantly different, which further illustrated the diversity of thyme genus, and its chemical compositions were different with species (Table 2). In addition, it was reported that the composition and content of EO compositions of thyme in different environments of the same species were also different [38]. And different extraction methods will affect the composition and content of EO [16]. Therefore, the effects of species, region and extraction method should be considered when studying the content and main compositions of thyme EO

(Supplementary Fig. S1; Table 2). There is also a certain relationship between the EO color and the EO composition. The EO color of the α-terpineol-type thyme is milky white, the EO color of the geranyl-type thyme is light yellow almost lucency, and the EO colors of the thymol-type and the carvacrol-type are golden yellow. At present, the most studied volatile compositions of thyme at China and European are thymol and carvacrol, because thymol and carvacrol are the compositions with the most content in most of thymes [4, 8, 9]. In this study, the chemical types of the seven Seven European thymes are mainly thymol, geraniol and terpineol, and the chemical types of the three Chinese wild thymes are carvacrol and thymol.

The phenomenon of male sterility in plants is very convenient for cross breeding and can save the complicated operation of emasculation. Darwin found thymes in southern England belong to one species have two kinds of flower types, one flower type is that there is both completely male and female function, the other flower type is a small flower which is no anther or anther is completely cut, and therefore is complete male sterility. In different ecological conditions, in different thyme species, the proportion of female plants is more than 50% [39]. Because thyme flowers are very small, it is very difficult to cross and emasculate, so if some thyme plants are female, it is very convenient to cross. Among the 10 thymes in this study, Tr, Tve, Tg, Tt, Ts, and Tl are male sterile, while Tq, Tqp, Tm, and Tvf are male fertile. Therefore, male sterile thyme can be used as the female parent and male fertile thyme as the male parent in cross breeding. Male sterility is jointly determined by genes and environmental factors. In this study, male sterile thyme was introduced from European, so the phenomenon of male sterility may be caused by maladjustment to the new environment.

Chinese thyme plants are widely distributed and rich in resources [14], so breeders should intensify the research on them, and cross Chinese wild thyme with European thyme to breed new varieties suitable for medicinal and edible purposes, so as to promote the development of thyme planting industry. Hybrid breeding is an important means of germplasm resource innovation [12]. Through interspecific hybridization, the advantages of both parents can be aggregated. In this study, the new thyme germplasm resources were constructed by crossing different chemical types and different plant types of thyme.

Ssr Development And Its Application In F1 Hybrid Progenies

Hybrid breeding is an important means of germplasm resource innovation. The advantages of parents can be combined through interspecific hybridization. In this study, new germplasm resources of thyme were constructed by hybridizing thyme with different chemical types and plant types. SSR molecular marker are simple, time-consuming, reproducible and stable, which has been widely used in different species authenticity identification of hybrid progenies [15, 40]. For early identification and selection of hybrid progeny is an important link in cross breeding, enhance the efficiency of the hybrid early by morphological observation identification of hybrid progenies which the results of the identification accuracy is not high. Therefore, molecular marker technology combined with morphological observation can be used to identify the hybrid progenies to improve the accuracy of identification.

SSR molecular marker identification process is firstly to screen out codominant SSR primers in parents, and then to identify in progenies [40]. However, there may be errors in SSR molecular marker identification of hybrid progenies by using one pair of primers in practice. For example, in this study, primers TqSSR119 and TqSSR124 were used to identify F_1 hybrid progenies of the hybrid combination Tl × Tvf. Appear different primers of identification of the same sample in different appraisal result, this may be due to the thyme itself highly heterozygous genes, lead to some of the parents can't codominance or appear some non-parental bands, also may be due to in the process of meiotic division to form gametes, DNA regions of marker loci may be due to the exchange between homologous chromosomes amplification loci change, thus the marker bands of the parents disappears in the progenies.

Generally, the phenotype of hybrid F_1 progenies is between the parents, or biased towards the female or male parent, which also provides the possibility for the breeding of excellent new varieties [41]. In this study, among the 14 F_1 lines of Tl × Tvf population, eight F_1 lines had the same leaves volatile organic compositions (VOC) as their female parent Tl, one line had the same leaves VOC as their male parent Tvf, and five lines had different leaves VOC from their parents, indicating that the genetic characteristics of the compositions were maternal. Among 11 F_1 lines of Tve × Tq population, five lines had the same leaves VOC as their female parent Tve, and six lines were different from their male parent Tq. Therefore, the inheritance of volatile compositions in the F_1 progenies of thyme are mainly maternal inheritance. However, some F_1 progenies also have volatile compositions that are absolutely dominant and not found in their parents, such as thymol and carvatol are found in the progenies of Tl × Tvf population, α-terpenol and α-terpenol acetate are found in the progenies of Tve × Tq population. These results indicated that new varieties related volatile compositions could be bred from thyme through cross breeding. Heterosis was largely determined by the heterogeneity between parents, and the greater the phenotypic differences between parents, the stronger the heterosis of their progenies would be.

Bioinformatics Analysis And Screen Of Terpene Synthase In Thyme

Monoterpenoid biosynthesis begins with geranyl diphosphate (GPP) which is the precursor of all monoerpenoids, and yields α-terpinyl cations, that highly unstable intermediates can then be converted to specific monoerpenoids by certain monoterpene synthases such as γ-terpene and α-terpenol, etc. [42]. The synthetic precursors of geraniol, linalool, myrcene, and ocimene are also GPP, which are formed under the catalysis of different monoterpene synthases. In addition, CYP71D178, CYP71D180, and CYP71D181, which belong to cytochrome P450 (CYP) monooxygenase, are also involved in further modification of *p*-cymene framework to produce thymol and carvonol [43], and the synthetic pathway of some monoterpenes is shown in Fig. 6B. The gene function of γ-terpinene synthase TcTPS02 and α-terpineol synthase TcTPS05 were validated in *T. caespititius* [28]. The γ-terpinene synthase TvTPS2 and three cytochrome P450 (TvCYP71D179, TvCYP71D180, and TvCYP71D507) were validated in *T. vulgaris* [44].

In this study, the screened TPSs were expressed in thyme population by qRT-PCR and verified by combining the relative content of the TPS catalyzed product in thyme. Two γ-terpinene synthases Tq02G002290.1 and Tq13G005250.1 were verified in Tve × Tq population, whose parental chemotypes were thymol and carvatol. α-Terpineol synthase Tq03G001560.1 and geraniol synthase Tq04G005190.1 were validated in Tl × Tvf population with parental chemotypes were α-terpineol and geraniol. The results showed that the relative expression levels of *Tq02G002290.1*, *Tq13G005250.1*, *Tq03G001560.1*, and *Tq04G005190.1* were

consistent with the relative content of catalytic products in thyme some F_1 lines. Therefore, this method is only a preliminary screen of the gene function, and provides a basis for further gene function verification. The above results laid the foundation for the creation of new germplasm resources, construction of genetic linkage map, location of QTL, MAS breeding, and insight into the mechanism of monoterpenoids biosynthesis in thymes.

Conclusions

Essential oil of thyme has antibacterial, antifungal, and antiviral properties. In order to obtain EOs different terpene compositions, we cross Chinese wild thymes with European thymes to breed new germplasm resources suitable for medicinal and edible purposes, so as to promote the development of thyme industry. Two F_1 populations of thyme were obtained to construct genetic linkage map and locate QTL of terpene compositions. We screened monoterpene synthase in thyme to lay the foundation of insight into the mechanism of monoterpenoids biosynthesis in thyme.

Methods

Plant materials

Three Chinese wild thymes *T. quinquecostatus*, *T. quinquecostatus* var. *przewalskii*, *T. mongolicus* and seven European thymes *T. vulgaris* 'Fragrantissimus', *T. vulgaris* 'Elsbeth', *T. guberlinesis* 'lijin', *T. serpyllum* 'Aureus', *T. thracicus*, *T. longicaulis*, and *T. rotundifolius*. Taken the male sterile line as the female parent and the male fertile line as the male parent, we obtained two F_1 populations of *T. longicaulis* (Tl) × *T. vulgaris* 'Fragrantissimus' (Tvf) and *T. vulgaris* 'Elsbeth' (Tve) × *T. quinquecostatus* (Tq) by crossing in 2019, China. All the thyme materials were grown in an experimental farm, Institute of Botany, Chinese Academy of Sciences (IB-CAS), Beijing, China.

Dna Extraction

The leaves of *T. longicaulis* (female parent), *T. vulgaris* 'Fragrantissimus' (male parent) and their 14 F_1 lines, and *T. vulgaris* 'Elsbeth' (female parent), *T. quinquecostatus* (male parent) and their 11 F_1 lines were collected from plants, all samples were immediately frozen in liquid nitrogen or dry ice and stored at -80°C before DNA extraction. DNA was extracted from the plant using a DNA Secure Plant Kit (Tiangen, China). The DNA concentration and quality were assessed by 1% agarose gel electrophoresis and with a 2.0 Fluorometer (Life Technologies, CA, USA).

Ssr Amplification

All SSR-polymerase chain reaction (PCR) were performed on PCR system (Bio-Rad,

Hercules, CA, USA). PCR was performed in a 10 μl reaction volume containing 2 μl (20 ng/μl) genomic DNA, 3 μl 2 × Taq PCR Master Mix II (Tiangen, China), 2 μl forward and reverse primer mixture, and 3 μl ddH_2O, its reaction procedure is as follows: 94°C for 3 min, followed by 6 cycles of 45 s at 94°C, 1 min at 55–65°C, and 1 min at 72°C, and followed by 9 cycles of 45 s at 94°C, 1 min at 50–58°C, and 1 min at 72°C, followed by 19 cycles of 30 s at 94°C, 50°C for 30 s, and 1 min at 72°C, with a final extension of 72°C for 5 min. Amplification products were analyzed by electrophoresis in 8.0% (w/v) denaturing polyacrylamide gel in TBE buffer for 1 h on the DYY-6C electrophoresis apparatus (Beijing Liuyi Instrument Factory, China) under 220 V constant voltage. Fragments were then visualized by silver staining (Silver sequence staining reagents, Promega, Madison, USA) and sized with 50 base pairs DNA ladder marker (Tiangen, China) [40]. SSR primer sequences were listed in Supplementary Table S2.

Rna Extraction And Cdna Synthesis

In order to isolate RNAs, leaves of *T. longicaulis*, *T. vulgaris* 'Fragrantissimus', *T. vulgaris* 'Elsbeth', *T. quinquecostatus*, and their F_1 lines were frozen in liquid nitrogen and stored at -80 °C. Total RNA was extracted from these tissues using the Easy Spin RNA extraction kit (Sangon Biotech, Shanghai, China), treated with DNase I and further purified with RNA clean kit (Promega, Madison, WI, USA). The concentration of each RNA sample was checked using a NanoDrop spectrophotometer (Thermo Fisher Scientific Inc., USA) and a 2.0 Fluorometer (Life Technologies, CA, USA). The RNA integrity was checked using a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA). cDNAs were synthesized using a HiScript Reverse Transcriptase Kit (Vazyme, China) according to the manufacturer's instructions [32].

Real-time Quantitative Pcr (Qrt-pcr)

To analyze γ-terpinene, α-terpineol, and geraniol synthase gene expression patterns, qRT-PCR was performed. Total RNAs from *T. vulgaris* 'Elsbeth', *T. quinquecostatus*, *T. longicaulis*, *T. vulgaris* 'Fragrantissimus', and their F_1 lines were isolated and reverse transcribed, as described above. The γ-terpinene, α-terpineol, and geraniol synthase gene primers were designed by primer3 (http://primer3.ut.ee), and are described in Supplementary Table S8. The total reaction volume was 20 μl containing 0.8 μl of each primer, 1 μl of template cDNA, 10 μl 2 × T5 Fast qPCR Mix (TSINGKE, China), and 7 μl of ddH_2O, 0.4 μl 50 × ROX Reference Dye (TaKaRa, China). The PCR program was initiated with a preliminary step of 1 min at 95°C, followed by 40 cycles at 95°C for 10 s, 50–60°C for 5 s, and 72°C for 15 s. qRT-PCR was carried out with a Ssofast EvaGreen Suppmix Kit using a CFX96 instrument (Bio-Rad, USA). Relative quantification of transcripts was performed using the 18sRNA. The difference in relative expression levels of TPS was calculated from the $2^{-\Delta\Delta Ct}$ value after the normalization of thyme TPSs data. All analyses were performed triplicate [37].

Essential Oil (Eo) Extraction

The essential oil (EO) from ten thymes were isolated by steam distillation. The steam distillation was performed at 180–200 ℃ for 90 min. The EO yield (%) was calculated as volume (ml) of the isolated oil per 100 g of the dry plant material. The isolated EOs were dried using anhydrous sodium sulfate and stored at 4 ℃, until tested [11].

Analysis Of Eos By Gas Chromatography-mass Spectrometry (Gc-ms)

Gas chromatographic analyses of ten thyme EOs were performed on an Agilent 7890A-7000B gas chromatograph (Agilent, USA) and equipped with an Agilent 5975C MS detector (Agilent, USA). Using HP-5MS (30 m, 250m ID, 0.25 μm film thickness) capillary column, separation of volatiles were used with the following temperature programme: 5 min at 60°C, then 4°C/min to 220°C, then 60°C/min to 250°C, held for 5 min. Injector and detector temperature: 250°C; carrier gas: He; flow rate: 1 m/min; split ratio: 1:10; acquisition range: 50–500 m/z in electron-impact mode; ionization voltage: 70 eV, and the injected volume of the sample was 1μl. The determination of the percentage content of each compositions was based on the normalization of the GC peak areas. The identification of the OEs compositions were based on the comparison of their retention indices (RIs), relative to a homologous series of n-alkanes (C7–C40) and mass spectra from the National Institute of Standards and Technology (NIST, 4.0) library and literature data [45].

Analysis Of Leaves Volatile Organic Compositions (Voc) By Headspace Solid-phase Microextraction (Hs-spme)

The leaves VOCs of *T. vulgaris* 'Elsbeth', *T. quinquecostatus*, *T. longicaulis*, *T. vulgaris* 'Fragrantissimus', and their F_1 lines were detected via headspace solid-phase microextraction (HS-SPME) as follows: 0.25 g of fresh leaf powder was weighed and immediately placed into a 20 ml headspace vial (Aligent, Palo Alto, CA, USA) containing 20 μl of internal standard solution (1 mg/ml, 3-Octanol, Cas#589-98-o, Aladdin, Shanghai, China). The vials were sealed using crimp-top caps with TFE-silicone headspace septa (Agilent, Palo Alto, CA, USA). Subsequently, each vial was immediately incubated at 40°C for 30 min, then a 100 μm coating fiber polydimethylsiloxane (Supelco, Inc., Bellefonte, PA, USA) was exposed to the headspace to absorb the volatiles for 30 min. All volatile compositions on the coating fiber were then analyzed using gas chromatography (GC). A Model 7890A GC instrument and a 7000B mass spectrometer (Agilent, Palo Alto, CA, USA) were used to perform GC-MS analysis of leaves VOCs [46].

The chromatographic conditions were as follows: the injector and transfer line temperatures were regulated at 250 ℃ and 250 ℃, respectively; column temperature was initially kept at 50 ℃ for 3 min, then gradually increased to 150 ℃ at 4 ℃/min for 2 min, and finally raised to 250 ℃ at 8 ℃/min for 5 min; carrier gas (helium) flow was 1 ml/min; injection was performed in splitless mode. Identification of volatile compositions were performed comparing their retention times with those of authentic standards. Ionization voltage was 70 eV; source temperature was 250 ℃; mass spectra were scanned in the range 35–500 m/z. Agilent MassHunter 5.0 was used to analyze the chromatograms and mass spectra. The volatile compositions were identified by comparing the retention times of individual peaks and identified the mass spectra using the mass spectra databases NIST 4.0 and literature data [47].

Density Of Glandular Trichomes

The density of the glandular trichomes were evaluated using a stereomicroscope (Leica DVM6, Germany). The ImageJ software was used to count the glandular trichomes and measure the leaf area. The density was calculated based on three plants.

Statistical analysis

All samples were prepared and analyzed in triplicate, and data were expressed as the mean ± standard deviation. Statistical analyses were performed using the variance (ANOVA) test. Duncan's test was used to determine the significance of differences between the groups. Differences at $P < 0.05$ were considered significant. SPSS 18 (SPSS Inc., Chicago, IL, USA) was used for the analysis.

Abbreviations

DMAPP
Dimethylallyl diphosphate
EO
Essential oil
GC
Gas chromatography
GC-MS
Gas chromatography-mass spectrometry
GPP
Geranyl diphosphate
Hi-C
Chromatin conformation capture
HiFi

High-fidelity
HS-SPME
Headspace solid-phase microextraction
IB-CAS
Institute of Botany, Chinese Academy of Sciences
IPP
Isopentenyl diphosphate
MAS
Molecular marker-assisted selection
Mb
Megabases
MEP
2-C-methyl-D-erythritol-4-phosphate
MVA
Mevalonate
PCA
Principal component analysis
PCR
Polymerase chain reaction
qRT-PCR
Real-time quantitative PCR
QTL
Quantitative trait locus
RIs
Retention indices
SSR
Simple sequence repeat
Tg
T. guberlinesis 'Iijin'
TI
T. longicaulis
Tm
T. mongolicus
TPS
Terpene synthase
Tq
T. quinquecostatus
Tqp
T. quinquecostatus var. *przewalskii*
Tr
T. rotundifolius
Ts
T. serpyllum 'Aureus'
Tt
T. thracicus
Tve
T. vulgaris 'Elsbeth'
Tvf
T. vulgaris 'Fragrantissimus'
VOC
Volatile organic compositions.

Declarations

Ethics approval and consent to participate

The data collection of plants were carried out with permission of related institution, and complied with national or international guidelines and legislation.

Consent for publication

Not applicable as the manuscript contains no individual identifying data.

Availability of data and material

The raw sequence data and gene sequence of *T. quinquecostatus* genome have been deposited in NCBI under project accession No. PRJNA690675. All supplementary figures and tables are provided in Supplemental information. All material are collected and grown in an experimental farm, Institute of Botany, Chinese Academy of Sciences (IB-CAS), Beijing, China.

Competing interests

Conflict of interest the authors declare that they have no conflict of interest.

Funding

This study was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA23080603).

Authors' contributions

M.Y.S. and L.Z. performed the experiments, analyzed the data and wrote the manuscript. Y.N.Z. and N.N.L. helped to analyze the volatile organic compositions of thymes. J.Z.Z., H.L. and H.T.B. helped to collect the samples and analyzed the data. L.S. was involved in designing the research and revising the manuscript. All authors read and approved the manuscript.

Acknowledgements

We thank PhD Yan Zhu and Lu Wang from Plant Science Facility of the Institute of Botany, Chinese Academy of Sciences for their excellent technical assistance of GC-MS and HS-SPME assistance on essential oil and leaves volatile organic compositions.

References

1. Peter KV. **Handbook of Herbs and Spices** (**Second edition**). **Woodhead Publishing Series in Food Science**, **Technology and Nutrition**. 2012;**1**:**499**–**525**.
2. Tohidi B, Rahimmalek M, Arzani A. **Variations in chemical composition and bioactive compounds of** *Thymus kotschyanus* **Boiss. & Hohen populations originated from different collection sites**. J Essent Oil Bear Pl. 2018;21:1272–83.
3. Lorenzo JM, Khaneghah AM, Gavahian M, Marszalek K, Es I, Munekata PES. **et al**. Understanding the potential benefits of thyme and its derived products for food industry and consumer health: From extraction of value-added compounds to the evaluation of bioaccessibility, bioavailability, anti-inflammatory, and antimicrobial activities. Crit Rev Food Sci Nutr. 2019;59:2879–95.
4. Salehi B, Abu-Darwish MS, Tarawneh AH, Cabral C, Gadetskaya AV, Salgueiro L. **et al**. *Thymus* **spp. plants-Food applications and phytopharmacy properties**. Trends Food Sci Tech. 2019;85:287–306.
5. Rota MC, Herrera A, Martinez RM, Sotomayor JA, Jordan MJ. **Antimicrobial activity and chemical composition of** <bi>Thymus vulgaris, Thymus zygis</bi> **and** <bi>Thymus hyemalis</bi> **essential oils**. **Food Control**. 2008;**19**:**681**–**687**.
6. Kargioglu M, Cenkci S, Serteser A, Konuk M, Vural G. Traditional uses of wild plants in the middle aegean region of Turkey. Hum Ecol. 2010;38:429–50.
7. Ocana A, Reglero G. **Effects of thyme extract oils (from** *Thymus vulgaris*, *Thymus zygis*, **and** *Thymus hyemalis*) **on cytokine production and gene expression of oxLDL-Stimulated THP-1-Macrophages**. Int J Obes. 2012;2012:104706.
8. Kowalczyk A, Przychodna M, Sopata S, Bodalska A, Fecka I. Thymol and thyme essential oil-new insights into selected therapeutic applications. Molecules. 2020;25:4125.
9. Nabavi SM, Marchese A, Izadi M, Curti V, Daglia M, Nabavi SF. Plants belonging to the genus Thymus as antibacterial agents: From farm to pharmacy. Food Chem. 2015;173:339–47.
10. Bigdeloo M, Hadian J, Nazeri V. **Composition of essential oil compounds from different populations of** *Thymus caramanicus* **Jalas**. J Appl Res Med Aromat Plants. 2017;7:95–8.
11. Kim M, Moon JC, Kim S, Sowndhararajan K. **Morphological, chemical, and genetic characteristics of Korean native thyme Bak-Ri-Hyang (***Thymus quinquecostatus* **Celak.)**. Antibiotics-Basel. 2020;9:289.
12. Carlen C, Schaller M, Carron CA, Vouillamoz JF, Baroffio CA. **The new** *Thymus vulgaris* **L. hybrid cultivar 'Varico 3' compared to five established cultivars from Germany, France and Switzerland**. Acta Hortic. 2010;860:161–6.
13. Shetty K, Carpenter TL, Kwok D, Curtis OF, Potter TL. **Selection of high phenolics-containing clones of thyme (***Thymus vulgaris* **L) using** *Pseudomonas sp*<bi>.</bi>. J Agric Food Chem. 1996;44:3408–11.
14. Li XW, Ian CH. Flora of China. Science Press (Beijing) and Missouri Botanical Garden Press. (St Louis). 1994;17:186–8.
15. Martins S, Simoes F, Mendonca D, Matos J, Silva AP, Carnide V. Western European wild and landraces hazelnuts evaluated by SSR markers. Plant Mol Biol Rep. 2015;33:1712–20.
16. Tohidi B, Rahimmalek M, Trindade H. **Review on essential oil, extracts composition, molecular and phytochemical properties of** *Thymus* **species in Iran**. Ind Crop Prod. 2019;134:89–99.
17. Trindade H, Costa MM, Sofia BLA, Pedro LG, Figueiredo AC, Barroso JG. **Genetic diversity and chemical polymorphism of** *Thymus caespititius* **from Pico, Sao Jorge and Terceira islands (Azores)**. Biochem Syst Ecol. 2008;36:790–7.

18. Trindade H, Costa MM, Lima SB, Pedro LG, Figueiredo AC, Barroso JG. **A combined approach using RAPD, ISSR and volatile analysis for the characterization of** *Thymus caespititius* **from Flores, Corvo and Graciosa islands (Azores, Portugal)**. Biochem Syst Ecol. 2009;37:670–7.
19. Younsi F, Rahali N, Mehdi S, Boussaid M, Messaoud C. Relationship between chemotypic and genetic diversity of natural populations of Artemisia herba-alba Asso growing wild in Tunisia. Phytochemistry. 2018;148:48–56.
20. Rustaiee AR, Yavari A, Nazeri V, Shokrpour M, Sefidkon F, Rasouli M. **Genetic diversity and chemical polymorphism of some** *Thymus* **species**. Chem Biodivers. 2013;10:1088–98.
21. Talebi M, Rahimmalek M, Norouzi M. **Genetic diversity of** *Thymus daenensis* **subsp** *daenensis* **using SRAP markers**. Biologia. 2015;70:453–9.
22. Zhou F, Pichersky E. More is better: the diversity of terpene metabolism in plants. Curr Opin Plant Biol. 2020;55:1–10.
23. Chen F, Tholl D, Bohlmann J, Pichersky E. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. Plant J. 2011;66:212–29.
24. Karunanithi PS, Zerbe P. Terpene synthases as metabolic gatekeepers in the evolution of plant terpenoid chemical diversity. Front Plant Sci. 2019;10:1166.
25. Alicandri E, Paolacci AR, Osadolor S, Sorgona A, Badiani M, Ciaffi M. **On the evolution and functional diversity of terpene synthases in the** *Pinus* **species: a review**. J Mol Evol. 2020;88:253–83.
26. Keszei A, Brubaker CL, Foley WJ. A molecular perspective on terpene variation in Australian Myrtaceae. Aust J Bot. 2008;56:197–213.
27. Lima AS, Lukas B, Novak J, Figueiredo AC, Pedro LG, Barroso JG. **et al**. **Genomic characterization of γ-terpinene synthase from** *Thymus caespititius*. Planta Med. 2011;77:PI9.
28. Lima AS, Schimmel J, Lukas B, Novak J, Barroso JG, Figueiredo AC. **et al**. **Genomic characterization, molecular cloning and expression analysis of two terpene synthases from** *Thymus caespititius* **(Lamiaceae)**. Planta. 2013;238:191–204.
29. Mendes MD, Barroso JG, Oliveira MM, Trindade H. **Identification and characterization of a second isogene encoding γ-terpinene synthase in** *Thymus caespititius*. J Plant Physiol. 2014;171:1017–27.
30. Müller-Uri F, Egerer-Sieber C, Rudolph K, Kreis W, Muller Y. **γ-Terpinene synthase of** *Thymus vulgaris*. Planta Med. 2015;81:PM–240.
31. Mashhadi MA, Majdi M, Maroufi A. **Expression analysis of biosynthetic genes of thymol and carvacrol in different tissues of thyme (***Thymus vulgaris*). Iran J Rangelands For Plant Breed. 2016;25:160–71.
32. Rudolph K, Parthier C, Egerer-Sieber C, Geiger D, Muller YA, Kreis W. **et al**. **Expression, crystallization and structure elucidation of γ-terpinene synthase from** *Thymus vulgaris*. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2016;72:16–23.
33. Majdi M, Malekzadeh-Mashhady A, Maroufi A, Crocoll C. **Tissue-specific gene expression patterns of genes associated with thymol/carvacrol biosynthesis in thyme (***Thymus vulgaris* **L.) and their differential changes upon treatment with abiotic elicitors**. Plant Physiol Biochem. 2017;115:152–62.
34. Filipe A, Cardoso JC, Miguel G, Anjos L, Trindade H, Figueiredo AC. **et al**. **Molecular cloning and functional characterization of a monoterpene synthase isolated from the aromatic wild shrub** *Thymus albicans*. J Plant Physiol. 2017;218:35–44.
35. Glas JJ, Schimmel BCJ, Alba JM, Escobar-Bravo R, Schuurink RC, Kant MR. Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. Int J Mol Sci. 2012;13:17077–103.
36. Schilmiller AL, Last RL, Pichersky E. Harnessing plant trichome biochemistry for the production of useful compounds. Plant J. 2008;54:702–11.
37. Zhao CH, Yu ZM. **da** Silva JAT, He CM, Wang HB, Si C, **et al**. **Functional characterization of a** *Dendrobium officinale* **geraniol synthase DoGES1 involved in floral scent formation**. Int J Mol Sci. 2020;21:7005.
38. Tohidi B, Rahimmalek M, Arzani A. **Essential oil composition, total phenolic, flavonoid contents, and antioxidant activity of** *Thymus* **species collected from different regions of Iran**. Food Chem. 2017;220:153–61.
39. Stahl-Biskup E, Saez F. **Thyme The genus** *Thymus* **(First edition)**. 81997699: Taylor & Francis Inc; 2002.
40. Li FY, Liu X, Wu SY, Luo QY, Yu BJ. **Hybrid identification for** *Glycine max* **and** *Glycine soja* **with SSR markers and analysis of salt tolerance**. PeerJ. 2019;7:e6483.
41. Duskova E, Dusek K, Smekalova K, Nosalkova M, Benicka S. **Genetic resources of** *Thymus vulgaris* **L. and** *T. vulgaris* × *T Marschallianus* **Willd in the Czech Republic** Julius Kuhn Archiv. 2016;453:56–8.
42. Dhandapani S, Tjhang JG, Jang IC. Production of multiple terpenes of different chain lengths by subcellular targeting of multi-substrate terpene synthase in plants. Metab Eng. 2020;61:397–405.
43. Crocoll C. **Biosynthesis of the phenolic monoterpenes, thymol and carvacrol, by terpene synthases and cytochrome P450s in oregano and thyme. DPhil Thesis.Friedrich-Schiller-Universität Jena, Jena**. 2011.
44. Krause ST, Liao P, Crocoll C, Boachon B, Forster C, Leidecker F. **et al**. The biosynthesis of thymol, carvacrol, and thymohydroquinone in Lamiaceae proceeds via cytochrome P450s and a short-chain dehydrogenase. Proc Natl Acad Sci USA. 2021;118:e2110092118.
45. Bendif H, Boudjeniba M, Miara MD, Biqiku L, Bramucci M, Lupidi G. **et al**. **Essential Oil of** *Thymus munbyanus* **subsp** *coloratus* **from Algeria: chemotypification and in vitro biological activities**. Chem Biodivers. 2017;14:e1600299.
46. Pontes M, Marques JC, Camara JS. **Headspace solid-phase microextraction-gas chromatography-quadrupole mass spectrometric methodology for the establishment of the volatile composition of** *Passiflora* **fruit species**. Microchem J. 2009;93:1–11.
47. Bendif H, Peron G, Miara MD, Sut S, Dall'Acqua S, Flamini G, **et al**. **Total phytochemical analysis of** <bi>Thymus munbyanus</bi> **subsp**. <bi>coloratus</bi> **from Algeria by HS-SPME-GC-MS, NMR and HPLC-MSn studies. J Pharmaceut Biomed**. 2020; **186**:**113330**.

Figures

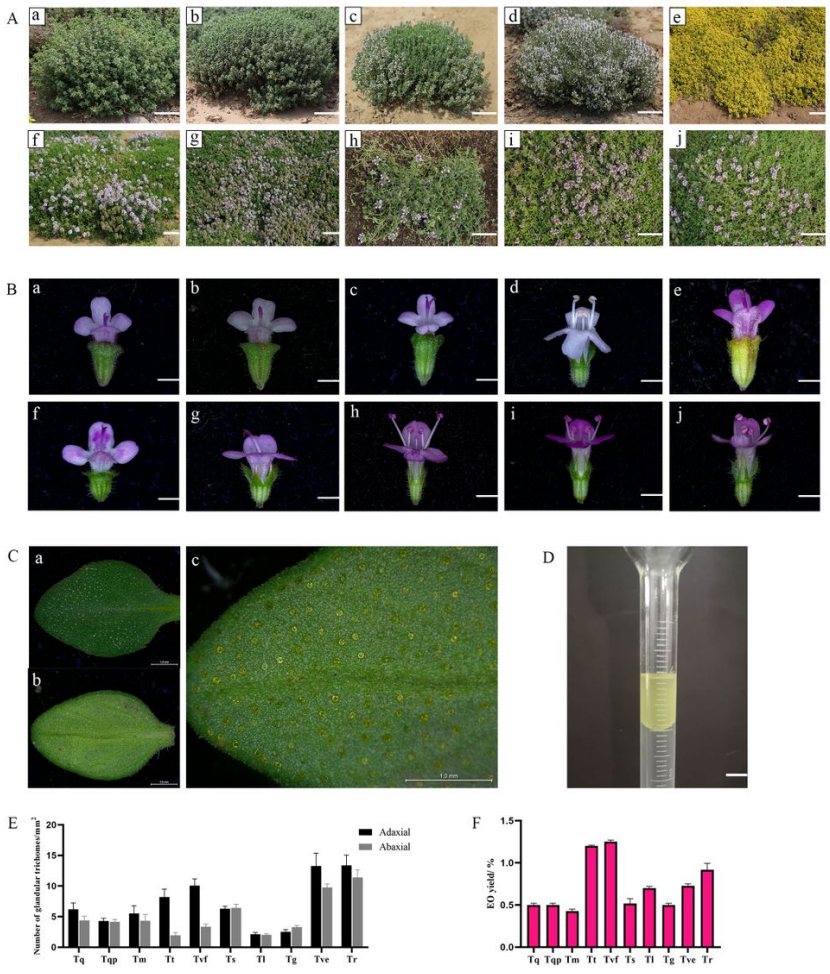

Figure 1

Evaluation phenotypes of three Chinese wild thymes and seven European thymes. A Plant types of three Chinese wild thymes and seven European thymes. **a** *T. rotundifolius* (Tr); **b** *T. vulgaris* 'Elsbeth' (Tve); **c** *T. thracicus* (Tt); **d** *T. vulgaris* 'Fragrantissimus' (Tvf); **e** *T. serpyllum* 'Aureus' (Ts); **f** *T. guberlinesis* 'lijin' (Tg); **g** *T. longicaulis* (Tl); **h** *T. quinquecostatus* (Tq); **i** *T. quinquecostatus* var. *Przewalskii* (Tqp); **j** *T. mongolicus* (Tm). **B** Flower fertility of three Chinese wild thymes and seven European thymes. **a** *T. rotundifolius* (Tr); **b** *T. vulgaris* 'Elsbeth' (Tve); **c** *T. thracicus* (Tt); **d** *T. vulgaris* 'Fragrantissimus' (Tvf); **e** *T. serpyllum* 'Aureus' (Ts); **f** *T. guberlinesis* 'lijin' (Tg); **g** *T. longicaulis* (Tl); **h** *T. quinquecostatus* (Tq); **i** *T. quinquecostatus* var. *Przewalskii* (Tqp); **j** *T. mongolicus* (Tm). Bars = 1 mm. **C** Glandular trichomes of adaxial and abaxial in leaf. **a** Adaxial; **b**Abaxial; **c** Glandular trichomes on the abaxial. **D** Images of thyme essential oil. **E** Glandular trichomes number of adaxial and abaxial in leaf. **F** The essential oil (EO) yield.

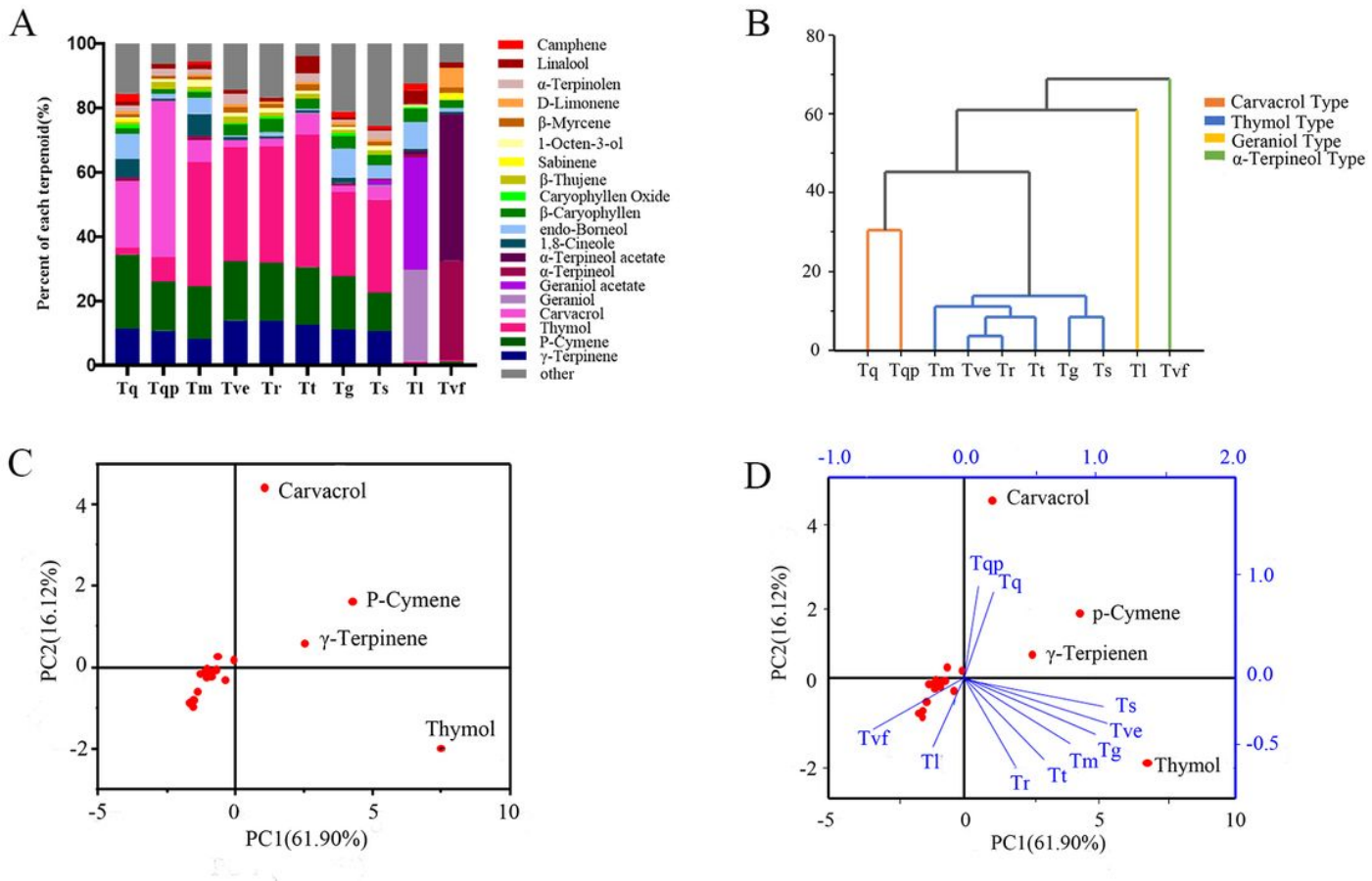

Figure 2

Relative content, cluster analysis and PCA analysis of essential oil compositions of 10 thymes. **A** Histogram of relative content of 10 thyme essential oil terpenoids. **B** Cluster analysis of 10 thymes using essential oil terpenoids compositions. **C** The loading plot of active essential oil terpenoids compositions. **D** The score plot of the tested thyme and essential oil compositions.

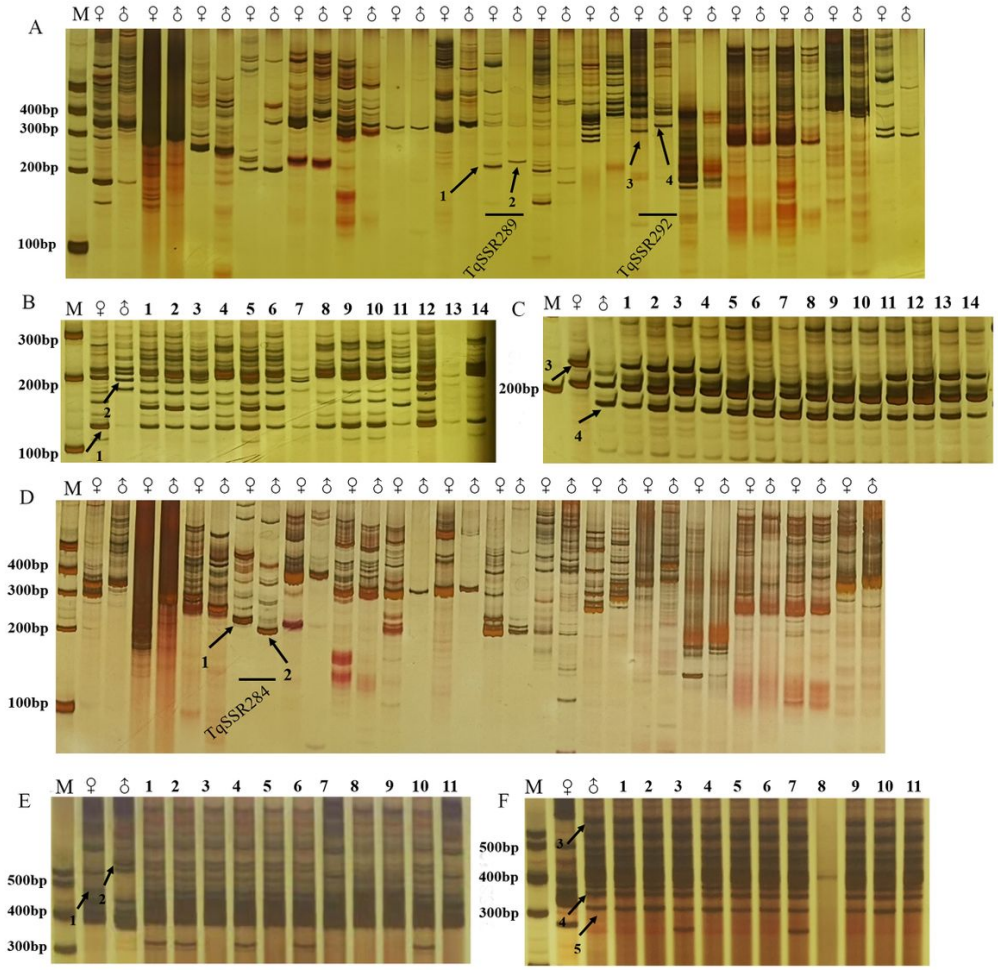

Figure 3

Amplification patterns of two population parents and F_1 lines in partial SSR primer screening in thymes. A Amplification patterns of parents in Tl × Tvf population of partial SSR primer screening. **B** SSR detection results of primers TqSSR119 in parental Tl, Tvf and hybrid F_1 lines. **C** SSR detection results of primers TqSSR124 in parental Tl, Tvf and hybrid F_1 lines. **D** Amplification patterns of parents in Tve × Tq population of partial SSR primer screening. **e** SSR detection results of primers TqSSR170 on parental Tve, Tq and hybrid F_1 lines. **f** SSR detection results of primers TqSSR104 on parental Tve, Tq and hybrid F_1 lines.

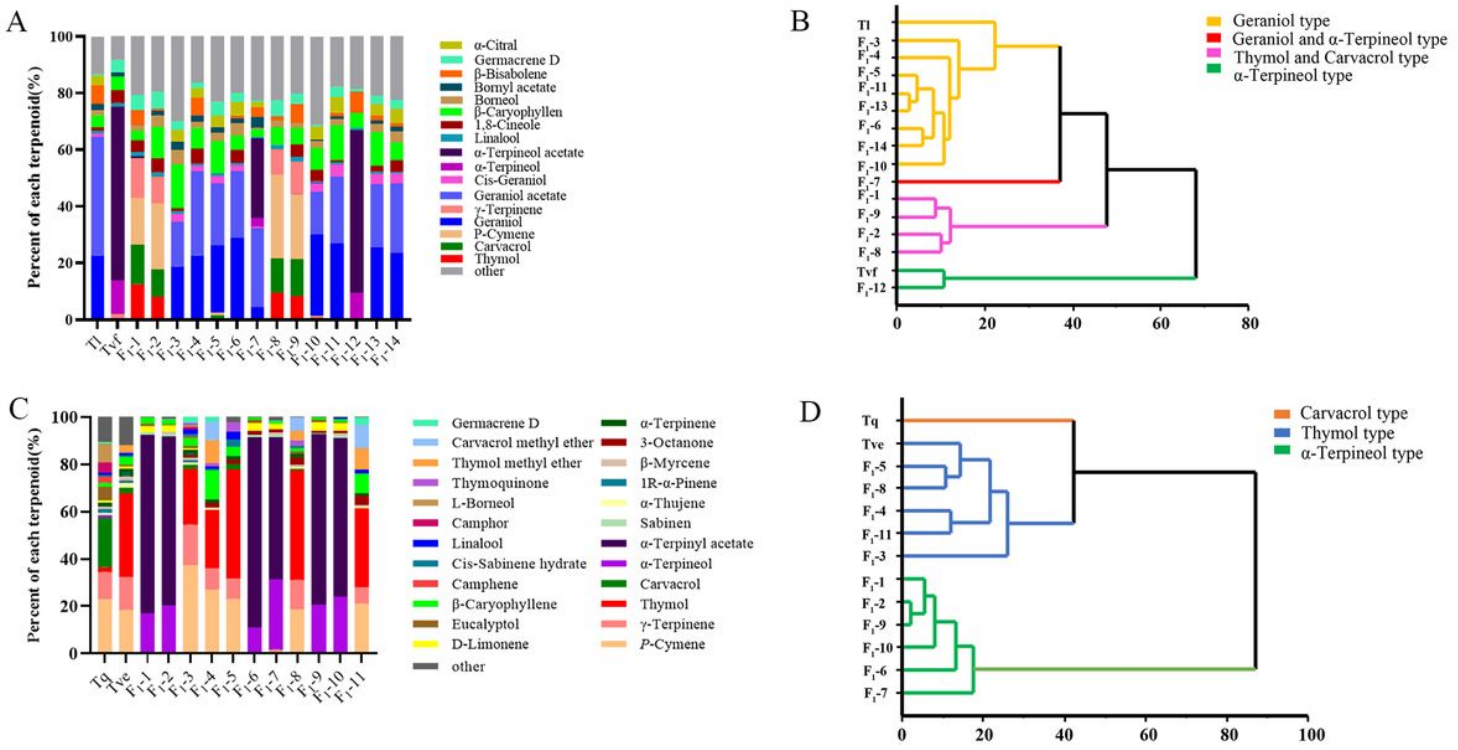

Figure 4

Analysis of volatile terpenoids in leaves of two populations parents and hybrid F_1 lines in thymes. A The volatile terpenoids histogram in leaves of Tl × Tvf parents and hybrid F_1 lines. **B** The volatile terpenoids cluster analysis in leaves of Tl × Tvf parents and hybrid F_1 lines. **C** The volatile terpenoids histogram in leaves of Tve × Tq parents and hybrid F_1 lines. **D** The volatile terpenoids cluster analysis in leaves of Tve × Tq parents and hybrid F_1 lines.

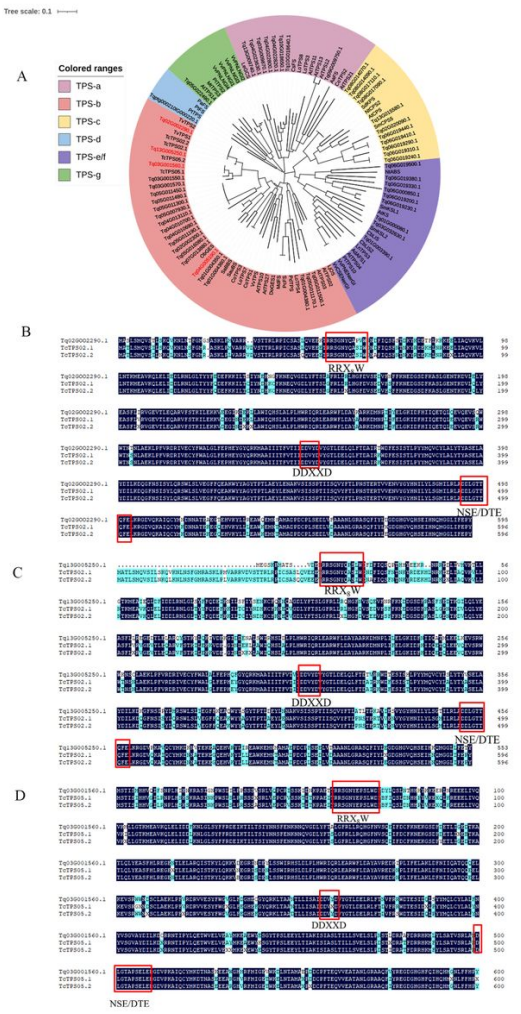

Figure 5

TPS phylogenetic tree and amino acid sequence alignment of TPS proteins in thymes. A 52 TPSs phylogenetic tree in thymes. **B** Amino acid sequence alignment of *Tq02G002290.1* with the reported amino acid sequence of γ-terpinene synthase *TcTPS02.1* and *TcTPS02.2* (Lima et al., 2013). **C** Amino acid sequence alignment of *Tq13G005250.1* with the reported amino acid sequence of γ-terpinene synthase *TcTPS02.1* and *TcTPS02.2* (Lima et al., 2013). **D** Amino acid sequence alignment of *Tq03G001560.1* with the reported amino acid sequence of α-terpineol synthase *TcTPS05.1* and *TcTPS05.2* (Lima et al., 2013). The conserved motifs (RRX8W, DDXXD and NSE/DTE) are highlighted with red boxes.

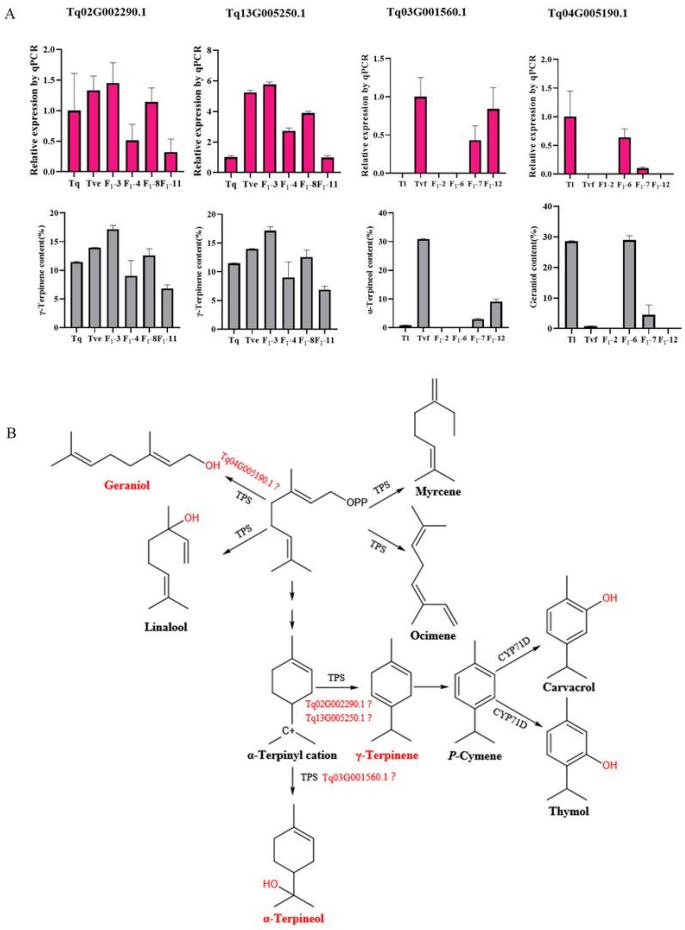

Figure 6

TPS target gene screening and biosynthesis pathway of some monoterpenoids. A Expression of TPS target gene in the leaves of Tve × Tq and Tl × Tvf population parents and part of F_1 lines. **B** The biosynthesis pathway of some monoterpenoids.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementalinformation.pdf