

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Genetic mapping and characterization of the globe artichoke (+)-germacrene A synthase gene, encoding the first dedicated enzyme for biosynthesis of the bitter sesquiterpene lactone cynaropicrin.

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/107602> since 2016-07-14T14:37:10Z

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)



UNIVERSITÀ DEGLI STUDI DI TORINO

This Accepted Author Manuscript (AAM) is copyrighted and published by Elsevier. It is posted here by agreement between Elsevier and the University of Turin. Changes resulting from the publishing process - such as editing, corrections, structural formatting, and other quality control mechanisms - may not be reflected in this version of the text. The definitive version of the text was subsequently published in *Plant Science*, Volume 190, 2012, DOI [10.1016/j.plantsci.2012.03.006](https://doi.org/10.1016/j.plantsci.2012.03.006).

You may download, copy and otherwise use the AAM for non-commercial purposes provided that your license is limited by the following restrictions:

- (1) You may use this AAM for non-commercial purposes only under the terms of the CC-BY-NC-ND license.
- (2) The integrity of the work and identification of the author, copyright owner, and publisher must be preserved in any copy.
- (3) You must attribute this AAM in the following format: Creative Commons BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>), [[+10.1016/j.plantsci.2012.03.006](https://doi.org/10.1016/j.plantsci.2012.03.006)]

Genetic mapping and expression of a globe artichoke (+)-germacrene A synthase, an enzyme involved in sesquiterpene lactone biosynthesis

Barbara Menin¹, Cinzia Comino¹, Ezio Portis¹, Andrea Moglia¹, Katarina Cankar^{2,3}, Harro J Bouwmeester³, Sergio Lanteri¹, Jules Beekwilder^{2*}

¹DIVAPRA, Plant Genetics and Breeding, University of Torino, via L. da Vinci 44, 10095 Grugliasco (TO) Italy

² Plant Research International, P.O. Box 16, 6700 AA Wageningen, The Netherlands

³ Laboratory of Plant Physiology, Wageningen University and Research Centre, 6708PB Wageningen, The Netherlands

* **Corresponding author:** jules.beekwilder@wur.nl

Telephone number: +31 317 480979

Fax number: +31 317 418094

Abstract

Globe artichoke (*Cynara cardunculus* var. *scolymus* L., Asteraceae) is a perennial crop traditionally consumed as a vegetable in the Mediterranean countries and rich in nutraceutically and pharmaceutically active compounds, including phenolic and terpenoid compounds. Its bitter taste is caused by its high content of sesquiterpene lactones (STLs), such as cynaropicrin. The biosynthetic pathway responsible for STL biosynthesis in globe artichoke is unknown, but likely proceeds through germacrene A, as has been shown for other Asteraceae species. Here, we investigated the accumulation of cynaropicrin in different tissues of globe artichoke, and compared it to accumulation of phenolic compounds. Cynaropicrin concentration was highest in old leaves. A putative germacrene A synthase (*GAS*) gene was identified in a set of ~19,000 globe artichoke unigenes. When heterologously expressed in *E. coli*, the putative globe artichoke *GAS* converted farnesyl diphosphate (FPP) into (+)-germacrene A. Among various tissues assayed, the level of globe artichoke *GAS* expression was highest in mature (six week old) leaves. A sequence polymorphism within a mapping population parent allowed the corresponding *GAS* gene to be positioned on a genetic map. This study reports the isolation, expression and mapping of a key gene involved in STL biosynthesis in *C. cardunculus*. This is a good basis for further investigation of this pathway.

Keywords: *Cynara cardunculus*, (+)-germacrene A, terpene synthase, cynaropicrin, sesquiterpene lactones

Introduction

Cynara cardunculus is a diploid ($2n=2x=34$) cross pollinated plant species belonging to the Asteraceae family. This species includes globe artichoke (var. *scolymus*), cultivated cardoon (var. *altilis*) and their progenitor wild cardoon [var. *sylvestris* (Lamk) Fiori] [1]. Globe artichoke is cultivated in the Mediterranean Basin and in various parts of the Americas and Asia. Its immature inflorescences (“heads”) are valued for consumption and is marketed fresh, canned or frozen. The phenolic compounds of globe artichoke, which have several medicinal properties, have been subject of study in the past years. In particular studies have been devoted to the biosynthesis of caffeoylquinic acids, their regulation under UV stress and to the identification of several genes from their biosynthetic pathway [2-6].

The characteristic bitterness of globe artichoke is mainly due to the presence of sesquiterpene lactones (STLs), of which the two major representatives are cynaropicrin and, at lower concentration, grosheimin and its derivatives [7-9]. Cynaropicrin is known as the major bitter principle in globe artichoke [7, 9], and, like many sesquiterpenes lactones, has various medicinal activities [10-14], among which cytotoxicity against several types of cancer cells [15]. Cynaropicrin belongs to the guaianolide STL subclass. Like other STLs, such as eudesmanolides and germacranolides, these compounds likely derive from the common precursor costunolide [16, 17]. Costunolide is known to be the product of germacrene A, by the action of two cytochrome P450 enzymes [17-19] (Fig. 1). Therefore, germacrene A biosynthesis is assumed to be the first committed step in the biosynthesis of STLs [20].

Germacrene A is formed by cyclization of farnesyl diphosphate (FPP). The cyclization reaction is catalysed by the terpene synthase (+)-germacrene A synthase (GAS) (Fig. 1). Germacrene A synthases from a number of Asteraceae species have been described, including chicory [20], lettuce [21], goldenrod [22], *Ixeris dentata* [23], *Artemisia annua* [24] and sunflower [25].

Linkage maps, created for genes in biosynthetic pathways in several species, can be used to locate known genes of a pathway within a specific genomic region [26, 27]. Such maps can be very instrumental in the breeding of new varieties with tailored content of bio-active compounds. Genetic mapping of germacrene A synthases on the chicory genome has recently been reported [28]. In globe artichoke, several genes involved in the phenylpropanoid pathway have been already located on a genetic *C. cardunculus* map, based on a two-way pseudo test-cross strategy in an F₁ population obtained by crossing two contrasting genotypes: globe artichoke ‘Romanesco C3’ and cultivated cardoon ‘Altilis 41’ [6, 29, 30].

Here we report the isolation of a germacrene A synthase gene from globe artichoke, which is

involved in the initial step leading to cynaropicrin biosynthesis. Gene information is used to define its genomic location, and to profile its pattern of expression in globe artichoke tissues, including *in vitro* calli. Gene expression is compared to cynaropicrin content of the tissues.

Materials and Methods

Plant material, RNA extraction and cDNA synthesis

Seeds of the globe artichoke variety ‘Concerto’ (obtained from Nunhems Seeds, The Netherlands) were germinated for two weeks between two layers of wetted filter paper, and transplanted into pots in a glasshouse held at 22–24°C for ten weeks. The same plants were then transplanted to the field up to the production of commercial heads (capitula), about 35 weeks after seed germination. Mature (six week old) leaves were collected from plants in greenhouse and stored in liquid nitrogen. Old leaves (twenty weeks old) as well as capitula outer bracts and receptacles were also harvested from the same plants grown in field. Calli were cultured from leaf explants of virus-free globe artichoke plants (variety ‘Romanesco C3’), as reported by Menin *et al.* [31].

Total RNA was extracted from 100 mg fresh tissue using the TRIzol reagent (Invitrogen, USA), following the manufacturer's instructions; its concentration determined by spectrophotometry, and its integrity checked by electrophoresis in 1% (w/v) formaldehyde-agarose gels [32]. The RNA was purified by adding one volume of 7.5 M LiCl to five volumes of crude sample, holding at -20°C for ~1h, centrifuging (14,000 g, 20min) and re-suspending the pellet in 10 µl DEPC H₂O. Reverse transcription was performed using the High Capacity RNA-to-cDNA Master Mix (Applied Biosystems, USA), following the manufacturer's instructions.

Identification and phylogeny of full-length sesquiterpene synthase cDNA

Two chicory GAS sequences (CiGAS_short and CiGAS_long,[20]) were used as BLAST search terms to query the globe artichoke unigene database containing ~19000 sequences [33]. Two ESTs, CL5386 and CCPW7274, showed a high similarity level with the beginning and the end of the queries, respectively. Primers, Probe_GAS_For and Probe_GAS_Rev (Table 1), were designed on CL5386 and CCPW7274 sequence respectively to investigate the full length. The two extended ESTs matched together originating one full-length cDNA sequence. The GAS open reading frame (ORF) was amplified from a cDNA template using the primer pair Expr_GAS_For and Expr_GAS_Rev (Table 1), which were designed to anneal

to each end of the transcript, while building in a *Bam*HI restriction site at the 5' end and a *Hind*III one at the 3' end. PCRs was performed using Phusion® Hot Start High-Fidelity DNA Polymerase (Finnzymes, Finland), following the manufacturer's instructions. Multiple sequence alignments were performed using ClustalW2 (www.ebi.ac.uk/Tools/clustalw2/index.html), applying standard parameters. The putative globe artichoke GAS (CcGAS) sequence was then aligned with a set of known GAS sequences [34], and phylogenetic analysis was conducted using MEGA v3.0 software [35]. The neighbour-joining method was used to construct a guided tree, supported by bootstrapping based on 10,000 replicates.

Heterologous expression of CcGAS in E. coli

The CcGAS ORF was restricted with *Bam*HI and *Hind*III, and ligated into the multiple cloning site of *Bam*HI and *Hind*III digested pACYCDuet™-1 vector (Novagen, Germany). The resulting recombinant pACYCDuet-GAS plasmid was transformed into competent *E. coli* strain BL21 A1 cells (Invitrogen, USA), previously made competent by treatment with 100 mM CaCl₂. A 500 µL aliquot of an over-night culture (LB medium containing 50 µg/mL chloramphenicol) of the transgenic *E. coli* was transferred into 50 mL 2xYT medium containing 50 µg/mL chloramphenicol and incubated at 37°C at 250 rpm until the optical density (A₆₀₀) had reached 0.6. Subsequently, gene expression was induced by the addition of 0.02% (w/v) arabinose, and the cultures were incubated overnight at 18°C and 250 rpm. A 25 mL sample of this culture was centrifuged (10 min, 10,000g), and the pellet re-suspended in 1 mL 50 mM Tris-HCl pH 8, 150 mM NaCl, 1.4 mM 2-mercaptoethanol, 10% (v/v) glycerol at 4°C. The cells were disrupted by sonication on ice (five bursts for 10 s, each separated by a 10 s pause), using an MSE Soniprep 150 device set to an amplitude of 14 µm. Particulate matter was removed by centrifugation (10 min, 13,000 g, 4°C), and the supernatant passed through a Qiagen Ni-NTA spin column, and eluted twice with 300 µL of Elution Buffer (50 mM NaH₂PO₄, 300 mM NaCl, 500 mM imidazole, pH 8.0), following the manufacturer's instructions (Qiagen) to purify the protein. A 200 µL volume of eluate was placed in a Slide-A-Lyzer Mini Dialysis Unit (10,000 MWCO; Pierce), and dialysed for 3 h against 1 L 50 mM Tris-HCl pH 8, 10% (v/v) glycerol, 1.4 mM 2-mercaptoethanol at 4°C. After dialysis, the eluate was immediately used for assaying enzyme activity.

Enzyme assay and GC-MS analysis

For the enzyme assay, 800 µL MOPSO buffer [15 mM 3-(N-morpholino)-2-hydroxypropane

sulphonic acid, pH 7, 1 mM MgCl₂, 0.1% (v/v) Tween 20, 1 mM ascorbic acid, 1 mM dithiothreitol, 12.5% (v/v) glycerol] was mixed with 100 µL of the purified enzyme solution and 5 µL 10mM farnesyl diphosphate (FPP) (Sigma-Aldrich, USA). This mixture was overlaid with 1 mL pentane, and incubated at 30°C with mild agitation for 2 h, after which the pentane layer was collected. The remaining aqueous phase was re-extracted with 1 mL ethylacetate. The ethylacetate extract and the pentane phase were combined, centrifuged at 4,000 g, dried over a sodium sulphate column and analyzed by GC-MS. The GC-MS analysis was performed using an Agilent Technologies system, comprising a 7980A GC system, a 597C inert MSD detector (70eV) and a Phenomenex Zebron ZB-5ms column of 30m length x 0.25 mm internal diameter and 0.25 µm stationary phase, with a Guardian pre-column (5 m). The sample (1 µL) was injected into the device chamber at 250°C; the injection was splitless and the ZB5 column was maintained at 45°C for 1 min, after which the temperature was increased at a rate of 10°C per min until it reached 300°C. Compounds were identified by their retention index and by their mass spectrum in comparison to a (+) germacrene A standard.

Expression analysis in globe artichoke tissues

The tissues chosen for the expression analysis were: mature (6 weeks old) and old (20 weeks old) leaves, outer bracts, receptacle and callus, collected as previously described. Primer 3 software (frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) was used to design primers (RT_GAS_For and RT_GAS_Rev, Table 1) for real time quantitative PCR (RT-qPCR) based on the GAS sequence. Actin was chosen as the reference gene, assayed with primers RT_ACT_For / RT_ACT_Rev (Table 1), designed from a partial globe artichoke actin sequence (GenBank accession AM744951). The first cDNA strand was synthesized from 1 µg of total RNA using the iScript cDNA Synthesis kit (Bio-Rad Laboratories, USA), following the manufacturer's instructions. Each tissue was sampled twice as biological replicates. The cDNA was diluted to obtain a threshold cycle (CT) value between 25 and 35. A 20 µL RT-qPCR was performed in triplicate for each sample; these reactions contained 1x iQ Supermix, 1x SYBR-Green I (iQTM, SYBR® GreenSupermix), 10 µM RT primers (Table 1) and 3 µL cDNA. PCR reactions were carried out in an iCycler Real-time PCR Detection System (Bio-Rad Laboratories, USA). The PCR conditions comprised an initial incubation of 95°C/5 min, followed by 35 cycles of 95°C/15 s and 60°C/60 s. In all experiments, appropriate negative controls containing no template were subjected to the same procedure. Melting curve analysis was performed at the end of the amplification process. Standard curves were analyzed using

iCycler iQ software. The slopes of the standard curves of the amplification system revealed a correlation coefficient >0.99 and an efficiency >0.90 (data not shown). The comparative threshold cycle method was applied, in which $\Delta\Delta Ct$ is calculated as $\Delta Ct_I - \Delta Ct_M$, where ΔCt_I is the Ct value for the target normalized against the reference gene and ΔCt_M is the Ct value for the calibrator, also normalized against the reference gene.

Biochemical analyses

Globe artichoke tissues were freeze dried before analysis. 10 mg of the same dry tissues used for RT-qPCR was ground to a fine powder, and extracted with 1 ml of 75% methanol with 0.1% formic acid by sonication for 20 min. Then the extracts were centrifuged for 5 min at 13,000 rpm and filtered through 0.2 μ m inorganic membrane filters (RC4, Sartorius, Germany). For identification and quantification of products, liquid chromatography, coupled to quadrupole time-of-flight mass spectrometry (LC-QTOF-MS) was performed using a Waters Alliance 2795 HPLC connected to a Waters 2996 PDA detector and subsequently a QTOF Ultima V4.00.00 mass spectrometer (Waters, MS technologies, UK) operating in positive ionization mode. The column used was an analytical column (Luna 3 μ C18/2 100A; 2.0 \times 150 mm; Phenomenex, USA) attached to a C18 pre-column (2.0 \times 4 mm; AJO-4286; Phenomenex, USA). Degassed eluent A, ultra pure water:formic acid (1000:1,v/v), and eluent B, acetonitril:formic acid (1000:1,v/v) were used at 0.19 ml/min. The gradient started at 5% B and increased linearly to 75% B in 45 min, after which the column was washed and equilibrated for 15 min before the next injection. The injection volume was 5 μ l.

Sequence polymorphism detection and linkage analysis

Sequence variations in the GAS gene were sought by comparing the allelic form present in the globe artichoke genotype 'Romanesco C3' with the cultivated cardoon 'Altilis 41', the parental genotypes of the F₁ progeny used for map development [29]. Parental genomic DNAs were amplified with primers SNP_GAS_For and SNP_GAS_Rev (Table 1), and amplicons were directly sequenced to facilitate SNP mining.

Subsequent genotyping was achieved using an allele specific primer (Inn_GAS_Rev, Table 1) in combination with two common primers (Out_GAS_For and Out_GAS_Rev), designed using the package available at cedar.genetics.soton.ac.uk/public_html/primer1.html. The primers were applied to a population of 94 F₁ progeny of the cross 'Romanesco C3' x 'Altilis 41', and the resulting PCR products were separated by 2% agarose gel electrophoresis. These

genotypic data were then incorporated into a pre-existing data set already used to generate a globe artichoke genetic map [29, 30]. Goodness-of-fit between observed and expected segregation ratios was tested by χ^2 , and only markers fitting or deviating only marginally from expectation ($\chi^2_{\alpha=0.1} < \chi^2 \leq \chi^2_{\alpha=0.01}$) were included for mapping. Linkage groups were established by JoinMap v4.0 [36] software, on the basis of a LOD threshold of 6.0, using the following parameter set: Rec=0.40, LOD=1.0, Jump=5. Map distances were converted to centiMorgans (cM) using the Kosambi mapping function [37]. Furthermore, a combined data set based exclusively on microsatellite and SNP markers was created to construct a consensus linkage map as reported by Portis *et al.* [29].

Results

Identification and phylogeny of full-length sesquiterpene synthase cDNA

To identify a putative germacrene A synthase, a collection of 19,000 unigenes from globe artichoke was interrogated. Two members of the globe artichoke unigene set (CL5386 and CCPW7274) showed a high level of sequence similarity (84% and 86% identity respectively) with chicory *GAS* genes, but did not overlap. PCR carried out using primers designed on these two unigenes (Probe_GAS_For and Probe_GAS_Rev, respectively Table 1) allowed the amplification of a single fragment, demonstrating that both unigenes were part of the same sequence. The full-length open reading frame was amplified from globe artichoke and the resulting 1791bp ORF (GenBank accession JN383985) was translated *in silico* into a 560 residue protein of predicted molecular weight 64.5KDa. A phylogenetic tree of germacrene A synthases indicated that globe artichoke germacrene A synthase maps in a branch of synthases containing also the short germacrene A synthase from chicory, but being clearly distinct from other members (Fig. 2).

Heterologous expression of CcGAS in E. coli

To assess the activity of globe artichoke GAS, the full-length cDNA was cloned in expression vector pACYCDUET-1, and expressed in *E. coli* BL21A1. The recombinant CcGAS protein was purified by his-tag affinity chromatography, and its activity with FPP was assayed by *in vitro* enzyme assays. Formation of a new product was detected in GC MS analysis, at a retention time of 12.50. This product was not detected when the reaction was performed with the purified extract of the negative control (*E. coli* transformed with empty vector) (Fig. 3 A). The retention time and mass spectrum of this product were identical to that of β -elemene

which was also formed when a germacrene A standard was injected (not shown). The novel product was thus identified as β -elemene (Fig. 3 B) that is known to be the cope-rearrangement product of germacrene A [17].

Gene expression of CcGAS in globe artichoke tissues

Different leaf tissues, floral tissues and callus were tested for expression of the *GAS* gene using RT-qPCR. The expression level detected was highest in mature (six week old) leaves, where the level was 640 fold higher than in old (twenty week old) leaves. The gene was also expressed in callus and the receptacle, at a level, respectively, about 4- and 9-fold higher than in old leaves (Fig. 4). No expression was detected in the bracts.

Detection of cynaropicrin and phenolic compounds in globe artichoke tissues

The sesquiterpene lactone cynaropicrin can be detected in globe artichoke extracts by liquid chromatography coupled to mass spectrometry (LC-MS) techniques [8]. LC-MS in combination with photodiode-array absorption measurements was deployed to detect cynaropicrin in the same set of tissues tested for *GAS* expression analysis. In this experiments, we also detected a number of phenolic compounds, such as dicaffeoylquinic acid, chlorogenic acid, luteolin-7-glucoside and luteolin-7-rutinoside, which are known to occur in globe artichoke. Table 2 shows the chromatographic and MS properties of these compounds. Cynaropicrin was detected in positive mode LC-ESI-MS as a compound eluting at 21.9 minutes (Fig. 5), showing characteristic ions at m/z 347 [$M+H$], m/z 245 [$M_{\text{fragment}}+H$], m/z 385 [$M+K$] and m/z 369 [$M+Na$] [8].

Accumulation of secondary metabolites may occur in specific plant tissues, depending on the class of compounds. In Table 3, it is shown that cynaropicrin can be detected in both mature (six week old) and old (twenty week old) leaves, and in receptacle and callus, but could not be detected in outer bract tissue. The highest content was found in the old and mature leaf tissues. Phenolic compounds show different accumulation patterns: dicaffeoylquinic acid and chlorogenic acid are highest in receptacle and bract tissues, while luteolin glucoside and in particular luteolin rutinoside are highest in old leaf, but much less present in mature leaf. Thus, cynaropicrin shows an accumulation pattern that is clearly distinct from phenolic compounds.

SNP detection and linkage analysis

To assess genetic diversity in *CcGAS* genes, cDNAs were amplified from globe artichoke

'Romanesco C3' and cultivated cardoon 'Altilis 41' genotypes. Analysis of the *GAS* nucleotide sequences revealed one single-nucleotide polymorphism (SNP) at position 603. This polymorphism was used to design a PCR amplification marker assay that showed that 'Altilis 41' is heterozygous for this SNP. The polymorphism segregated as a 1:1 ratio ($\chi^2=0.89$, $p>0.1$) in the mapping population, allowing the sequence to be mapped to linkage group 'Alt_5' in the cultivated cardoon map, between the AFLP loci e33/t89-510 and e39/m50-410 (Fig. 6). In all, 12 markers (four microsatellites, seven AFLPs and the present SNP locus) have been assigned to date on this linkage group, giving a genetic length of 55.1cM and a mean inter-marker distance of 4.82cM, with no pair of loci separated by >7cM. The presence of co-dominant loci shared between the cultivated cardoon and globe artichoke maps allowed the construction of the consensus SSR-based linkage group (LG) 10, with the *GAS* locus lying between the microsatellite loci aCELMS-14 and rCyEM-214 (Fig. 6).

Discussion

The STLs are associated with potent bio-activity, and are characteristic constituents of many Asteraceae species; it has been postulated that their *in vivo* role has been important for the evolutionary success of this large plant family [16, 25]. The compounds probably serve as a deterrent against herbivory, and may be responsible for the relative unpalatability (and occasional toxicity) of these species for grazing animals.

Much of the research focus applied to the STLs in globe artichoke has been directed to its medicinal value and its content in leaves [10, 14, 38]. We analyzed the content of phenolic compounds and cynaropicrin in several globe artichoke (*Cynara cardunculus* var. *scolymus* L.) tissues. Since cynaropicrin is the most abundant sesquiterpene lactone (STL) and the major bitter principle of this plant species [7, 9], we focused our attention on the STL biosynthetic pathway that is almost unknown in this species. Our biochemical analyses on globe artichoke tissues highlighted that cynaropicrin was mainly detected in mature and old leaf tissues (Table 3). In parts used for human consumption, such as the receptacle and the outer bracts, the content was much lower (receptacle) or undetectable (outer bracts) (Table 3). These results parallel to a large extent those reported by Schneider & Thiele [7]. In contrast to their observations, where cynaropicrin in leaves was observed to be reduced to about half during 20 weeks, we did not observe a decrease in cynaropicrin content during this period, but rather a slight increase (Table 3). Possibly this could result from different growth conditions as well as genotypes analysed.

The specific absence of cynaropicrin in floral parts could be the result of domestication of the globe artichoke, since the bitterness of cynaropicrin could be considered as a negative organoleptic aspect. Flowers from other *Asteraceae*, such as sunflower and *Artemisia annua*, are known to be very rich in STLs. In these species, specialized organs, such as glandular trichomes, are formed on floral parts in which STLs are accumulated [25]. In chicory and lettuce, the STLs are known to accumulate in laticifers throughout the plant [39]. Cynaroideae such as globe artichoke do not have laticifers or glandular trichomes, but secretory canals are generally present in Cardueae roots [40], that sometimes continue into the stem. In *Cynara cardunculus* the secretory canals have been found in the stems [41]; it remains to be established if STLs accumulate in globe artichoke secretory ducts.

In this paper we isolated and characterized the cDNA for sesquiterpene synthase from globe artichoke (CcGAS). *In vitro* enzyme assays showed that CcGAS protein is indeed active as a germacrene A synthase, thus suggesting its involvement in STL biosynthesis [20]. This role is confirmed by the expression analysis, which showed a correlation between the CcGAS expression level and the cynaropicrin content. The expression of the active enzyme was found to be absent in tissues where cynaropicrin was not detected, which further supports a role for this enzyme in cynaropicrin biosynthesis. Furthermore strongly reduced expression of CcGAS in old leaves, where high levels of cynaropicrin were found, could be explained by accumulation of the compound in mature leaves, which is apparently stored and not subject to further metabolism or degradation. Notably, the STL was clearly detectable in *in vitro* grown callus tissue, and also expression of GAS in *in vitro* grown calli could be observed. Callus is not a very relevant tissue for commercial extracts or consumption, but may be useful as a model to study the response of gene expression and/or enzyme activities to external stimuli.

Several GAS isoforms have been identified in various *Asteraceae* species: for example, there are two in both chicory (CiGAS_short and CiGAS_long, [20]) and lettuce (LTC1 and LTC2, [21]), and three in the glandular trichome of sunflower (HaGAS1, HaGAS2 and HaGAS3, [25]). Thus, it is possible that also globe artichoke includes more than one form of this enzyme. However, so far we have been unable to identify other homologues among the available globe artichoke EST dataset.

Despite the ecological importance and potential medicinal use of the STLs, the biosynthetic pathway in plants has not yet been elucidated to completion. However, the genes for the first dedicated steps of the biosynthetic pathway – in particular GAS, germacrene A oxidase (GAO) and costunolide synthase (COS) - have been isolated and functionally characterized in a range of *Asteraceae* species [34, 42].

The genetic placement of gene-derived markers can make a direct contribution to our understanding of the inheritance of plant traits such as the content of STLs. Moreover it forms an essential link between molecular plant knowledge and breeding of new varieties of globe artichoke. The CcGAS locus was mapped in the cultivated cardoon genome, as the polymorphism was in a heterozygous state in the cultivated cardoon ‘Altilis 41’ parent. This mapping both improved the quality of the ‘Alt_5’ linkage group map co-dominant GAS_snp603 marker and filled a gap in the consensus linkage group 10 one. Such mapping would allow looking for metabolic gene clusters. Moreover, populations with variation in STL content can be characterized by molecular tools, which would expand the breeding potential for taste and health aspects of globe artichoke.

Our future research activities will be focused on the analysis of the *in vivo* role of CcGAS, using a combination of forward genetic approaches, on the identification of other candidate genes involved in STL synthesis, with particular attention to the downstream steps from GAS, such as germacrene A oxidase and costunolide synthase. The genetic placement of characterized gene-derived markers, such as the one described here, may contribute in identifying QTLs involved in sesquiterpene lactones accumulation and highlight if the mapped GAS gene underlines a major QTL involved in cynaropicrin biosynthesis. An interesting example of this kind of approach has been recently provided in *Artemisia annua*, where QTLs that contribute to sesquiterpene-lactone content have been mapped in candidate genes from the STL biosynthetic pathway [43].

References

- [1] S. Lanteri, E. Portis, Globe artichoke and Cardoon. In *Vegetables I: Asteraceae, Brassicaceae, Chenopodiaceae, and Cucurbitaceae*. Edited by Prohens J and Nuez F Springer (2008) 49-74.
- [2] C. Comino, S. Lanteri, E. Portis, A. Acquadro, A. Romani, A. Hehn, R. Larbat, F. Bourgaud, Isolation and functional characterization of a cDNA coding a hydroxycinnamoyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus* L, *BMC Plant Biol.* 7 (2007) 14.
- [3] A. Moglia, S. Lanteri, C. Comino, A. Acquadro, R. de Vos, J. Beekwilder, Stress-induced biosynthesis of dicaffeoylquinic acids in globe artichoke, *J. Agric. Food Chem.* 56 (2008) 8641-8649.
- [4] C. Comino, A. Hehn, A. Moglia, B. Menin, F. Bourgaud, S. Lanteri, E. Portis, The isolation and mapping of a novel hydroxycinnamoyltransferase in the globe artichoke chlorogenic acid pathway, *BMC Plant Biol.* 9 (2009) 30.
- [5] A. Moglia, C. Comino, E. Portis, A. Acquadro, R. De Vos, J. Beekwilder, S. Lanteri, Isolation and mapping of a C3'H gene (CYP98A49) from globe artichoke, and its expression upon UV-C stress, *Plant Cell Rep.* 28 (2009) 963-974.
- [6] B. Menin, C. Comino, A. Moglia, Y. Dolzhenko, E. Portis, S. Lanteri, Identification and mapping of genes related to caffeoylquinic acid synthesis in *Cynara cardunculus* L., *Plant Science* 179 (2010) 338-347.
- [7] G. Schneider, K. Thiele, Die Verteilung des Bitter-stoffes Cynaropicrin in der Artischocke, *Planta Medica* 26 (1974) 174-183.
- [8] J. Fritsche, C. Beindorff, M. Dachtler, H. Zhang, J. Lammers, Isolation, characterization and determination of minor artichoke (*Cynara scolymus* L.) leaf extract compounds, *European Food Research and Technology* 215 (2002) 149-157.
- [9] G. Cravotto, G. Nano, A. Binello, P. Spagliardi, G. Seu, Chemical and biological modification of cynaropicrin and grosheimin: a structure-bitterness relationship study, *Journal of the Science of Food and Agriculture* 85 (2005) 1757-1764.
- [10] H. Shimoda, K. Ninomiya, N. Nishida, T. Yoshino, T. Morikawa, H. Matsuda, M. Yoshikawa, Anti-hyperlipidemic sesquiterpenes and new sesquiterpene glycosides from the leaves of artichoke (*Cynara scolymus* L.): structure requirement and mode of action, *Bioorganic & Medicinal Chemistry Letters* 13 (2003) 223-228.
- [11] J. Cho, A. Kim, J. Jung, T. Chun, M. Rhee, E. Yoo, Cytotoxic and pro-apoptotic activities of cynaropicrin, a sesquiterpene lactone, on the viability of leukocyte cancer cell lines, *European Journal of Pharmacology* 492 (2004) 85-94.
- [12] E. Schinor, M. Salvador, I. Ito, S. de Albuquerque, D. Dias, Trypanocidal and antimicrobial activities of *Moquinia kingii*, *Phytomedicine* 11 (2004) 224-229.
- [13] F. Emendörfer, F. Emendörfer, F. Bellato, V. Noldin, V. Cechinel-Filho, R. Yunes, F. Delle Monache, A. Cardozo, Antispasmodic activity of fractions and cynaropocrin from *Cynara scolymus* on guinea-pig ileum, *Biological & Pharmaceutical Bulletin* 28 (2005) 902-904.
- [14] K. Ishida, R. Kojima, M. Tsuboi, Y. Tsuda, M. Ito, Effects of artichoke leaf extract on acute gastric mucosal injury in rats, *Biological & Pharmaceutical Bulletin* 33 (2010) 223-229.
- [15] K. Yasukawa, H. Matsubara, Y. Sano, Inhibitory effect of the flowers of artichoke (*Cynara cardunculus*) on TPA-induced inflammation and tumor promotion in two-stage carcinogenesis in mouse skin, *J. Nat. Med.* 64 (2010) 388-391.
- [16] F.C. Seaman, Sesquiterpene lactones as taxonomic characteristics in the Asteraceae, *The Botanical Review* 48 (1982) 123-595.
- [17] J. De Kraker, M. Franssen, M. Joerink, A. de Groot, H. Bouwmeester, Biosynthesis of

- costunolide, dihydrocostunolide, and leucodin. Demonstration of cytochrome P450-catalyzed formation of the lactone ring present in sesquiterpene lactones of chicory, *Plant Physiology* 129 (2002) 257-268.
- [18] N. Ikezawa, J. Gopfert, D. Nguyen, S. Kim, P. O'Maille, O. Spring, D. Ro, Lettuce costunolide synthase (CYP71BL2) and its homolog (CYP71BL1) from sunflower catalyze distinct regio- and stereoselective hydroxylations in sesquiterpene lactone metabolism, *Journal of Biological Chemistry* 286 (2011) 21601-21611.
- [19] Q. Liu, M. Majdi, K. Cankar, M. Goedbloed, T. Charnikhova, F.W.A. Verstappen, R.C.H. de Vos, J. Beekwilder, S. van der Krol, H. Bouwmeester, Reconstitution of the costunolide biosynthetic pathway in yeast and *Nicotiana benthamiana*, *PLoS One* 6 (2011) e23255.
- [20] H. Bouwmeester, J. Kodde, F. Verstappen, I. Altug, J. de Kraker, T. Wallaart, Isolation and characterization of two germacrene A synthase cDNA clones from chicory, *Plant Physiology* 129 (2002) 134-144.
- [21] M. Bennett, J. Mansfield, M. Lewis, M. Beale, Cloning and expression of sesquiterpene synthase genes from lettuce (*Lactuca sativa* L.), *Phytochemistry* 59 (2002) 255-261.
- [22] I. Prosser, A. Phillips, S. Gittings, M. Lewis, A. Hooper, J. Pickett, M. Beale, (+)-(10R)-germacrene A synthase from goldenrod, *Solidago canadensis*; cDNA isolation, bacterial expression and functional analysis, *Phytochemistry* 59 (2002) 691-702.
- [23] M. Kim, Y. Chang, M. Bang, N. Baek, J. Jin, C. Lee, S. Kim, cDNA isolation and characterization of (+)-germacrene A synthase from *Ixeris dentata* form. *albiflora* Hara, *Journal of Plant Biology* (48) 2005 178-186.
- [24] C. Berteau, A. Voster, F. Verstappen, M. Maffei, J. Beekwilder, H. Bouwmeester, Isoprenoid biosynthesis in *Artemisia annua*: cloning and heterologous expression of a germacrene A synthase from a glandular trichome cDNA library, *Archives of Biochemistry and Biophysics* 448 (2006) 3-12.
- [25] J. Gopfert, G. MacNevin, D. Ro, O. Spring, Identification, functional characterization and developmental regulation of sesquiterpene synthases from sunflower capitulate glandular trichomes, *Bmc Plant Biology* 9 (2009) 86.
- [26] T. Thorup, B. Tanyolac, K. Livingstone, S. Popovsky, I. Paran, M. Jahn, Candidate gene analysis of organ pigmentation loci in the Solanaceae, *Proceedings of the National Academy of Sciences of the United States of America* 97 (2000) 11192-11197.
- [27] X. Chen, F. Salamini, C. Gebhardt C, A potato molecular-function map for carbohydrate metabolism and transport, *Theoretical and Applied Genetics* 102 (2001) 284-295.
- [28] T. Cadalen, M. Morchen, C. Blassiau, A. Clabaut, I. Scheer, J. Hilbert, T. Hendriks, M. Quillet, Development of SSR markers and construction of a consensus genetic map for chicory (*Cichorium intybus* L.), *Molecular Breeding* 25 (2010) 699-722.
- [29] E. Portis, G. Mauromicale, R. Mauro, A. Acquadro, D. Scaglione, S. Lanteri, Construction of a reference molecular linkage map of globe artichoke (*Cynara cardunculus* var. *scolymus*), *Theor. Appl. Genetics* 120 (2009) 59-70.
- [30] E. Portis, A. Acquadro, D. Scaglione, G. Mauromicale, R. Mauro, C.A. Taylor, S.J. Knapp, S. Lanteri, Construction of an SSR-based linkage map for *Cynara cardunculus*. In *8th Plant Genomics European Meeting*, 7- 10 October (2009) Lisbon.
- [31] B. Menin, S. Lanteri, C. Comino, A. Moglia, M. Barba, T.W.J.M. Van Herpen, J. Beekwilder, Setting up a callusogenesis and transformation system for globe artichoke [abstracts] *Journal of Biotechnology* 150S (2010) 497.
- [32] J. Sambrook, E.F. Fritsch, T. Maniatis, *Molecular cloning: a laboratory manual*. Second edition, Cold Spring Harbor: Laboratory Press, New York, USA. (1989).

- [33] D. Scaglione, A. Acquadro, E. Portis, C.A. Taylor, S. Lanteri, S.J. Knapp, Ontology and diversity of transcript-associated microsatellites mined from a globe artichoke EST database, *BMC Genomics* 10 (2009) 454.
- [34] J. Degenhardt, T. Kollner, J. Gershenzon, Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants, *Phytochemistry* 70 (2009) 1621-1637.
- [35] S. Kumar, K. Tamura, M. Nei, MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment, *Briefings in Bioinformatics* 5 (2004) 150-163.
- [36] J.W. Van Ooijen, JoinMap® v.4, Software for the calculation of genetic linkage maps in experimental populations, Wageningen, Netherlands: B.V. Kyazma (2006).
- [37] D. Kosambi, The estimation of map distances from recombination values, *Ann Eugen* 12 (1944) 172–175.
- [38] A. Hay, M. Hamburger, K. Hostettmann, J. Hoult, Toxic inhibition of smooth-muscle contractility by plant-derived sesquiterpenes caused by their chemically reactive alpha-methylenebutyrolactone functions, *British Journal of Pharmacology* 112 (1994) 9-12.
- [39] T.A. Van Beek, P. Maas, B.M. King, E. Leclercq, A.G.J. Voragen, A.E. De Groot, Bitter sesquiterpene lactones from chicory roots, *J. Agric. Food. Chem.* 38 (1990) 1035–1038.
- [40] M. Dittrich, *Cynarae*-systemic review, In *The biology and chemistry of the Compositae*, Edited by V.H. Heywood, J.B. Harborne, B.L. Turner, (1977)999-1015.
- [41] A. Moth Wiklund, *Arcyna*, a new genus segregated from *Cynara* (*Compositae*), *Willdenowia* 33 (2003) 63-68.
- [42] D. Nguyen, J. Gopfert, N. Ikezawa, G. MacNevin, M. Kathiresan, J. Conrad, O. Spring, D. Ro, Biochemical conservation and evolution of germacrene A oxidase in Asteraceae, *Journal of Biological Chemistry* 285 (2010) 16588-16598.
- [43] I.A. Graham¹, K. Besser, S. Blumer, C. A. Branigan, T. Czechowski, L. Elias, I. Guterman, D. Harvey, P.G. Isaac, A.M. Khan, T. R. Larson, Y. Li, T. Pawson, T. Penfield, A.M. Rae, D.A. Rathbone, S. Reid, J. Ross, M.F. Smallwood, V. Segura, T. Townsend, D. Vyas, T. Winzer, D. Bowles, The Genetic Map of *Artemisia annua* L. Identifies Loci Affecting Yield of the Antimalarial Drug Artemisinin, *Science* 327 (2010) 328-331.

Table 1: Primer sequences used to amplify segments of globe artichoke *GAS*.

Primer	Sequence (5'- 3')
Probe_GAS_For	CCTTCCATCAAGGGATGC
Probe_GAS_Rev	GCTCCGATGCACCAATTT
Expr_GAS_For	GGGTTTGGATCCAATGGCTGCAGTAGAAGCTAAC
Expr_GAS_Rev	CGAAGCTTCTACATTGGTAAAGAGGGAACAA
RT_GAS_For	CTGTCATCTTGGATATGTTC
RT_GAS_Rev	GATCTCGTCTTATCCAGTG
RT_Act_For	TACTTTCTACAACGAGCTTC
RT_Act_Rev	ACATGATTTGAGTCATCTTC
SNP_GAS_For	CATGCCATGGCTGCAGTAGAAGCTAAC
SNP_GAS_Rev	TCGAGTCTCGGTGAAAAGCTG
Out_GAS_For	TGCGAGATTATGATGAAGCTG
Out_GAS_Rev	GAAACCATGGTGCAGTAGCA
Inn_GAS_Rev	TCCCCTTTACTCCACTTAACCAG

Table 2: Chromatographic and spectral properties of compounds detected in globe artichoke tissues.

Peak	RT ^a (min)	[M+H] ^b (<i>m/z</i>)	fragments (<i>m/z</i>)	λ max ^c (nm)	Elemental composition	Calc [M+H] ^d (<i>m/z</i>)	Identity
1	10.9	355.1008	163	325	C ₁₆ H ₁₈ O ₉	355.1023	chlorogenic acid
2	15.5	595.1632	449, 287	349, 254	C ₂₇ H ₃₀ O ₁₅	595.1657	luteolin rutinoside
3	16.1	449.1087	287	349, 254	C ₂₁ H ₂₀ O ₁₁	449.1078	luteolin- glucoside
4	17.4	517.1338	499	329, 245	C ₂₅ H ₂₄ O ₁₂	517.1340	dicafeoyl- quinic acid
5	21.9	347.1503	245, 385, 369, 675	212	C ₁₉ H ₂₂ O ₆	347.1489	cynaropicrin

^a RT: retention time; ^b [M+H]: observed molecular ion; ^c λ max: absorption maximum; ^d Calc [M+H]: calculated molecular mass.

Table 3: Relative concentrations of metabolites in globe artichoke tissues.

	Mature leaf	Old leaf	Receptacle	Outer bract	Callus
chlorogenic acid	39 ± 2	91 ± 5	100^a ± 3	74 ± 5	69 ± 5
luteolin-7-rutinoside	3 ± 0.1	100 ± 8	22 ± 5	5 ± 0.3	4 ± 0.4
luteolin-7-glucoside	0.5 ± 0.1	100 ± 17	73 ± 14	8 ± 0.4	3 ± 0.9
dicaffeoyl quinic acid	60 ± 2	81 ± 5	100 ± 6	90 ± 4	62 ± 10
cynaropicrin	69 ± 2	100 ± 13	3 ± 0.2	0 ± 0	11 ± 3

^a Concentrations were compared by measuring mass signals of the molecular ion ([M+H]) in different tissues and normalisation to the highest concentration (100%, indicated in bold).

Figure legends

Fig. 1: Proposed pathway for the biosynthesis of (+)-costunolide, an intermediate in sesquiterpene lactones (STLs) synthesis. GAS: The cyclization of farnesyl diphosphate (FPP) to (+)-germacrene A by GAS; GAO: the oxidation of germacrene A to germacrene A 1(10),4,11(13)-trien-12-oic acid by germacrene A oxidase; COS: costunolide synthase. (+)-costunolide is the branching point from where STLs divide into the three major sesquiterpene groups: germacranolides, eusmanolides and guaianolides. Cynaropicrin is synthesized in subsequent steps that lead to change the typical tricyclic guaianolide skeleton (adapted from de Kraker *et al.*, [17]).

Fig. 2: Guided phylogenetic analysis of sesquiterpene synthases, based on MEGA v4.0 software. The length of the lines indicates the relative distances between nodes. Protein sequences used for the alignment: chicory Ci_GAS_short (AAM21659); lettuce Ls_GAS1 (AAM11626); *Crepidiastrum sonchifolium* Cs_GAS (ABB00361); sunflower Ha_GAS3 (ACZ50512); lettuce Ls_GAS2 (AAM11627); *Artemisia annua* Aa_GAS (ABE03980); sunflower Ha_GAS1 (ACA14463); sunflower Ha_GAS2 (ABY49939); globe artichoke Cc_GAS (JN383985, this paper); chicory Ci_GAS_long (AAM21658); *Ixeris dentata* var. *albiflora* Id_guaiadiene_synth (AAL92481); *Solidago canadensis* Sc_casc_synth (AAT72931); sunflower Ha_γ_cadin_synth. γ_cadinene synthase (AAY41422); *Solidago canadensis* Sc_GAS (CAC36896); *Pogostemon cablin* Pc_GAS (AAS86321); tobacco Nt_TEAS (AAA19216).

Fig. 3: GC-MS analysis of sesquiterpene product of the GAS gene. GC-MS analysis on reaction product with farnesyl diphosphate as substrate, using: A) pACYC Duet plasmid with recombinant enzyme (CcGAS) and pACYC Duet plasmid empty as negative control. B) Spectrum for peak identification by NIST as β-elemene which is a cope-rearrangement of (+)-germacrene A.

Fig. 4: GAS expression in globe artichoke. Gene activity in mature (6 week old) and old (20 week old) leaves, outer bracts, receptacle and calli was measured by qRT-PCR, using actin as a reference gene, and cDNA synthesized from RNA extracted from old leaves as calibrator. GAS expression in bracts was below the detection threshold.

Fig. 5: LC-MS chromatogram showing the total ion count of mass signals (positive mode; 312 nm) of mass signals recorded in an extract from a mature leaf. Indicated peaks (1: chlorogenic acid; 2: luteolin rutinoside; 3: luteolin glucoside; 4: dicaffeoylquinic acid; 5: cynaropicrin) are described in Table 1.

Fig. 6: Linkage mapping of *GAS*. The original cultivated cardoon linkage group ('Alt_5') [29] is placed on the left and changed marker orders, compared to the new developed map, are indicated by dotted lines. The position of *GAS* is shown as a grey box. Asterisks indicate markers showing significant levels of segregation distortion (*: $0.1 > P \geq 0.05$). 'Alt_5' linkage group (in blue) is then aligned with the consensus linkage group LG_10 (in green) obtained by the comparison with its globe artichoke homologue C3_7 (in yellow, on the right). This alignment is based exclusively on common microsatellite loci and SNP markers. 'r-' and 'a-' indicate markers segregating only in 'Romanesco C3' and 'Altilis 41', respectively.

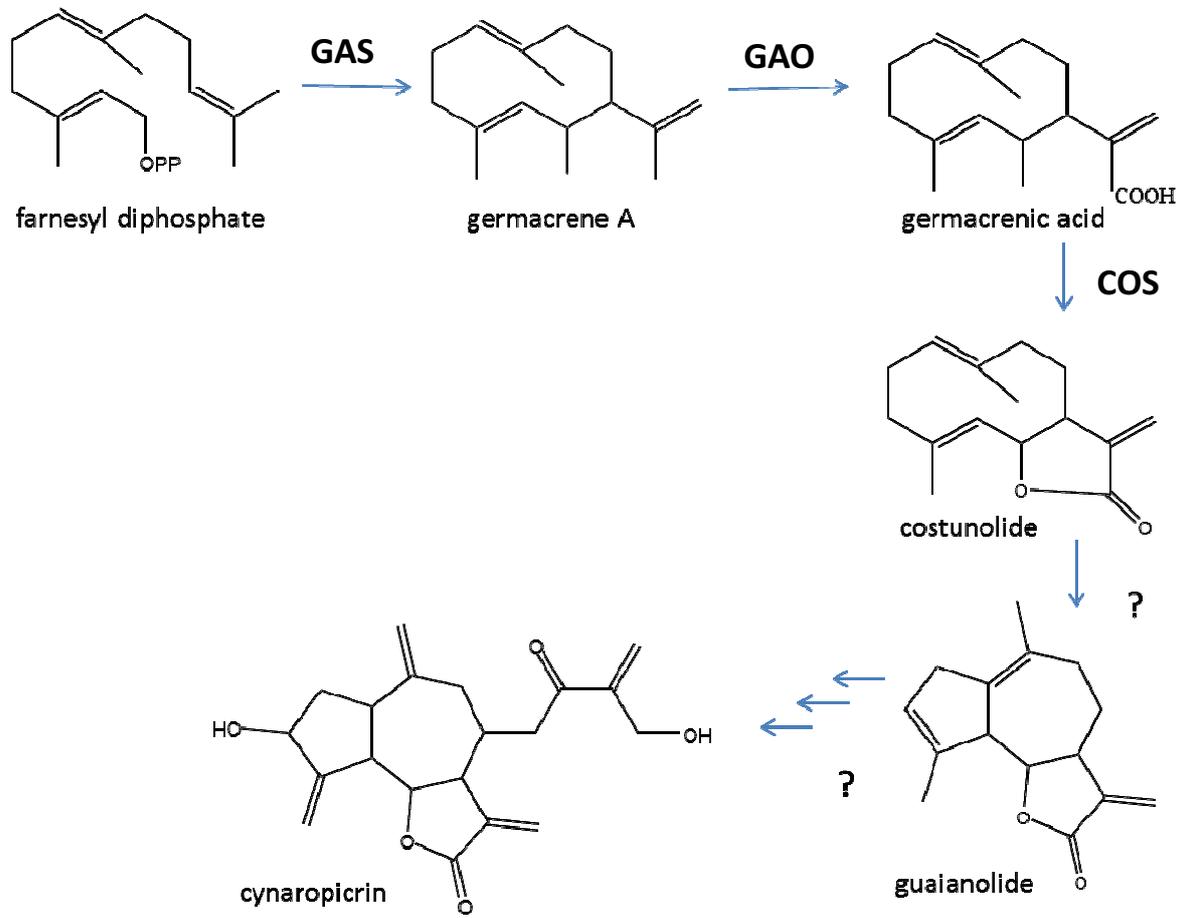


Figure 1

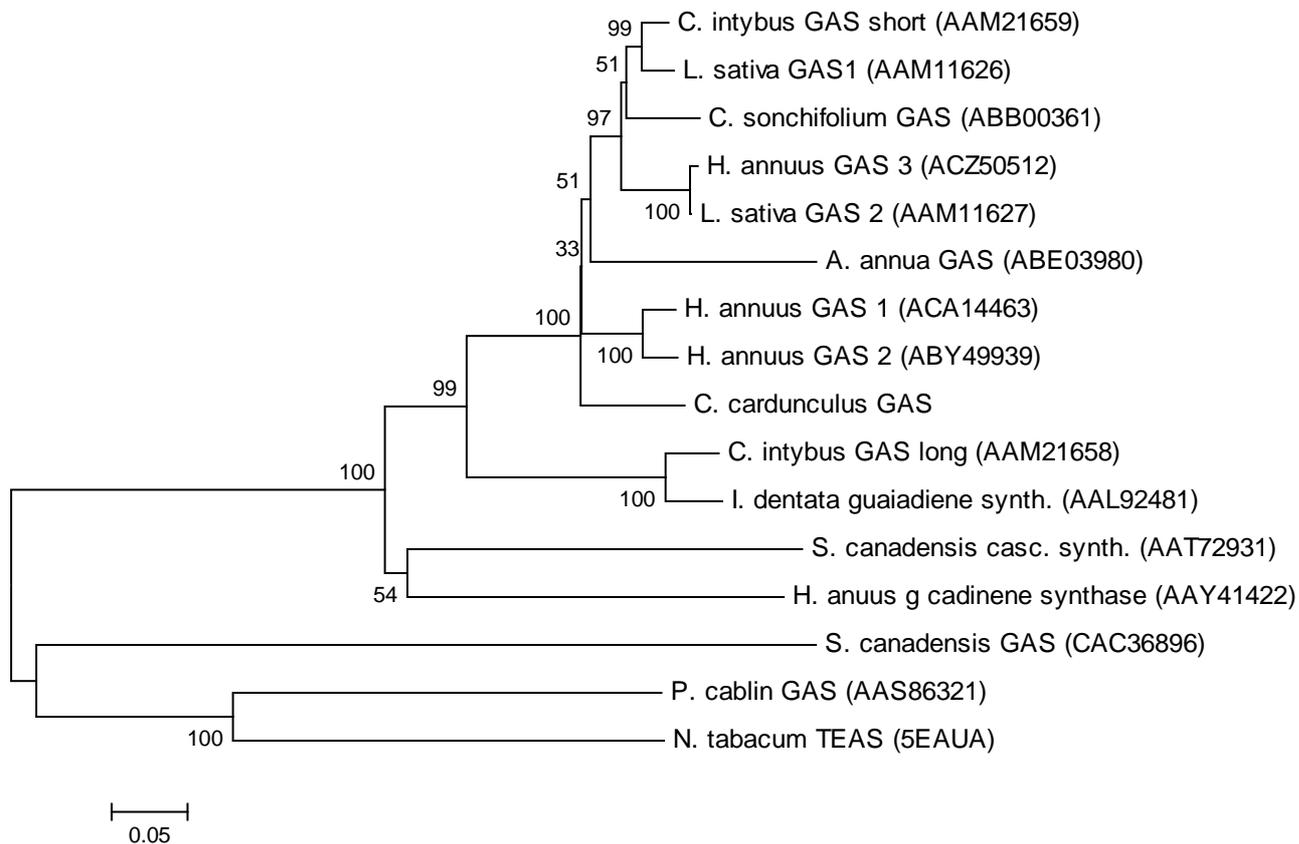


Figure 2

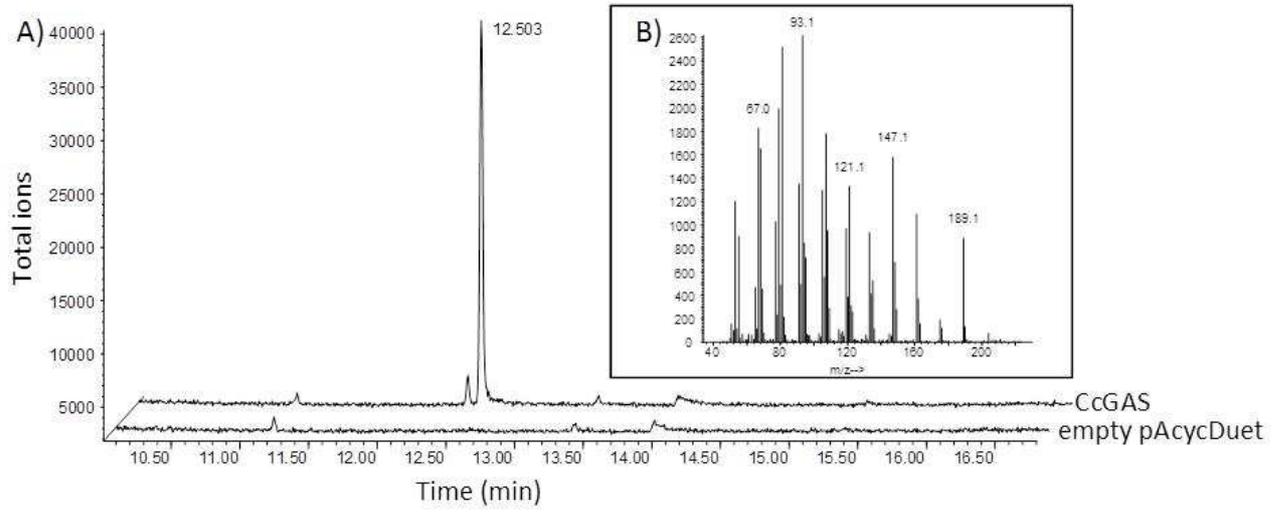


Figure 3

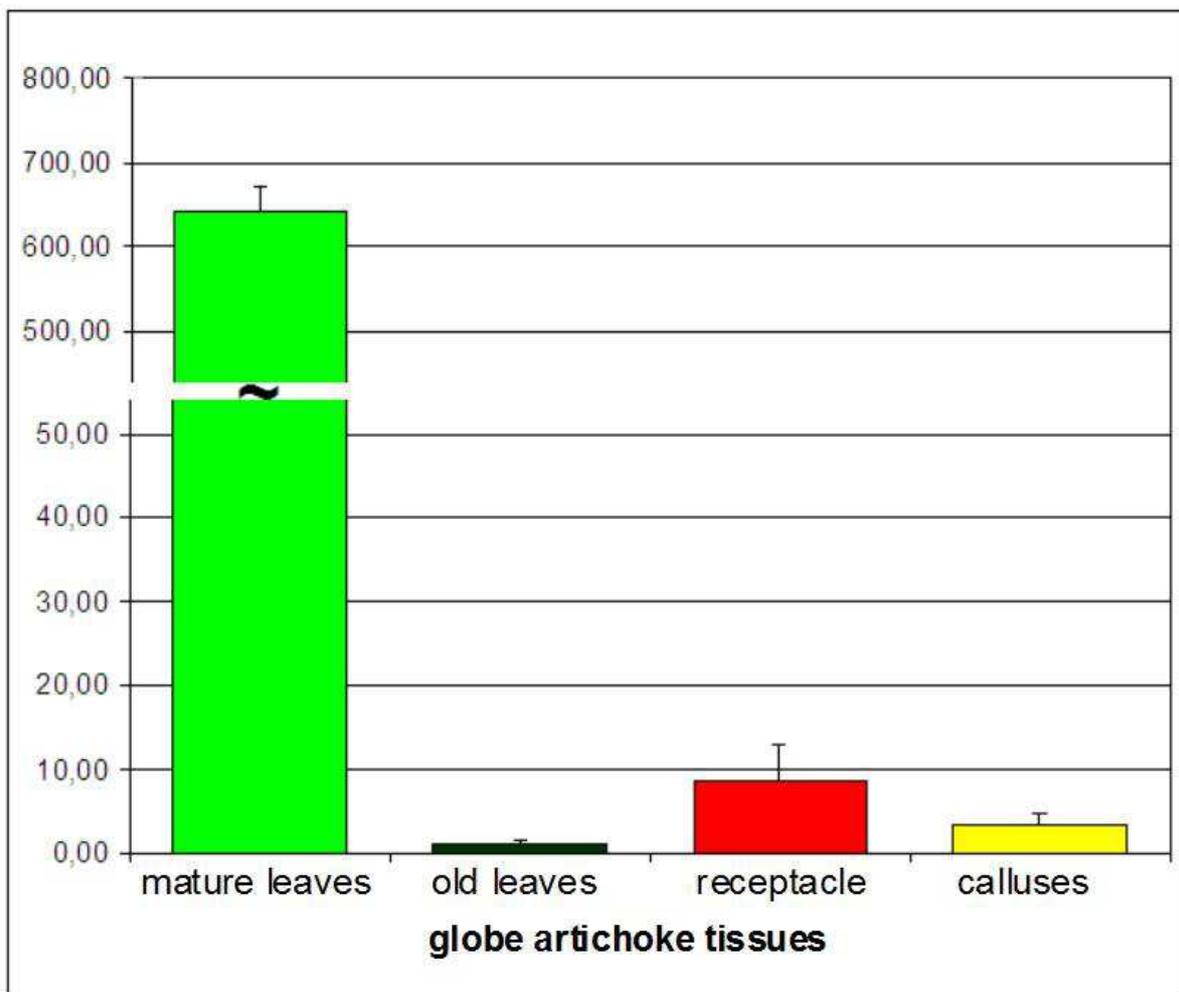


Figure 4

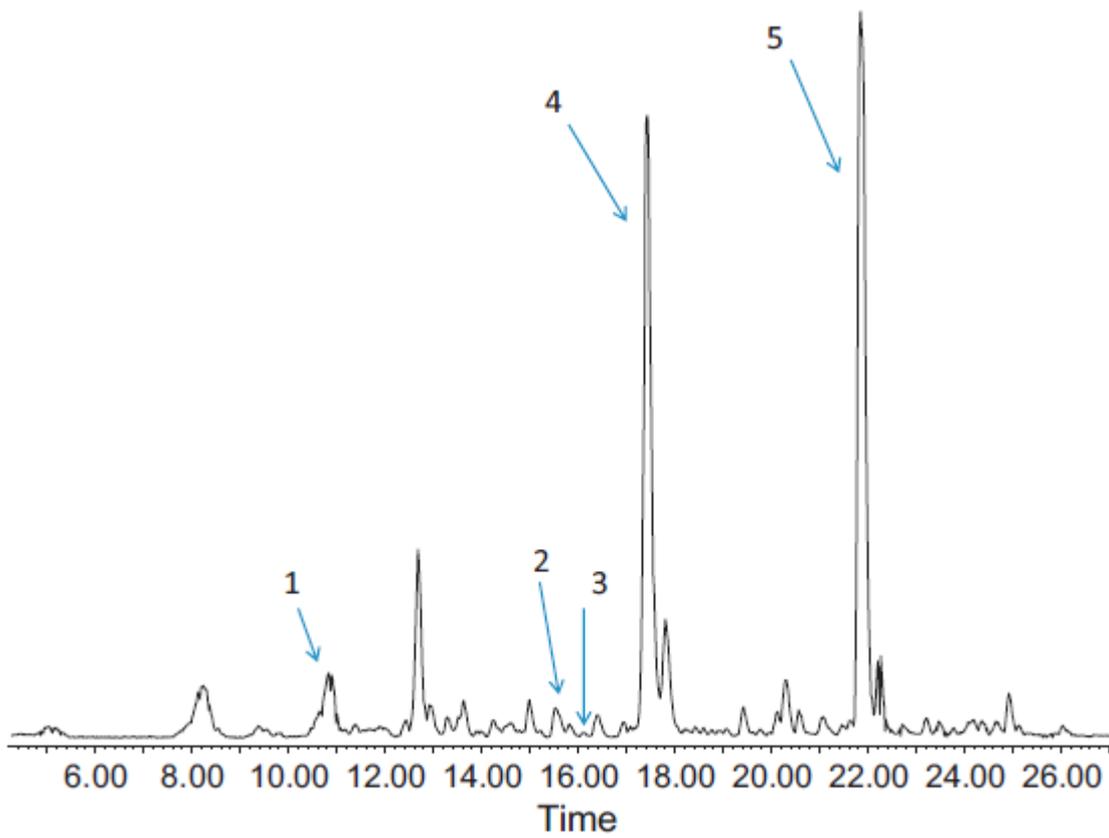


Figure 5

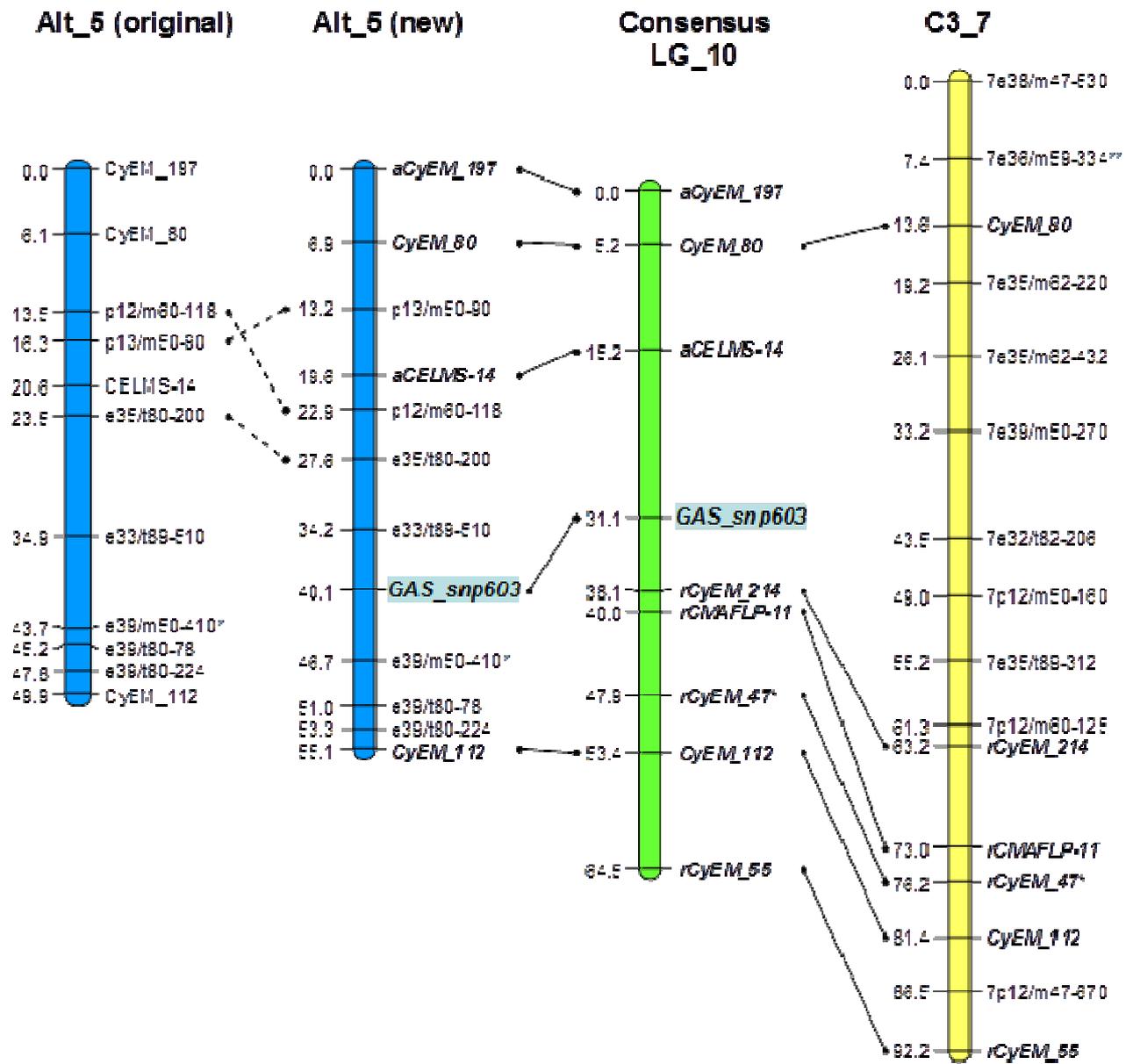


Figure 6