Terpene chemotypes in *Gossypium hirsutum* (wild cotton) from the Yucatan Peninsula, Mexico

Mary V. Clancy [a], Marine Mamin [a], Galien Flückiger [a], Teresa Quijano-Medina [b], Biiniza Pérez-Niño [b], Luis Abdala-Roberts [b], Ted C.J. Turlings [a,*], Carlos Bustos-Segura [a]

[a] *University of Neuchâtel, Institute of Biology, Fundamental and Applied Research in Chemical Ecology, Rue Emile-Argand 11, CH-2000, Neuchâtel, Switzerland*
[b] *Departamento de Ecología Tropical, Campus de Ciencias Biológicas y Agropecuarias, Universidad Autónoma de Yucatán, Km. 15.5 Carretera Mérida-Xtmakuil s/n, Mérida, Yucatán, 97200, Mexico*

ARTICLE INFO

ABSTRACT

Cultivated plants of *Gossypium hirsutum* Cav. (cotton) consistently emit low levels of volatile organic compounds, primarily mono- and sesquiterpenoids, which are produced and stored in pigment glands. In this study, we provide a comprehensive evaluation of the terpene profiles of wild *G. hirsutum* plants sourced from sites located throughout natural distribution of this species, thus providing the first in-depth assessment of the scope of its intraspecific chemotypic diversity. Chemotypic variation can potentially influence resistance to herbivory and diseases, or interact with abiotic stress such as extreme temperatures. Under controlled environmental conditions, plants were grown from seeds of sixteen *G. hirsutum* populations collected along the coastline of the Yucatan Peninsula, which is its likely centre of origin. We found high levels of intraspecific diversity in the terpene profiles of the plants. Two distinct chemotypes were identified: one chemotype contained higher levels of the monoterpenes γ-terpinene, limonene, α-thujene, α-terpinene, terpinolene, and p-cymene, while the other chemotype was distinguished by higher levels of α- and β-pinene. The distribution of chemotypes followed a geographic gradient from west to east, with an increasing frequency of the former chemotype. Concurrent analysis of maternal plants revealed that chemotypes in wild *G. hirsutum* are highly heritable.

1. Introduction

Many plants constitutively produce a wide range of volatile organic compounds (VOCs) that play an important role in mediating the relationships between plants and their surrounding environment, and can influence their interactions with many species (Rosenkranz et al., 2021). Terpenes are a large chemical class of VOCs that are involved in a variety of functions that contribute to overall plant fitness, including defence and inter- and intra-specific communication (Gershenzon and Dudareva, 2007). They are found across the plant kingdom and display high levels of variation in structure and abundance both across and within plant families and species (Pichersky and Raguso, 2018). A large source of the chemical variation found in plants is due to the vast range of terpenoid compounds, of which approximately 80,000 are known to date (Christianson, 2017). They are derived from both the mevalonic acid and methylerythritol pyruvate pathways, and are produced by terpene synthases, which are capable of catalysing a wide assortment of structures from single substrates (Zhou and Pichersky, 2020). Terpenoid compounds can be stored constitutively in some species, often making up the majority of the volatile blend of a healthy unstressed plant; when plants are challenged, for example by herbivory, they will usually enhance the production/accumulation of the same compounds they store constitutively (Lange and Turner, 2013). The high levels of variation in terpenoid compounds across and within plant species has allowed for the practice of grouping plants by their chemical phenotype, an application known as chemotyping, which normally has a discrete and genetic basis (Despinasse et al., 2020).

High levels of intra-specific variation in constitutive volatile terpene compounds have been documented in several plant species. For instance, terpene chemotypes in the aromatic genus *Thymus* have been long recognised, and the genetic basis for them has been described (Trindade et al., 2018). In the conifer *Pinus pinaster*, where terpenes are stored in resin canals, two distinct chemotypes have been described (Arrabal et al., 2012). Additionally, some *Nepeta* species (including *Nepeta cataria*, commonly known as catnip) group into chemotypes, with three discrete groups distinguished by their volatile oil components

* Corresponding author.
 E-mail address: ted.turlings@unine.ch (T.C.J. Turlings).

(Gomes et al., 2021). In turn, chemotypic variation has been found to be correlated with herbivory levels. For instance, *Melaleuca alternifolia* plants exhibiting terpinolene chemotypes (i.e. high levels of terpinolene relative to other phenotypes) were found to suffer from less feeding damage by *Paropsisterna tigrina*, a specialist herbivore, than plants belonging to the eucalyptol chemotypes (Bustos-Segura et al., 2015). In addition, *Tanacetum vulgare* (tansy) plants in a field site grouped into four chemotypes, which were found to influence colonisation by aphids (Clancy et al., 2016). Furthermore, plant chemotype effects on herbivores can in turn affect higher trophic levels. Senft et al. (2019) compared the effects of these tansy chemotypes on aphid population dynamics and the abundance of patrolling ants through manipulation studies, and found that ants had a preference for aphids that were on specific chemotypes. How terpene profiles affect a plant's vulnerability to herbivores and interactions with other trophic levels can be of great importance to crop management.

Gossypium hirsutum Cav. (Malvaceae), commonly called 'upland cotton', is an important cash crop that is grown in arid and semi-arid regions worldwide. It is frequently attacked by multiple pest species (e.g. *Aphis gossypii* (cotton aphid), *Helicoverpa zea* (cotton bollworm), *Spodoptera exigua* (beet armyworm); Bottrell and Adkisson, 1977), and requires heavy use of pesticides, earning it the title of world's dirtiest crop (Environmental Justice Foundation in collaboration with Pesticide Action Network UK, 2007). Many efforts have been made to develop alternative strategies to handle pest management of *G. hirsutum*. Transgenic Bt lines expressing Cry-endotoxin were developed, however suppression of lepidopteran feeding made them more susceptible to other pest species. In particular, the reduction in damage by chewing herbivores on Bt lines led to a reduction of the induction of other defensive terpenoids, which favoured the cotton aphid (Hagenbucher et al., 2013). Interestingly, it has also been shown that resistance induced in *G. hirsutum* by phytophagous mites (*Tetranychus urticae* and *T. turkestani*) can lead to reduced mite population growth (Karban and Carey, 1984). In response to enemy attack, *Gossypium* spp. produces a wide range of protective/defensive compounds. Gossypol and associated terpenoid aldehydes are stored in specialised pigment glands in the leaves, stems, seeds, and flower buds. These glands are induced by damage and have been associated with resistance against noctuid moths (Agrawal and Karban, 2000). Other defences in cotton include the production of extrafloral nectaries (which recruit plant "bodyguards" such as ants (Reyes-Hernández et al., 2022)), and the induction of volatile emissions following herbivory (Loughrin et al., 1994). In damaged plants, the foliar volatile profile is comprised of compounds that are emitted both constitutively and in response to insect feeding, and can act as cues for natural enemies of herbivores, such as host-seeking parasitoids. It has been shown that specialist and generalist parasitoid species (*Microplitis croceipes* and *Cotesia marginiventris* respectively) are attracted to the volatiles released by damaged cotton plants (McCall et al., 1993). *G. hirsutum* plants constitutively produce a range of mono- and sesquiterpenoid compounds which are stored in the pigment glands. Low levels of these compounds are consistently emitted from undamaged plants (Paré and Tumlinson, 1997).

Though several *Gossypium* species have been independently domesticated in different regions of the world, wild populations persist to date. *Gossypium hirsutum* was domesticated in Mesoamerica, at least 4000 years ago (Yuan et al., 2021). Wild populations of *G. hirsutum* grow naturally along the coastal shrubland of the Yucatan Peninsula (Abdala-Roberts et al., 2019), which is also likely the centre of origin and domestication of this species (d'Eeckenbrugge and Lacape, 2014). In this study, we aimed to assess qualitative and geographic variation in constitutive stored volatile terpenes among wild *G. hirsutum* populations in this region, and to assess whether these plants can be grouped according to their chemical phenotype. To this end, we analysed the chemical profiles of plants grown from seeds collected from sixteen populations along the Yucatan Peninsula in Mexico. We also analysed the chemical constituents of the maternal plants in order to estimate heritability of the chemical phenotype.

2. Results and discussion

Due to seed dormancy and germination rates differing among populations, the number of replicates was not the same for all populations. The 16 populations were grouped into four geographic zones: Celestún (Far West), Sisal (West), Chicxulub (Centre), and Coloradas (East; see Fig. 1). In total, we analysed 349 individual plants and identified 44 compounds. Through comparison to database searches, authentic standards, and retention indices, 17 of the compounds were determined to be monoterpenes (MTs), and 19 sesquiterpenes (SQTs). Five green leaf volatiles (GLVs), two diterpenes, and one triterpene were also identified in the liquid extracts (Table 1, see Table S1 for absolute values; see Table S2 for full list of compounds).

2.1. Terpene chemistry and chemotyping of wild G. hirsutum populations

A correlation analysis (Fig. 2) performed on the relative abundance of mono- and sesquiterpenes in the leaf solvent extracts revealed two distinct highly supported (approximately unbiased (AU) p < 0.05) groups of monoterpenes. One group comprised the monoterpenes γ-terpinene, limonene, α-thujene, α-terpinene, terpinolene, and p-cymene (the "γ-terpinene group", designated chemotype class A), while the other consisted of the pinene-type compounds α- and β-pinene (the "α-pinene group", designated chemotype class B). The sesquiterpene compounds α-humulene and β-caryophyllene were also strongly intercorrelated. These strong correlations suggest a shared or highly linked biosynthesis. A theorised pathway of the synthesis of these compounds highlights their shared biochemical precursors (Degenhardt et al., 2009, Fig. 3). Taking the relative abundance of only monoterpenes, a negative relationship was observed between the proportions of γ-terpinene and α-pinene groups with two clusters of plants identified (Fig. 4a). Representative chromatograms from plants belonging to chemotype classes A and B respectively can be seen in Fig. 4b. Chemotype class A was composed of plants where the γ-terpinene group compounds made up 4% or more of the total monoterpene concentration; chemotype class B comprised plants where the γ-terpinene group compounds made up less than 4% of total monoterpenes (Fig. 4c). The major difference is the strong reduction or absence of the γ-terpinene cassette of monoterpenes in chemotype class B. In a principal coordinate analysis (PCoA) on relative abundances of monoterpenes, with samples separating based on their compositional similarity (using a Euclidean dissimilarity matrix), the chemotype classes separated along the first axis (Fig. 5), which explained ca. 65% of the variation. Plants belonging to chemotype class B group together, while plants in class A are spread along the first axis. Around 42% of overall variance in the monoterpene profiles of the plants was explained by chemotype class (adonis: R2 = 0.419, p = 0.001).

Most modern varieties of *G. hirsutum* produce and store a range of mono- and sesquiterpenes, including α- and β-pinene, limonene, camphene, β-myrcene, α-humulene, and β-caryophyllene (Elzen et al., 1985; Minyard et al., 1965; Yang et al., 2013), and plants producing significant amounts of γ-terpinene are rare. Only one naturalised variety has been reported to emit γ-terpinene in large quantities (Loughrin et al., 1995). Although modern cultivated varieties came from diverse landraces, this could indicate that the main sources for domestication events of *G. hirsutum* were obtained from B chemotype wild plants, or that perhaps the artificial selection throughout the domestication process selected against chemotype A plants due to pleiotropic interactions or linkage with traits of interest.

The large diversity of terpenoid compounds found in plants is due to terpene synthases (TPSs), which are a diverse family of enzymes that catalyse terpenoid compounds from single substrates (Karunanithi and Zerbe, 2019). The genes *GhTPS1* and *GhTPS2*, identified and characterised in *Gossypium hirsutum* (cultivar CCRI12) (Huang et al., 2013),
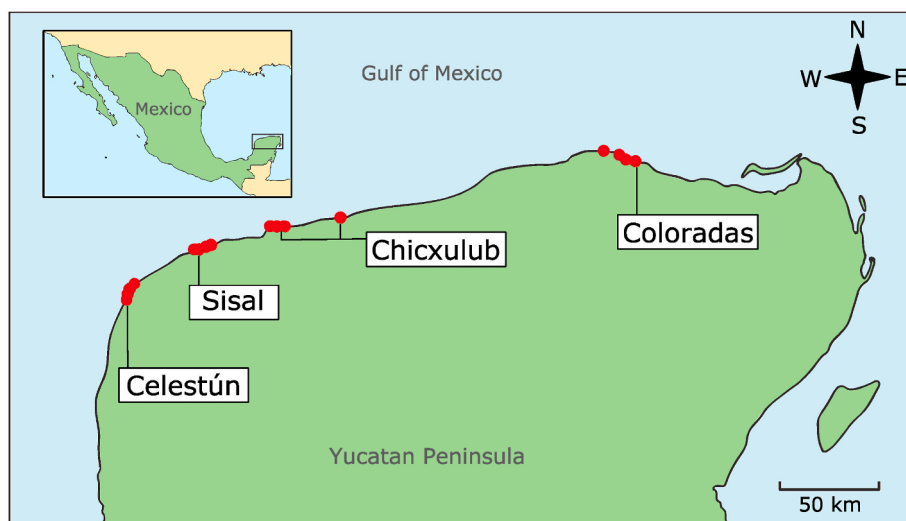
Fig. 1. Map showing locations of the wild *Gossypium hirsutum* populations from which seeds were collected along the Yucatan Peninsula.

Table 1
Relative abundance of monoterpenes and sesquiterpenes used in analyses grouped by geographic zones. Compounds marked with an asterisk were verified by comparison to authentic standards.

	Compound	RI	Celestún (Far West)		Sisal (West)		Chicxulub (Centre)		Coloradas (East)	
			Composition (%)	(N = 86)	Composition (%)	(N = 71)	Composition (%)	(N = 80)	Composition (%)	(N = 116)
			Mean	SE	Mean	SE	Mean	SE	Mean	SE
Monoterpenes	α-Thujene	929	0.1	0.0	0.2	0.0	0.3	0.0	0.4	0.0
	α-Pinene*	938	26.3	0.4	24.0	0.5	24.5	0.6	22.8	0.5
	Camphene	955	0.4	0.0	0.4	0.0	0.5	0.0	0.4	0.0
	Sabinene	977	1.4	0.1	1.4	0.1	1.5	0.1	1.3	0.0
	β-Pinene	983	4.8	0.1	4.6	0.1	4.8	0.1	4.4	0.1
	β-Myrcene*	989	20.9	0.5	17.4	0.6	20.4	0.6	18.7	0.4
	α-Phellandrene	1008	0.1	0.0	0.1	0.0	0.2	0.0	0.2	0.0
	α-Terpinene	1020	0.0	0.0	0.1	0.0	0.2	0.0	0.2	0.0
	p-Cymene	1028	0.0	0.0	0.2	0.0	0.2	0.0	0.2	0.0
	Limonene*	1033	1.6	0.1	2.1	0.2	2.7	0.2	3.0	0.1
	β-Phellandrene	1035	1.2	0.0	1.0	0.1	0.9	0.1	0.7	0.1
	β-Ocimene	1047	2.5	0.3	3.2	0.3	3.0	0.2	4.5	0.2
	Isoterpinolene	1057	0.1	0.0	0.1	0.0	0.1	0.0	0.0	0.0
	γ-Terpinene*	1063	0.7	0.2	2.7	0.5	4.4	0.7	5.9	0.4
	Terpinolene	1090	0.0	0.0	0.2	0.0	0.3	0.0	0.3	0.0
	Citronellal	1155	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Bornyl acetate	1288	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
Sesquiterpenes	γ-Elemene	1347	0.8	0.0	0.9	0.0	0.8	0.0	0.6	0.0
	α-Cubebene	1359	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	α-Copaene	1389	0.2	0.0	0.4	0.1	0.1	0.0	0.2	0.0
	β-Elemene	1401	0.1	0.0	0.2	0.0	0.1	0.0	0.1	0.0
	α-Gurjunene	1425	0.1	0.0	0.1	0.0	0.1	0.0	0.1	0.0
	β-Caryophyllene*	1432	23.5	0.4	22.6	0.6	19.5	0.5	22.6	0.3
	β-Gurjunene	1448	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Alloaromadendrene	1453	0.4	0.0	0.6	0.1	0.5	0.0	0.3	0.0
	Humulene	1465	6.2	0.1	6.1	0.2	5.3	0.2	5.9	0.1
	Unknown #1	1473	0.3	0.0	0.3	0.0	0.4	0.0	0.2	0.0
	Bicyclogermacrene	1498	6.8	0.4	8.5	0.5	7.7	0.5	5.3	0.2
	β-Acorneol	1519	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	β-Cadinene	1524	0.3	0.0	0.5	0.0	0.3	0.0	0.2	0.0
	Germacrene D-4-ol	1603	0.1	0.0	0.2	0.0	0.1	0.0	0.0	0.0
	Spathulenol	1610	0.1	0.0	0.1	0.0	0.1	0.0	0.1	0.0
	Unknown #2	1628	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	Unknown #3	1668	0.1	0.0	0.2	0.0	0.1	0.0	0.0	0.0
	Farnesol (E,E)	1749	0.6	0.0	0.9	0.1	0.4	0.0	0.6	0.0
	Unknown #4	2033	0.3	0.0	0.5	0.0	0.4	0.0	0.5	0.0

encode for active TPSs that are expressed in young developing leaves, and may be involved in the production of constitutive terpenoids that are stored in glands. It has been found that GhTPS1 mainly produces the sesquiterpenes β-caryophyllene and humulene at a ratio of ∼4:1.

GhTPS2 was found to mainly produce the monoterpenes α-pinene and β-pinene, at the ratio of ∼6:1. These sets of compounds were highly correlated in our dataset, with β-caryophyllene and humulene being present at a ratio of 3.8:1 (SE = 0.013), and α-pinene and β-pinene at a

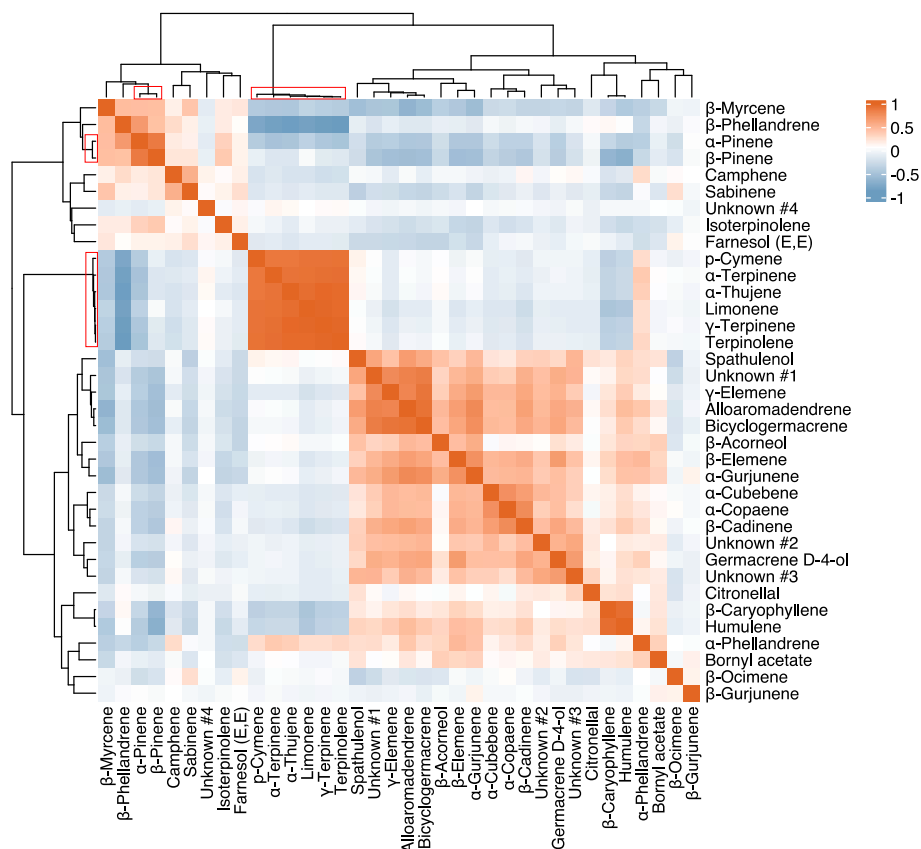**Fig. 2.** Correlation analysis of all mono- and sesquiterpenes analysed in the wild *Gossypium hirsutum* plants. Red rectangles indicate highly supported groups of monoterpenes (approximately unbiased (AU) p < 0.05). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

ratio of 5.3:1 (SE = 0.027). The highly similar product ratios between our dataset and those described in Huang et al. (2013), suggests that GhTPS1 and GhTPS2 contribute to the production of these compounds in wild *G. hirsutum*. Nonetheless, other TPSs may also be involved, as multiple *G. hirsutum* TPSs have been shown to produce several terpenes at varying levels (including the compounds that comprise the "γ -terpinene group" (Huang et al., 2018)). Eighty-five TPS genes have been putatively identified in *G. hirsutum* to date (Zhang et al., 2022). Our results suggest that one or more enzymes present or overexpressed in plants of chemotype A are multifunctional TPSs that synthesise the monoterpenes of the γ-terpinene group.

2.2. Association between chemotype and geographic location

The correlation between geographic location of the sites and chemical distance was not very strong but significantly different from zero (Mantel test, r = 0.165, p = 0.001), indicating that as geographic distance increased, the chemical profiles of the plants became more dissimilar. An association between site location and γ-terpinene content was observed, with more plants in western sites containing low amounts of compounds belonging to the γ-terpinene group, and plants in eastern sites containing higher proportions of γ-terpinene group compounds (Fig. 6). Future studies could examine the genetic structuring of these populations to determine if these chemotypic differences are related to genetic differentiation (i.e., isolation among populations) and/or adaptation to their local environment.

2.3. Chemotypes of mature plants

In addition to the plants grown from seed, we also assessed the chemical profiles of 165 mature plants from which a subset (34) were

also maternal plants. Thirty-seven compounds were detected in the mature plants, with 20 compounds identified as MTs, 10 SQTs, 4 GLVs, 1 hydrocarbon, and 2 nitrogen-containing volatiles (see Table S3). Mature plants were classified using the same chemotyping process as described above. PcoA ordination of all mature plants harvested *in situ* shows grouping by chemotype class, separating along axis 1 (Fig. S1; variation explained by first two ordinations: 82.6% and 15%). More than two-thirds of the variation was explained by chemotype group (adonis: R2 = 0.698, p = 0.001). Terpenoid production is known to be affected by a variety of environmental stresses, including exposure to extreme temperatures, water stress (flooding, drought), and salinity, as well as herbivory (Gouinguené and Turlings, 2002; Isah, 2019). A differential regulation in terpenoid production because of environmental pressures could explain why chemotype explained a higher percentage of total variation in the chemical profiles of the mature plants compared to plantlets; abiotic and biotic stress are known to affect the levels of VOCs in plants (Holopainen and Gershenzon, 2010). Plant ontogeny can also influence the terpenoid profile. Although the age of the mature plants was unknown, they had been surveyed for several years and were well established flowering plants. Meanwhile, offspring plants were seedlings with only four developed leaves (around one month old), and it is possible that mature *G. hirsutum* plants exhibit a more defined chemotype. For example, in *Melaleuca alternifolia*, six terpene chemotypes can be clearly differentiated in mature trees, however young plants only have four identifiable chemotypes, and two chemotypes with high terpinolene concentration are indistinguishable in young plants (Bustos-Segura et al., 2015).

2.4. Heritability of chemotypes

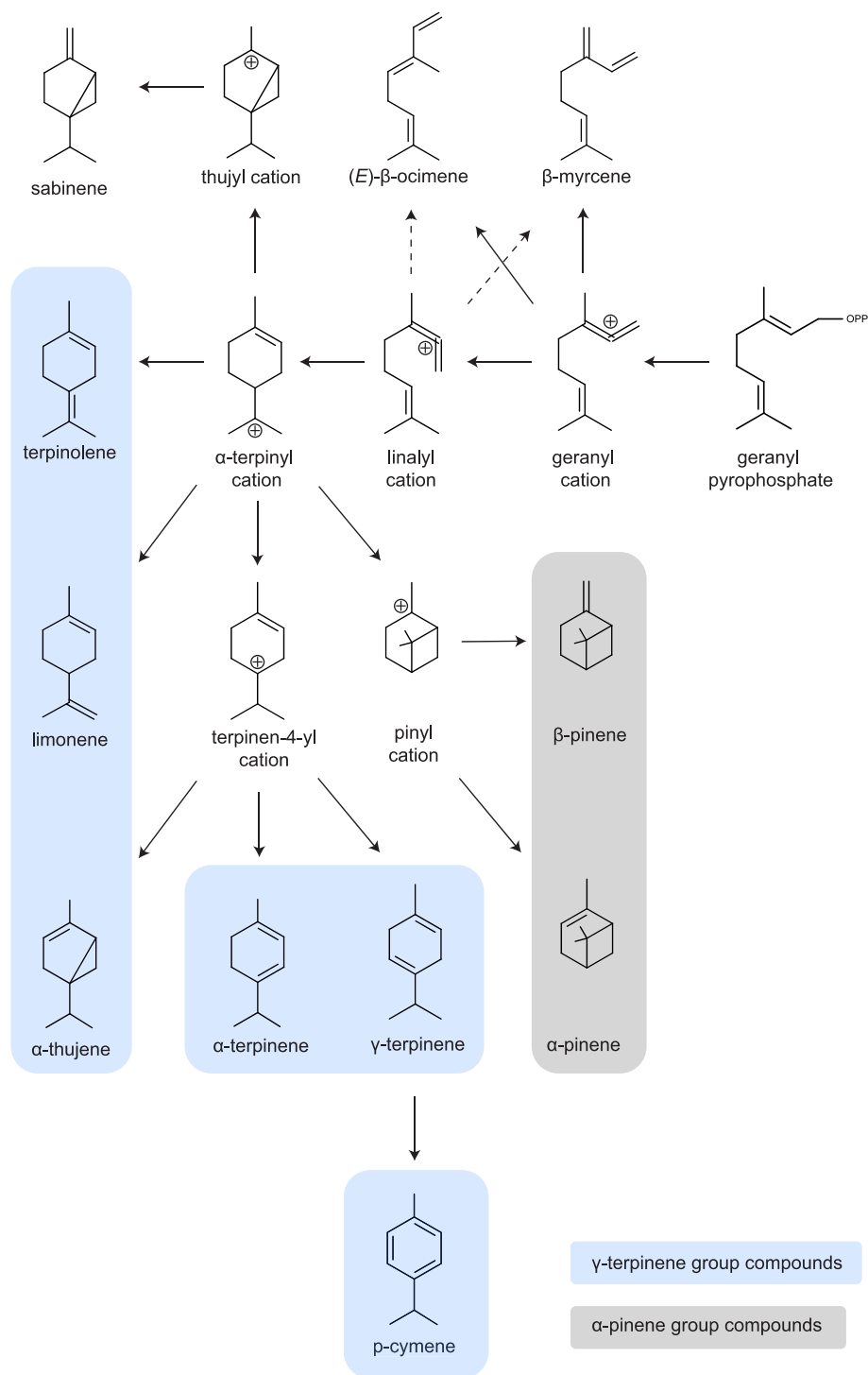Most of the progeny had the same chemotype as their parent plant.

Fig. 3. Theorised monoterpenoid biosynthesis pathway. Compounds highlighted in blue comprise the γ-terpinene compound group, those highlighted in grey comprise the α-pinene compound group. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Offspring of plants with chemotype class A matched their maternal plant 80% of the time (63/79 plants), and plants in chemotype class B matched their maternal plant almost 93% of the time (64/69). In wild *G. hirsutum* populations, both self-pollination and cross-pollination have been described by Wegier et al. (2011). Velázquez-López et al. (2018) reported an outcrossing rate of 0.72 for a metapopulation of wild *G. hirsutum* plants from the Yucatan Peninsula, indicating that although the level of self-pollination is important, there is also a large contribution from cross-pollination to the next generation. Thus, as we only knew the chemotype of the maternal plant and not the paternal plant,

this could explain why there was some disparity between maternal and offspring chemotypes. Broad sense heritability estimated with the animal model was 0.85 (CI = 0.75, 0.95). In concordance with what we observed, chemotype is generally considered to be heritable in plants (Hare, 2011). In the tree species *Eucalyptus globulus*, O'Reilly-Wapstra et al. (2011) found moderate to high broad-sense heritability of foliar terpenes, and Karban et al. (2014) found that in sagebrush (*Artemisia tridentata*) chemotypes are highly heritable between parent and offspring.
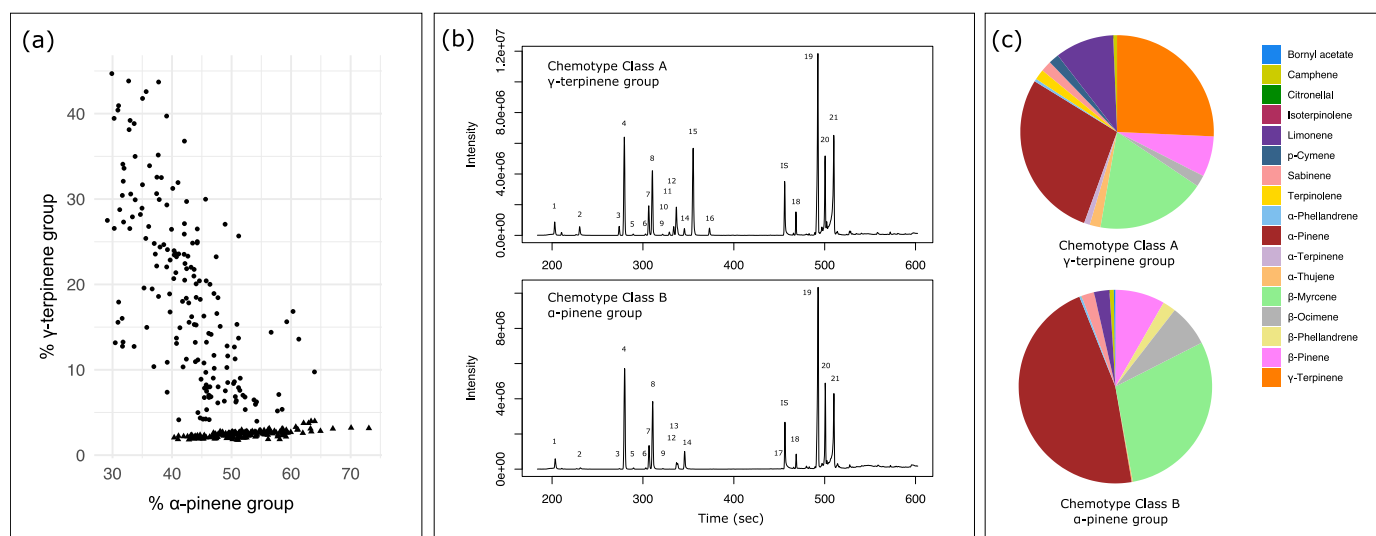
Fig. 4. Monoterpene composition of the two proposed cotton chemotypes. **(a)** Scatterplot of the relative abundance of the pinene group (α-Pinene + β-Pinene) against the relative abundance of the summed values of the γ-Terpinene group. **(b)** Representative chromatograms of plants belonging to chemotype classes A (upper panel) and B (lower panel). 1: (E)-2-Hexenal; 2: 2,4-Hexadienal (E,E); 3: α-Thujene; 4: α-Pinene; 5: Camphene; 6: Sabinene; 7: β-Pinene; 8: β-Myrcene; 9: α-Phellandrene; 10: α-Terpinene; 11: p-Cymene; 12: Limonene; 13: β-Phellandrene; 14: β-Ocimene; 15: γ-Terpinene; 16: Terpinolene; 17: Bornyl acetate; 18: γ-Elemene; 19: β-Caryophyllene; 20: α-Humulene; 21: Bicyclogermacrene. **(c)** visualisation of proportion of monoterpene compounds from panel (b), for each chemotype class.
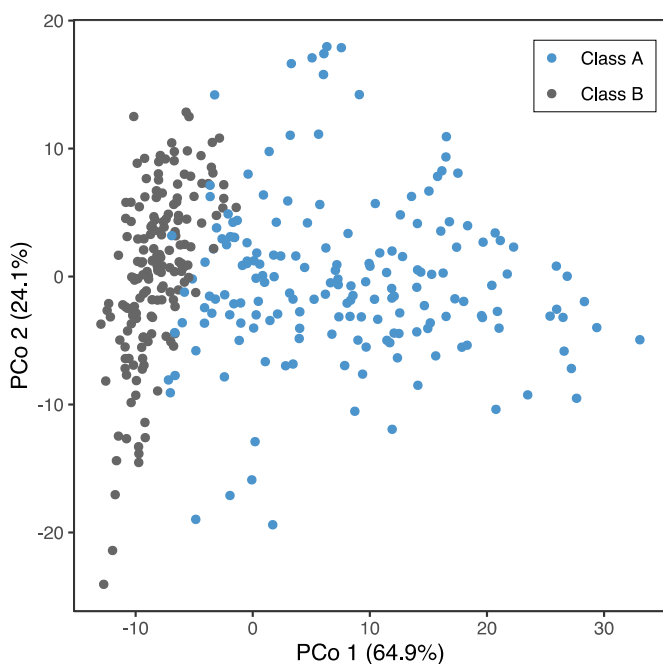

Fig. 5. Principal coordinate analysis on relative abundance of monoterpenes in plants grown from seed, showing samples separated based on their compositional similarity.

2.5. Concentration of leaf terpenoids

To determine the degree to which biological and geographic factors can explain the variation in the concentration of compound classes, we used generalised linear mixed models. Approximately a quarter of the total variance (from 21 to 29%) in concentrations of total volatiles, total monoterpenes, total sesquiterpenes, and total green leaf volatiles was significantly explained by 'genotype' (Table 2), whereas 'population' explained only a small fraction of the variance and was not statistically significant. Concentrations of total volatiles and total monoterpenes were significantly higher in plants belonging to chemotype Class A
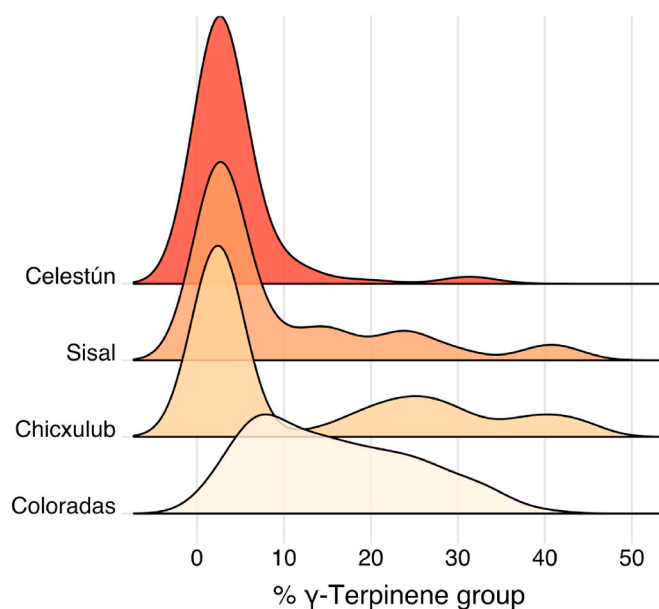

Fig. 6. Ridgeline plot showing the distribution of the summed values of the γ-terpinene compound group (γ-terpinene, limonene, α-thujene, α-terpinene, terpinolene, and p-cymene; as % relative to total monoterpenes in each plant). In order from top to bottom: the plots coloured red (Celestún) and orange (Sisal) are located at the west of the peninsula. The plot coloured yellow (Chicxulub) is in the centre, and the cream coloured plot (Coloradas) is located at the east of the peninsula. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

(VOCs: $\chi^2_{(1)} = 4.9101$, p = 0.027, MTs: $\chi^2_{(1)} = 8.298$, p = 0.004). No significant difference was observed in total concentrations of sesquiterpenes or green leaf volatiles between chemotypes (SQTs: $\chi^2_{(1)} = 1.5618$, p = 0.2114, GLVs: $\chi^2_{(1)} = 1.646$, p = 0.1995).

At the time of harvesting, all leaves were at a similar developmental stage (4th leaf fully unfurled). However, leaf area showed considerable variation with a bimodal distribution (excess mass test, p = 0.004, Fig. S2); therefore, all analyses on concentration included leaf area as a

Table 2
Variance between genotype and population random factors. The percentage of total variance explained is also shown. Variance components were obtained from generalised linear mixed models evaluating the variance of total concentrations of all volatiles, total monoterpenes, total sesquiterpenes, and total green leaf volatiles. Values shown in brackets specify the standard deviation of each component. Significant values are shown in bold. ***p < 0.0001.

Parameter	Genotype	Population	Residual variance	Genotype (%)	Population (%)	Total Variance
Total VOC	**0.03*****	0.01	0.09	23.20	5.10	0.13
	(0.16)	(0.08)	(0.3)			
Total MT	**0.04*****	0.01	0.13	22.22	5.56	0.18
	(0.20)	(0.09)	(0.37)			
Total SQT	**0.06*****	0.02	0.20	21.43	7.14	0.08
	(0.25)	(0.12)	(0.45)			
Total GLV	**0.03*****	0.00	0.07	29.23	4.61	0.10
	(0.17)	(0.07)	(0.26)			

covariate. Plants with smaller leaves contained significantly higher summed concentrations of all compounds found (summed GLVs, MTs and SQTs; $\chi^2_{(1)} = 184.9882$, p < 0.001). Total monoterpene and sesquiterpene concentrations were significantly higher in smaller leaves (MTs: $\chi^2_{(1)} = 585.716$, $p < 0.001$, SQTs: $\chi^2_{(1)} = 339.2854$, p < 0.001), whereas total concentrations of green leaf volatiles were higher in larger leaves (GLVs: $\chi^2_{(1)} = 51.733$, p < 0.001). Opitz and colleagues had found that accumulation of terpenoids in glands in cotton was strongly influenced by leaf area and developmental stage, with the two youngest/smallest leaves having the highest total terpenoid concentration (Opitz et al., 2008). They also showed that younger leaves have a much higher gland density than older leaves, but contain less total terpenoids per gland. Eisenring et al. (2017) made the observation that smaller leaves have a higher density of glands than larger leaves of the same age. We found that smaller leaves contained higher concentrations of MTs and SQTs and lower concentrations of GLVs in relation to larger leaves.

2.6. Ecological implications of chemical diversity in G. hirsutum

The chemotype of a plant plays an important role in the outcome of environmental adaptation, and volatile-mediated plant-insect and plant-plant interactions. For instance, in *Artemisia tridentata* (sagebrush), plants with the same chemotype respond more strongly to volatiles emitted by each other than to volatiles from individuals with a different chemotype (Karban et al., 2014). Moreover, when *Tanacetum vulgare* (tansy) plants were challenged by insect herbivores from two feeding guilds, distinct chemotypes responded differently. Following caterpillar feeding, some chemotypes exhibit a stronger volatile response when also pre-treated with aphids, while the opposite was observed for other chemotypes (Clancy et al., 2020). In *Thymus vulgaris* (thyme), phenolic chemotypes show reduced tolerance to freezing compared to non-phenolic chemotypes (Thompson et al., 2013) and occur more frequently in a region that has undergone regional warming (where occurrence of extreme winter freezing events has dwindled). Accordingly, the observed geographic gradient in the frequency of the two identified chemotypes across wild *G. hirsutum* populations calls for further work testing whether the terpene chemotypes are associated with a pattern of local adaptation to biotic (herbivores and/or pathogens) or abiotic factors.

There have been numerous investigations concerned with describing the volatile terpenoids found in *G. hirsutum*. While chemotypes have not, to our knowledge, been described in wild cotton to date, differences in the amounts of herbivory-induced plant volatiles (HIPVs) released by different cultivars have been recorded (Hagenbucher et al., 2016). Magalhães et al. (2020) investigated VOC differences between seven *G. hirsutum* genotypes and found few qualitative and quantitative differences. Röse and colleagues (Röse and Tumlinson, 2005) showed that cotton plants respond in specific ways to herbivory by systemically releasing distinct blends of volatile compounds, and suggest that plants that emit higher amounts of induced defence chemicals are better at defending themselves. Loughrin et al. (1995) showed that upon feeding damage by beet armyworm larvae (*Spodoptera exigua*), the naturalised

variety TX2259 was found to emit much greater quantities of HIPVs than five other cultivated lines (almost sevenfold higher). Moreover, this variety released higher amounts of γ-terpinene. Thus, it would be interesting to explore if wild (and naturalised) cotton emits higher amounts of HIPVs than cultivated plants of *G. hirsutum*, and whether the constitutive terpene chemotype is also linked to quantitative differences in the volatile emissions.

HIPVs in *G. hirsutum* have also been implicated in interactions with the third trophic level. *Campoletis sonorensis* and *Microplitis croceipes* (two parasitoid species) were found to respond to volatiles from undamaged cotton plants (Elzen et al., 1987), yet volatiles emitted by caterpillar-damaged plants seem to provoke a considerably stronger response (De Moraes et al., 1998; Turlings et al., 1995). Importantly, cotton volatiles are also involved in plant-plant signalling; studies have shown that cotton plants exposed to the volatiles of a nearby damaged plant are more resistant to herbivory (Llandres et al., 2018; Renou et al., 2011). Therefore, chemotypic variation could be a factor that not only affects the attraction of natural enemies of herbivores to damaged or undamaged plants but also be a key determinant of the effectiveness of airborne signalling between plants.

3. Conclusions

In brief, we demonstrate that wild *G. hirsutum* plants from 16 naturally occurring populations along the Yucatan peninsula could be grouped into two chemotypes based on differences in their stored monoterpene profiles. The identification of chemotypes in wild *G. hirsutum* could have implications for cotton direct defences against herbivory, and indirect defences through the attraction of natural enemies or plant-plant signalling. We recommend performing experiments exploring the consequences of the reported terpene variation on the different levels of plant defence and its effects on the outcome of associated interactions. In combination with functional characterisation of terpene synthases in wild cotton chemotypes such research would provide valuable information into the chemical ecology of cotton and may reveal opportunities to enhance the control of cotton pests without the current excessive use of pesticides.

4. Experimental

4.1. Plant material

Gossypium hirsutum Cav. (Malvaceae) is a perennial shrub that is native to Mexico and is distributed throughout the Caribbean Basin and Central America (d'Eeckenbrugge and Lacape, 2014). It grows up to 2 m tall under natural conditions. Wild populations of the species can be found growing in coastal shrubland along the Mexican Yucatan Peninsula coast. Seeds from 16 populations of wild *G. hirsutum* plants located around the Yucatan Peninsula (Fig. 1) were collected in 2019 and 2020. From each population, approximately 100 seeds from six maternal plants were harvested at each site. Offspring from the same mother are either full or half-siblings and referred to here as a genotype. The seeds

were shipped to the University of Neuchâtel, Switzerland, where the experiment was performed between July–November 2020.

4.2. Experimental protocol

4.2.1. Greenhouse seedling experiment

To improve germination rates, the seeds were pre-treated as follows before planting: seeds were delinted and carefully scratched at the chalazal end using a mini drill (Dremel 3000, WI, USA) with a sandpaper adapter. They were then soaked in tap water at 35 °C for 19 h before being planted in trays containing soil (Profi Substrat, Einheitserde, Germany). The trays were kept in phytotrons (GroBanks CLF Plant Climatics, Germany) under the following conditions until seedlings emerged; 14 h: 30 °C, 65 μmol m^{-2} s^{-1}, and 10 h at 25 °C in the dark. Once cotyledons were fully developed, seedlings were transplanted to individual pots (250 ml) and moved to the greenhouse, where they were placed randomly. Plants were then grown until the fourth true leaf had emerged and was fully developed.

4.2.2. Sampling of mature plants in situ

In the winter of 2020, the leaves of 165 mature plants from 14 of the 16 populations previously described, plus three additional populations, were sampled *in situ*. Of these, 34 plants from 11 populations produced seeds that were used in the main experiment. As some plants were lost to flooding, hurricanes, and urbanisation, not all maternal plants could be sampled. We only harvested undamaged leaves which were immediately frozen on dry ice upon collection. The frozen samples were ground to a powder under liquid nitrogen and stored at −80 °C in the Biotechnology Laboratory at the Faculty of Chemical Engineering of the Autonomous University of Yucatan (UADY). In 2021, the frozen samples were transported to the University of Neuchâtel, where hexane extraction and GC-MS analysis were performed as described below.

4.3. Sample preparation and GC-MS analysis

The fourth leaf of each plant was harvested, photographed, and weighed, then immediately frozen in liquid nitrogen. The frozen leaf samples were then ground into a powder and stored at −80 °C. For each sample, one ml of hexane was added to approximately 100 mg frozen leaf material and then vortexed and stored at 4 °C for 24 h. The hexane extract (800 μl) was removed and transferred to an amber glass 1.5 ml vial with screw top PTFE/silicone lid. Nonyl acetate was added as an internal standard (200 ng in 10 μl).

The samples were analysed by GC-MS (GC type: 8890A, MS type: 5977B MSD, both Agilent Technologies, Palo Alto, CA, USA), using a HP-5MS capillary column (30 m × 250 μm × 0.25 μm, Agilent Technologies). 1.5 μl of sample was autoinjected in splitless mode (inlet at 250°C) at a constant flow rate of He at 0.9 ml min^{-1} at the following temperature program: 40 °C for 0 min, ramping at 15 °C min^{-1} to 130 °C, then 60 °C min^{-1} to 190 °C, then 30 °C min^{-1} to 290 °C and holding for 0.5 min. Identification of all measured VOCs was carried out through comparison of obtained mass spectra with those of authentic commercially available standards, NIST version 2.3 library spectra, an in-house VOC mass spectra library, and the Kovats retention index library (Lucero et al., 2009). Peaks were aligned and quantified using the PyMassSpec package in Python forked from the PyMS repository (O'Callaghan et al., 2012).

4.4. Statistical analysis

Analyses were performed using the relative abundance of mono-terpenes and sesquiterpenes together, unless explicitly stated otherwise. Statistical analyses were performed using R (version 4.1.3). Hierarchical clustering of mono- and sesquiterpenes was performed in pvclust (Suzuki and Shimodaira, 2006) using the correlation distance method, Ward.D2 method, and 10,000 bootstrap replications to identify

compounds that were highly correlated in order to distinguish chemotype groupings (Bustos-Segura et al., 2017). The correlation heatmap was created using the ComplexHeatmap R package (Gu et al., 2016). We performed Principal Coordinate Analysis using functions from the R packages vegan and stats (vegdist, cmdscale (Oksanen et al., 2020; R Core Team, 2013);) using relative abundances of mono- and sesquiterpenes. We assessed the effect of geographic location on plant chemical profiles using a Mantel test (R package 'vegan'), which tested the correlation between the spatial and chemical distances (Haversine distance and Bray-Curtis dissimilarity). Permutational multivariate analysis of variance (PERMANOVA) using the 'adonis' function (R package 'vegan') with 999 permutations was used to test for differences between chemotypes. Distribution analysis of leaf area was performed using the R package 'multimode' (Ameijeiras-Alonso et al., 2021), and ridgeline plots were produced using 'ggridges' (R package ggridges (Wilke, 2021);). Generalised linear mixed-effects models were performed using the R package 'lme4' (Bates et al., 2015), using a Gamma distribution with a log link function to analyse the variation in VOCs concentration. Concentrations of total terpenes, total monoterpenes, total sesquiterpenes or total GLVs were used as dependent variables, and we assessed the effects of chemotype class and leaf area (cm^2) as a covariate (fixed effects); population and genotype were included as random effects. Likelihood ratio tests and the Akaike information criterion were used to determine which random factors significantly affected the model. Heritability was estimated with an animal model using a generalised linear mixed model with a Bayesian approach (R package 'MCMCglmm'; Hadfield, 2010). The model included chemotype as a response binomial variable (chemotype class A or B) and the pedigree of maternal and offspring plants as a random factor. A threshold distribution family was used (de Villemereuil, 2018). We ran MCMC for 2×10^6 iterations and thinning interval of 200 after a burn-in of 1×10^5 which yielded an effective sample size of 8075. Heritability and confidence intervals were estimated using the posterior distribution of the variance components from the model (Falconer and Mackay, 1996):

$$H^2 = \frac{V_A}{V_A + V_R}$$

where H^2: broad sense heritability; V_A: additive variance; V_R: residual variance.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.phytochem.2022.113454.

References

Abdala-Roberts, L., Quijano-Medina, T., Moreira, X., Vázquez-González, C., Parra-Tabla, V., Berny Mier Y Terán, J.C., Grandi, L., Glauser, G., Turlings, T.C.J., Benrey, B., 2019. Bottom-up control of geographic variation in insect herbivory on wild cotton (*Gossypium hirsutum*) by plant defenses and climate. Am. J. Bot. 106, 1059–1067. https://doi.org/10.1002/ajb2.1330.

Agrawal, A., Karban, R., 2000. Specificity of constitutive and induced resistance: pigment glands influence mites and caterpillars on cotton plants. Entomol. Exp. Appl. 96, 39–49.

Ameijeiras-Alonso, J., Crujeiras, R.M., Rodriguez-Casal, A., 2021. Multimode: an R package for mode assessment. J. Stat. Software 97, 1–32. https://doi.org/10.18637/jss.v097.i09.

Arrabal, C., García-Vallejo, M.C., Cadahia, E., Cortijo, M., de Simón, B.F., 2012. Characterization of two chemotypes of *Pinus pinaster* by their terpene and acid patterns in needles. Plant Systemat. Evol. 298, 511–522. https://doi.org/10.1007/s00606-011-0562-8.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Software 67, 1–48. https://doi.org/10.18637/jss.v067.i01.

Bottrell, D.G., Adkisson, P.L., 1977. Cotton insect pest management. Annu. Rev. Entomol. 22, 451–481.

Bustos-Segura, C., Dillon, S., Keszei, A., Foley, W.J., Külheim, C., 2017. Intraspecific diversity of terpenes of *Eucalyptus camaldulensis* (Myrtaceae) at a continental scale. Aust. J. Bot. 65, 257–269.

Bustos-Segura, C., Külheim, C., Foley, W., 2015. Effects of terpene chemotypes of *Melaleuca alternifolia* on two specialist leaf beetles and susceptibility to myrtle rust. J. Chem. Ecol. 41, 937–947. https://doi.org/10.1007/s10886-015-0628-0.

Christianson, D.W., 2017. Structural and chemical biology of terpenoid cyclases. Chem. Rev. 117, 11570–11648. https://doi.org/10.1021/acs.chemrev.7b00287.

Clancy, M.V., Haberer, G., Jud, W., Niederbacher, B., Niederbacher, S., Senft, M., Zytynska, S.E., Weisser, W.W., Schnitzler, J.-P., 2020. Under fire-simultaneous volatilome and transcriptome analysis unravels fine-scale responses of tansy chemotypes to dual herbivore attack. BMC Plant Biol. 20, 1–18.

Clancy, M.V., Zytynska, S.E., Senft, M., Weisser, W.W., Schnitzler, J.-P., 2016. Chemotypic variation in terpenes emitted from storage pools influences early aphid colonisation on tansy. Sci. Rep. 6, 1–12.

d'Eeckenbrugge, G.C., Lacape, J.-M., 2014. Distribution and differentiation of wild, feral, and cultivated populations of perennial upland cotton (*Gossypium hirsutum* L.) in Mesoamerica and the Caribbean. PLoS One 9.

De Moraes, C.M., Lewis, W.J., Paré, P.W., Alborn, H.T., Tumlinson, J.H., 1998. Herbivore-infested plants selectively attract parasitoids. Nature 393, 570–573. https://doi.org/10.1038/31219.

de Villemereuil, P., 2018. Quantitative genetic methods depending on the nature of the phenotypic trait. Ann. N. Y. Acad. Sci. 1422, 29–47. https://doi.org/10.1111/nyas.13571.

Despinasse, Y., Moja, S., Soler, C., Jullien, F., Pasquier, B., Bessière, J.-M., Baudino, S., Nicolè, F., 2020. Structure of the chemical and genetic diversity of the true lavender over its natural range. Plants 9. https://doi.org/10.3390/plants9121640.

Eisenring, M., Meissle, M., Hagenbucher, S., Naranjo, S.E., Wettstein, F., Romeis, J., 2017. Cotton defense induction patterns under spatially, temporally and quantitatively varying herbivory levels. Front. Plant Sci. 8.

Elzen, G.W., Williams, H.J., Bell, A.A., Stipanovic, R.D., Vinson, S.B., 1985. Quantification of volatile terpenes of glanded and glandless *Gossypium hirsutum* L. cultivars and lines by gas chromatography. J. Agric. Food Chem. 33, 1079–1082. https://doi.org/10.1021/jf00066a015.

Elzen, G.W., Williams, H.J., Vinson, S.B., Powell, J.E., 1987. Comparative flight behavior of parasitoids *Campoletis sonorensis* and *Microplitis croceipes*. Entomol. Exp. Appl. 45, 175–180. https://doi.org/10.1111/j.1570-7458.1987.tb01078.x.

Environmental Justice Foundation in collaboration with Pesticide Action Network UK, 2007. The Deadly Chemicals in Cotton.

Falconer, D.S., Mackay, T.F.C., 1996. Introduction to Quantitative Genetics. UK Longman Group, Essex.

Gershenzon, J., Dudareva, N., 2007. The function of terpene natural products in the natural world. Nat. Chem. Biol. 3, 408–414. https://doi.org/10.1038/nchembio.2007.5.

Gomes, E.N., Zorde, M., Patel, H., Lyu, W., Wu, Q., Simon, J.E., 2021. Chemodiversity in *Nepeta* spp.: a literature review on comparative germplasm studies with focus on iridoids and other terpenes. J. Med. Act. Plants 10, 82–115.

Gouinguené, S.P., Turlings, T.C., 2002. The effects of abiotic factors on induced volatile emissions in corn plants. Plant Physiol. 129, 1296–1307.

Gu, Z., Eils, R., Schlesner, M., 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32, 2847–2849. https://doi.org/10.1093/bioinformatics/btw313.

Hadfield, J.D., 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. J. Stat. Software 33, 1–22. https://doi.org/10.18637/jss.v033.i02.

Hagenbucher, S., Birgersson, G., Chattington, S., Anderson, P., 2016. Domestication influences choice behavior and performance of a generalist herbivore. Perspect. Plant Ecol. Evol. Systemat. 23, 63–72. https://doi.org/10.1016/j.ppees.2016.09.001.

Hagenbucher, S., Wäckers, F.L., Wettstein, F.E., Olson, D.M., Ruberson, J.R., Romeis, J., 2013. Pest trade-offs in technology: reduced damage by caterpillars in Bt cotton benefits aphids. Proc. R. Soc. B Biol. Sci. 280, 20130042 https://doi.org/10.1098/rspb.2013.0042.

Hare, J.D., 2011. Ecological role of volatiles produced by plants in response to damage by herbivorous insects. Annu. Rev. Entomol. 56, 161–180. https://doi.org/10.1146/annurev-ento-120709-144753.

Holopainen, J.K., Gershenzon, J., 2010. Multiple stress factors and the emission of plant VOCs. Spec. Issue Induc. Biog. Volatile Org. Compd. Plants 15, 176–184. https://doi.org/10.1016/j.tplants.2010.01.006.

Huang, X., Xiao, Y., Köllner, T.G., Zhang, W., Wu, Junxiang, Wu, Juan, Guo, Y., Zhang, Y., 2013. Identification and characterization of (E)-β-caryophyllene synthase and α/β-pinene synthase potentially involved in constitutive and herbivore-induced terpene formation in cotton. Plant Physiol. Biochem. 73, 302–308. https://doi.org/10.1016/j.plaphy.2013.10.017.

Huang, X.-Z., Xiao, Y.-T., Köllner, T.G., Jing, W.-X., Kou, J.-F., Chen, J.-Y., Liu, D.-F., Gu, S.-H., Wu, J.-X., Zhang, Y.-J., Guo, Y.-Y., 2018. The terpene synthase gene family in *Gossypium hirsutum* harbors a linalool synthase GhTPS12 implicated in direct defence responses against herbivores. Plant Cell Environ. 41, 261–274. https://doi.org/10.1111/pce.13088.

Isah, T., 2019. Stress and defense responses in plant secondary metabolites production. Biol. Res. 52, 39. https://doi.org/10.1186/s40659-019-0246-3.

Karban, R., Carey, J.R., 1984. Induced resistance of cotton seedlings to mites. Science 225, 53–54. https://doi.org/10.1126/science.225.4657.53.

Karban, R., Wetzel, W.C., Shiojiri, K., Ishizaki, S., Ramirez, S.R., Blande, J.D., 2014. Deciphering the language of plant communication: volatile chemotypes of sagebrush. New Phytol. 204, 380–385.

Karunanithi, P.S., Zerbe, P., 2019. Terpene synthases as metabolic gatekeepers in the evolution of plant terpenoid chemical diversity. Front. Plant Sci. 10.

Lange, M.B., Turner, G.W., 2013. Terpenoid biosynthesis in trichomes—current status and future opportunities. Plant Biotechnol. J. 11, 2–22. https://doi.org/10.1111/j.1467-7652.2012.00737.x.

Llandres, A.L., Almohamad, R., Brévault, T., Renou, A., Téréta, I., Jean, J., Goebel, F.-R., 2018. Plant training for induced defense against insect pests: a promising tool for integrated pest management in cotton. Pest Manag. Sci. 74, 2004–2012.

Loughrin, J.H., Manukian, A., Heath, R.R., Tumlinson, J.H., 1995. Volatiles emitted by different cotton varieties damaged by feeding beet armyworm larvae. J. Chem. Ecol. 21, 1217–1227.

Loughrin, J.H., Manukian, A.R.A., Heath, R.R., Turlings, T.C., Tumlinson, J.H., 1994. Diurnal cycle of emission of induced volatile terpenoids by herbivore-injured cotton plant. Proc. Natl. Acad. Sci. 91, 11836–11840.

Lucero, M., Estell, R., Tellez, M., Fredrickson, E., 2009. A retention index calculator simplifies identification of plant volatile organic compounds. Phytochem. Anal. 20, 378–384. https://doi.org/10.1002/pca.1137.

Magalhães, D.M., Borges, M., Laumann, R.A., Caulfield, J.C., Birkett, M.A., Blassioli-Moraes, M.C., 2020. Inefficient weapon—the role of plant secondary metabolites in cotton defence against the boll weevil. Planta 252, 94. https://doi.org/10.1007/s00425-020-03497-w.

McCall, P.J., Turlings, T.C.J., Lewis, W.J., Tumlinson, J.H., 1993. Role of plant volatiles in host location by the specialist parasitoid Microplitis croceipes cresson (Braconidae: Hymenoptera). J. Insect Behav. 6, 625–639. https://doi.org/10.1007/BF01048128.

Minyard, J.P., Tumlinson, J.H., Hedin, P.A., Thompson, A.C., 1965. Isolation and identification, constituents of cotton bud terpene hydrocarbons. J. Agric. Food Chem. 13, 599–602.

O'Callaghan, S., De Souza, D.P., Isaac, A., Wang, Q., Hodkinson, L., Olshansky, M., Erwin, T., Appelbe, B., Tull, D.L., Roessner, U., Bacic, A., McConville, M.J., Likić, V. A., 2012. PyMS: a Python toolkit for processing of gas chromatography-mass spectrometry (GC-MS) data. Application and comparative study of selected tools. BMC Bioinf. 13, 115. https://doi.org/10.1186/1471-2105-13-115.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2020. Vegan: Community Ecology Package. R Package Version 2, pp. 5–7.

Opitz, S., Kunert, G., Gershenzon, J., 2008. Increased terpenoid accumulation in cotton (*Gossypium hirsutum*) foliage is a general wound response. J. Chem. Ecol. 34, 508–522.

O'Reilly-Wapstra, J.M., Freeman, J.S., Davies, N.W., Vaillancourt, R.E., Fitzgerald, H., Potts, B.M., 2011. Quantitative trait loci for foliar terpenes in a global eucalypt species. Tree Genet. Genomes 7, 485–498. https://doi.org/10.1007/s11295-010-0350-6.

Paré, P.W., Tumlinson, J.H., 1997. De novo biosynthesis of volatiles induced by insect herbivory in cotton plants. Plant Physiol. 114, 1161–1167.

Pichersky, E., Raguso, R.A., 2018. Why do plants produce so many terpenoid compounds? New Phytol. 220, 692–702. https://doi.org/10.1111/nph.14178.

R Core Team, 2013. n.d. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. URL. http://www.R-project.org/.

Renou, A., Téréta, I., Togola, M., 2011. Manual topping decreases bollworm infestations in cotton cultivation in Mali. Crop Protect. 30, 1370–1375. https://doi.org/10.1016/j.cropro.2011.05.020.

Reyes-Hernández, M., Angulo-Pérez, D., Quijano-Medina, T., Moreira, X., Parra-Tabla, V., Vásquez-Bolaños, M., Abdala-Roberts, L., 2022. An experimental test of ant effects on herbivory and pathogen infection on wild cotton (*Gossypium hirsutum* L.). Arthropod-Plant Interact. 16, 77–86. https://doi.org/10.1007/s11829-021-09876-8.

Röse, U.S., Tumlinson, J.H., 2005. Systemic induction of volatile release in cotton: how specific is the signal to herbivory? Planta 222, 327–335.

Rosenkranz, M., Chen, Y., Zhu, P., Vlot, A.C., 2021. Volatile terpenes – mediators of plant-to-plant communication. Plant J. 108, 617–631. https://doi.org/10.1111/tpj.15453.

Senft, M., Clancy, M.V., Weisser, W.W., Schnitzler, J.-P., Zytynska, S.E., 2019. Additive effects of plant chemotype, mutualistic ants and predators on aphid performance and survival. Funct. Ecol. 33, 139–151. https://doi.org/10.1111/1365-2435.13227.

Suzuki, R., Shimodaira, H., 2006. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics 22, 1540–1542. https://doi.org/10.1093/bioinformatics/btl117.

Thompson, J., Charpentier, A., Bouguet, G., Charmasson, F., Roset, S., Buatois, B., Vernet, P., Gouyon, P., 2013. Evolution of a genetic polymorphism with climate change in a Mediterranean landscape. Proc. Natl. Acad. Sci. 110, 2893–2897. https://doi.org/10.1073/pnas.1215833110.

Trindade, H., Pedro, L.G., Figueiredo, A.C., Barroso, J.G., 2018. Chemotypes and terpene synthase genes in Thymus genus: state of the art. Ind. Crop. Prod. 124, 530–547. https://doi.org/10.1016/j.indcrop.2018.08.021.

Turlings, T.C., Loughrin, J.H., McCall, P.J., Röse, U.S., Lewis, W.J., Tumlinson, J.H., 1995. How caterpillar-damaged plants protect themselves by attracting parasitic wasps. Proc. Natl. Acad. Sci. 92, 4169–4174. https://doi.org/10.1073/pnas.92.10.4169.

Velázquez-López, R., Wegier, A., Alavez, V., Pérez-López, J., Vázquez-Barrios, V., Arroyo-Lambaer, D., Ponce-Mendoza, A., Kunin, W.E., 2018. The mating system of the wild-to-domesticated complex of *Gossypium hirsutum* L. Is mixed. Front. Plant Sci. 9.

Wegier, A., Pińeyro- Nelson, A., Alarón, J., Gálvez-Mariscal, A., Álvarez-Buylla, E.R., Pińero, D., 2011. Recent long-distance transgene flow into wild populations conforms to historical patterns of gene flow in cotton (*Gossypium hirsutum*) at its centre of origin. Mol. Ecol. 20, 4182–4194. https://doi.org/10.1111/j.1365-294X.2011.05258.x.

Wilke, C.O., 2021. Ggridges: Ridgeline Plots in "Ggplot2". R package version 0.5.3.

Yang, C.-Q., Wu, X.-M., Ruan, J.-X., Hu, W.-L., Mao, Y.-B., Chen, X.-Y., Wang, L.-J., 2013. Isolation and characterization of terpene synthases in cotton (*Gossypium hirsutum*). Phytochemistry 96, 46–56. https://doi.org/10.1016/j.phytochem.2013.09.009.

Yuan, D., Grover, C.E., Hu, G., Pan, M., Miller, E.R., Conover, J.L., Hunt, S.P., Udall, J.A., Wendel, J.F., 2021. Parallel and intertwining threads of domestication in allopolyploid cotton. Adv. Sci. 8, 2003634 https://doi.org/10.1002/advs.202003634.

Zhang, C.-P., Zhang, J.-L., Sun, Z.-R., Liu, X.-Y., Shu, L.-Z., Wu, H., Song, Y., He, D.-H., 2022. Genome-wide identification and characterization of terpene synthase genes in *Gossypium hirsutum*. Gene 828, 146462. https://doi.org/10.1016/j.gene.2022.146462.

Zhou, F., Pichersky, E., 2020. More is better: the diversity of terpene metabolism in plants. Phys. Met. 55, 1–10. https://doi.org/10.1016/j.pbi.2020.01.005.