IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Accumulation of cynaropicrin in globe artichoke and localization of enzymes involved in its biosynthesis

This is the author's manuscript

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/1546262 since 2016-10-18T12:18:17Z

Published version:

DOI:10.1016/j.plantsci.2015.07.020

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)

17 September 2026

IRIS AperTO

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, 239, 2015, 10.1016/j.plantsci.2015.07.020.

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.2015.07.020

The publisher's version is available at:
http://linkinghub.elsevier.com/retrieve/pii/S0168945215300261

When citing, please refer to the published version.

Link to this full text:
http://hdl.handle.net/2318/1546262

This full text was downloaded from iris - AperTO: https://iris.unito.it/

iris - AperTO

University of Turin's Institutional Research Information System and Open Access Institutional Repository

Accumulation of cynaropicrin in globe artichoke and localization of enzymes involved in its biosynthesis

Eljounaidi K[a], Comino C[a], Moglia A[a]*, Cankar K[b,c], Genre A[d], Hehn A[e,f], Bourgaud F[e,f], Beekwilder J[b], Lanteri S[a]

[a]Department of Agricultural, Forest and Food Sciences, University of Torino, Via L. da Vinci 44, 10095 Grugliasco, Italy

[b]Plant Research International, P.O. Box 16, 6700 AA Wageningen, The Netherlands

[c]Laboratory of Plant Physiology, Wageningen University, P.O. Box 658, 6700AR Wageningen, The Netherlands

[d]Department of Life Sciences and Systems Biology, University of Turin, Viale Mattioli 25, 10125 Turin, Italy

[e]Université de Lorraine, UMR 1121 Agronomie et Environnement, 2 avenue de la Forêt de Haye, TSA 40602, 54518 Vandoeuvre-lès-Nancy, France

[f]INRA, UMR 1121 Agronomie et Environnement, 2 avenue de la Forêt de Haye, TSA 40602, 54518 Vandoeuvre-lès-Nancy, France

* author for correspondence

Andrea Moglia

Department of Agricultural, Forest and Food Sciences

andrea.moglia@unito.it

Largo Paolo Braccini 2, Grugliasco (TO), 10095

+39-0116708810

Keywords: globe artichoke, cynaropicrin, sesquiterpene lactone, trichomes

ABSTRACT

Globe artichoke (*Cynara cardunculus* var. *scolymus*) belongs to the Asteraceae family, in which one of the most biologically significant class of secondary metabolites are sesquiterpene lactones (STLs). In globe artichoke the principal STL is the cynaropicrin, which contributes to approximately 80% of its characteristic bitter taste. Cynaropicrin content was assessed in globe artichoke tissues and was observed to accumulate in leaves of different developmental stages. In the receptacle, a progressive decrease was observed during inflorescence development, while the STL could not be detected in the inflorescence bracts. Almost undetectable amounts were found in the roots and inflorescence stems at the commercial stage. Cynaropicrin content was found to correlate with expression of genes encoding CcGAS, CcGAO and CcCOS, which are involved in the STL biosynthesis.

A more detailed study of leaf material revealed that cynaropicrin predominantly accumulates in the trichomes, and not in the apoplastic cavity fluids. Analysis of the promoter regions of CcGAO and CcCOS revealed the presence of L1-box motifs, which confers trichome-specific expression in Arabidopsis, suggesting that cynaropicrin is not only stored but also synthesized in trichomes. A transient expression of GFP fusion proteins was performed in *Nicotiana benthamiana* plants: the CcGAS fluorescence signal was located in the cytoplasm while the CcGAO and CcCOS localized to the endoplasmatic reticulum.

1. Introduction

Globe artichoke (*Cynara cardunculus* var. *scolymus*) belongs to the Asteraceae family. Its immature inflorescences (heads or *capitula*) are consumed as a vegetable. The inflorescences are surrounded by many bracts, of which the most inner as well as the receptacle are fleshy and edible. The leaves of globe artichoke and its related taxa, the cultivated (var. *altilis*) and the wild (var. *sylvestris*) cardoon, have been used since ancient times as herbal medicine. Therapeutic effects of *Cynara* extracts have been observed in several clinical trials [1-3] and associated with the content of phenylpropanoids (mono- and di-caffeoylquinic acids, flavonoids) and sesquiterpene lactones (STLs) [4, 5]. STLs constitute a class of terpenoids which are remarkably diverse in terms of their structure and properties, which play a key role in plant environment interaction and possess allelochemical, deterrent and insecticidal properties [6, 7].

STLs, which have primarily been reported from the Compositae family [7, 8], derive from farnesyl diphosphate and have a characteristic α-methylene-γ-lactone moiety on their C15 backbone [9]. In globe artichoke and related taxa, two predominant STLs have been described, including cynaropicrin, which contributes to approximately 80% to their peculiar bitter taste, and grosheimin, which is present in much lower amounts [10]. Cynaropicrin has been demonstrated to possess a variety of biological activities. Cho et al. [11] report an inhibitory effect on the production of proinflammatory cytokines and its immune-modulatory properties, suggesting its potential application to diseases such as virus-induced chronic inflammation and leukocyte cancer cells. In recent studies it was also evidenced that cynaropicrin prevents skin photo-aging and possesses a marked effect on mucosal injuries, preventing acute gastritis [12]. In addition, grosheimin was found to display a strong antioxidant and anti-allergic activity [13].

STLs, which often display cytotoxic properties, are generally separated from the cellular metabolism and stored in specialized cells and compartments such as cavities in *Solidago canadensis*, ducts in *Ambrosia trifida*, laticifers in *Lactuca sativa* or glandular trichomes in *Artemisia annua* and *Helianthus annuum* [14-17]. Glandular trichomes are specialized epidermal structures, formed by two rows of four to six cells and a terminal head. They cover the surface of many plant tissues among which *C. cardunculus* leaves [18].

The first three committed steps in the STL biosynthetic pathway have been elucidated in several *Asteraceae* species [19-23], including *C. cardunculus* [24, 25], and involve three key enzymes: the terpene synthase germacrene A synthase (CcGAS in *C. cardunculus*) and two cytochrome P450s: the germacrene A oxidase (CcGAO) and the costunolide synthase (CcCOS). The objective of this work was to investigate the accumulation and localization of cynaropicrin in *C. cardunculus* and its correlation with the transcriptional expression of the three key genes involved in STL biosynthesis.

In addition, the subcellular localization of CcGAS, CcGAO and CcCOS has been investigated by using confocal fluorescence microscopy.

2. Materials and methods

2.1 Plant material

Globe artichoke plants (F1 hybrid 'Concerto', Nunhems) were grown in the experimental field at Carmagnola (Torino) up to the production of commercial inflorescences (heads). Plant material was harvested at different developmental stages, as described below, and stored at -80°C until further uses: i) leaves from 6 weeks, 20 weeks and 1 year old plants; ii) whole immature inflorescences at the stages before (stage 1) and after (stage 2) the development of their stem as well as at the commercial stage (stage 3), from which also the receptacle and external bracts were collected); iii) stems of the primary head and roots at the commercial stage of the inflorescence.

Seeds of the globe artichoke Romanesco varietal types 'Concerto' and 'C3' and of cultivated (accession 'Altilis 41') and wild (accession 'Creta 4') cardoon were germinated for two weeks between two layers of wetted filter paper; plantlets were then transplanted into pots in a glasshouse and grown in a climate room with 16 h day length and day/night at the temperature of 24/18°C. Leaves from 14 weeks old plants were harvested and stored at -80°C until further uses.

2.2 Extraction of metabolites and LC-QTOF-MS analysis

Secondary metabolite extraction and liquid chromatography-quadrupole time of flight-mass spectrometry (LC-QTOF-MS) analyses were performed according to the protocols previously described [24]. Globe artichoke plant tissues at different developmental stages (leaves, receptacle, external bracts, roots and stems), trichomes and apoplastic fluids as well as leaf tissues of globe artichoke C3 and cultivated and wild cardoon were lyophilized and ground to a fine powder in liquid nitrogen. Three biological replicates were used for each sample. Each replicate (50 mg) was extracted with 1 ml of 75% methanol and 0.1% formic acid and sonicated (125 W, 20 kHz) for 15 min. The extracts were then centrifuged at 20000 x g for 5 min at 22°C and filtered through a 0.2 μm inorganic membrane filter (RC4, Sartorius, Germany), fitted into a disposable syringe, and transferred to a glass vial. The LC-QTOF-MS platform consisted of a Waters Alliance 2795 HT HPLC system equipped with a Luna C18(2) pre-column (2.0 × 4 mm) and an analytical column (2.0 × 150 mm, pore size of 100 Å, particle size of 3 μm; Phenomenex), connected to an Ultima V4.00.00 QTOF mass spectrometer (Waters, MS Technologies). Degassed eluent A, ultra-pure water: formic acid (1000:1, v/v), and eluent B, acetonitrile: formic acid (1000:1, v/v) were used at a flow rate of 0.19 mL min^{-1}. The gradient started at 5% B and increased linearly to 75% B in 45 min, afterwards the column was washed with 100% B and equilibrated at 5% A for 15 min before the next injection. The injection volume was 5 μL, the ionization was performed using an electrospray

ionization source, and masses were detected in the positive mode. Absolute quantification of the cynaropicrin was performed using a standard curve of an authentic standard of cynaropicrin (Analyticon, *m/z* 347 [M^+H]). The identification of phenylpropanoids was carried out by comparing retention times and masses with those reported in our previous publication [26] based on the use of the original standards: luteolin glucoside and rutinoside from Apin (Abingdon), cynarin (1,3-dicaffeoylquinic acid) from Carl Roth (Karlsruhe) and chlorogenic acid from SIGMA. Grosheimin was identified based on its accurate mass (*m/z* 263.516 [M^+H]).

A completely randomized design (CRD) was used for the comparison of the cynaropicrin concentration, metabolites relative concentration and relative gene expression (RGE) in different tissues. Mean comparison was conducted using Duncan's test. All the data were statistically analyzed using SPSS statistical software.

2.3 Apoplast and trichome analyses

The metabolic profiles of the leaf apoplast and glandular trichomes were examined by isolating the apoplastic fluids from globe artichoke leaf tissues as described by Joosten [27]. Fresh globe artichoke leaves, sampled from the 3 weeks old plants, were submerged into water and subjected to vacuum pressure at a rate of 20 kPa. When the air bubbling from leaves started to decrease, the vacuum was released to allow the water entering into the leaves. The latter were then placed into a 10 ml plastic vessel and centrifuged immediately at 3000 g for 10 min at 5°C. The collected apoplast fractions, the residual leaf tissues as well as not treated leaves (control) were stored at -80°C until further analysis.

The glandular trichomes were extracted by the chloroform dipping method reported by Duke et al. [28]. Fresh young leaves from 3 weeks old plants were collected and immersed into 5 ml of chloroform for 30 seconds. The chloroform extracts were then evaporated with a stream of N_2. Chloroform dipped and non-dipped leaves (control), as well as the glandular trichome extracts, were stored at -80°C until further analysis. Metabolic analysis of trichome was performed as described in the previous paragraph.

The different samples were examined using a Leica TCS SP2 confocal microscope equipped with 20X dry objective. Cell autofluorescence was excited at 488 nm and imaged with emission window at 500–560 nm and 560-700 nm respectively. All images were captured at 1024 X 1024 pixels.

2.4 Subcellular localization studies of CcGAS, CcGAO and CcCOS

The CcGAS, CcGAO and CcCOS genes were amplified from cDNA using the primer sequences (listed in Supplemental Table S1) and recombined into the pDONR207 Entry vector through a

Gateway strategy. The amplicons were cloned into pK7WGF2 [29] producing pK7-35S:GFP:CcGAS, pK7-35S:GFP:CcGAO and pK7-35S:GFP:CcCOS respectively. As a control, the vector for expression of GFP alone pK7WGF2 (under the control of the 35S promoter), and an endoplasmic reticulum-targeted GFP-KDEL construct were also agro-infiltrated in *Nicotiana benthamiana* leaves. The expression constructs pK7-35S:GFP:CcGAS, 35S:GFP:CcGAO and 35S:GFP:CcCOS and the pK7WGF2 vector alone (control) were transformed into *Agrobacterium tumefaciens* strain C5801. The obtained recombinant *A. tumefaciens* strains were grown at 28°C and 220 rpm for 24 h in 5 mL of LB media containing spectinomycin (100 mg L^{-1}) and tetracycline (10 mg L^{-1}). After an overnight, the culture cells were centrifuged for 20 min at 4000 g and 20°C, resuspended into 10 mM MES buffer containing 10 mM $MgCl_2$ and 200 µM acetosyringone (4′-hydroxy-3′,5′-dimethoxyacetophenone, Sigma) to a final OD_{600} of 0.5-1.0, and incubated at room temperature under gentle shaking at 50 rpm for 3 h. Bacteria were infiltrated into leaves of 5 weeks old *N. benthamiana* plants, grown from seeds on soil in a climate chamber with 16 h light at 25°C (16 h)/25°C (8 h), using a 1 mL syringe by slowly injection into the abaxial side of the leaf. The expression and the localization of the three genes were analyzed at 4 days post-agroinfiltration in small leaf samples (~0.5 cm^2 leaf material from at least three independent agro-infiltrated plants) by confocal laser scanning. All images were acquired and processed using a Leica TCS SP2 confocal microscope and software (Leica Microsystems GmbH,Wetzlar, Germany). GFP and plastid fluorescence were excited at 488 nm and recorded at 500-525 nm and 600-640 nm, respectively. A scanning resolution of 1024 × 1024 pixels was chosen and serial optical sections were acquired with either 1 or 2 µm resolution along the *z*-axis.

2.5 Bioinformatic analyses

Prediction of the subcellular localization of CcGAS, CcGAO and CcCOS enzymes was carried out through prediction programs PREDOTAR version 1.03 (http://urgi.versailles.inra.fr/predotar/) [30] and TARGETP 1.1 Server [31] (http://www.cbs.dtu.dk/services/ TargetP/). Promoter sequences of CcGAS, CcGAO and CcCOS were obtained from globe artichoke genome [32]. TSSP software (http://linux1.soft-berry.com/berry.phtml) and PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) [33] database were used to analyze the promoter regions of the three genes.

2.6 Gene expression analysis in globe artichoke tissues and in transformed N. benthamiana

Real-Time quantitative PCR analysis was performed in two tissues: leaves from 20 weeks old plants and stems of the main head collected at the commercial stage of the capitula. RNA was

isolated from 100 mg of globe artichoke tissues using Trizol reagent (Invitrogen) according to the manufacturer's instructions. The total RNA was quantified and controlled for purity using a spectrophotometer and agarose gel electrophoresis. The reverse transcription reaction was carried out using the MMLV Reverse Transcriptase 1st-Strand cDNA Synthesis Kit (Euroclone) and 1 µg of total RNA. Primer 3 software (frodo.wi.mit.edu/cgi-bin/primer3/primer3 www.cgi) was used for qPCR primer design of CcGAS, CcGAO and CcCOS and primers have been designed with a Tm of 60°C (Supplemental Table S2)[24, 25].

As a housekeeping gene for globe artichoke, actin (amplified with the primer combination qPCR-CcAct_F and qPCR-CcAct_R, Supplemental Table S2) was chosen for its stability and level of expression, comparable to one of the genes of interest, and whose expression remains stable in all tissues [24]. 20 µl RT-qPCRs were performed in three biological replicates for each of the tested tissue in the presence of fluorescent dye (GoTaq® qPCR Master Mix, Promega). PCR reactions were carried out in 48-well optical plates using the iCycler Real-time PCR Detection System (Bio-Rad Laboratories, USA). The PCR conditions comprised an initial incubation of 95°C/5 min, followed by 40 cycles of 95°C/5 s and 60°C/45 s. Melting curve analysis was performed at the end of the amplification process in order to assess the specificity of the primers. In all experiments, appropriate negative controls containing no template were subjected to the same procedure. Standard curves were analyzed using iCycler iQ software. Amplicons were analysed by the comparative threshold cycle method, in which ΔΔCt is calculated as ΔCtI- ΔCtM, where ΔCtI is the Ct value for the any target gene normalized to the endogenous reference gene and ΔCtM is the Ct value for the calibrator, which is also normalized to the reference gene.

For the semi-quantitative RT-PCR analysis of expression of the CcGAS, CcGAO and CcGAS genes in transient transformed tobacco plants, cDNAs were synthetized from 1.0 µg RNA from leaves of 2 *gas*, *gao* and *cos* transformants and of two plants transformed with the empty pK7WGF2 vector, using High Capacity RNA-to-cDNA kit (Applied Biosystem) according to the manufacturer's instructions. Semi-quantitative RT-PCR amplifications were performed with specific primers for the globe artichoke *gas*, *gao* and *cos* (Supplemental Table S2) and tobacco elongation factor (EF) as housekeeping gene (qPCR-NtEF_F and qPCR-NtEf_R, Supplemental Table S2), under standard PCR amplification conditions.

3. Results

3.1 Phytochemical analysis of the different tissues

Cynaropicrin was quantified in several globe artichoke tissues by liquid chromatography coupled to mass spectrometry (LC-MS) analysis on methanol extracts of freeze dried plant material, and comparison to an original standard. This compound was detected by the LC-ESI-MS (positive mode) at 21.67 min showing characteristic ions at *m/z* 347 [M^+H], *m/z* 245 [$Mfragment^+H$], *m/z* 385 [M^+K] and *m/z* 369 [M^+Na] [10, 24]. Significant differences in cynaropicrin content between the analyzed tissues were observed ($p < 0.05$) (Fig. 1A). The highest concentrations (9.6 ± 0.4 mg g^{-1} DW) were detected in leaf tissues. About 0.4 mg g^{-1} DW of cynaropicrin was observed in the inflorescence receptacle at the earliest stage of development and then its content gradually decreased with the progress of receptacle development to reach a very low concentration when the receptacle reached the commercial stage (0.05 mg g^{-1} dry weight). No cynaropicrin was detected in the external bracts, regardless of the stage of inflorescence development and only traces were found in the roots and stems.

As expected, a wide number of other compounds, such as chlorogenic acid, 1,3-di-caffeoylquinic acid, luteolin rutinoside, luteolin glucoside and grosheimin, were detected in significant quantities in all the different globe artichoke tissues as shown in Table 1 and Supplementary Figure 1. Grosheimin was only detected in leaf tissues, where also the highest concentrations of chlorogenic acid, luteolin rutinoside, luteolin glucoside were found. Inflorescence receptacles at the early developmental stage showed relatively high amounts of chlorogenic acid as well as the highest content of 1,3-di-caffeoylquinic acid.

Quantification of cynaropicrin was also performed in leaf extracts from different taxa of *C. cardunculus* such as two globe artichoke genotypes (Concerto and Romanesco), wild cardoon (Creta 4) and cultivated cardoon (Altilis 41). In 14 weeks old leaves, only minor variation in cynaropicrin content was observed, with only a 17% difference between the highest (in wild cardoon) and the lowest (in F1 globe artichoke 'Concerto') concentrations (Fig. 1B).

The leaf and stem tissues, characterized by the highest and lowest cynaropicrin levels, were also tested for expression of CcGAS, CcGAO and CcCOS genes using RT-qPCR. The gene expression of GAS, GAO and CcCOS was 120, 62 and 10 fold higher in leaf than in the stem tissues (Fig. 2).

3.2 Cynaropicrin localization in globe artichoke leaves

The concentration of cynaropicrin in apoplastic fluids and glandular trichomes was investigated. The apoplast fluid was found to contain relatively low concentrations of cynaropicrin, while

methanol extracts of leaves with and without apoplast showed hardly any differences in cynaropicrin content (Fig. 3A). LC-MS analyses conducted on chloroform dipped and non-dipped leaves (control), as well as on the glandular trichome extracts showed that leaves without trichomes contained significantly (p<0.05) less amounts of cynaropicrin (Fig. 3B), while a high amount of this molecule was highlighted in trichome extracts. This result suggests that trichomes are the accumulation site of cynaropicrin. Globe artichoke leaves were also observed at the confocal microscope before and after their dipping in chloroform (Fig. 3C-F). The treatment did not cause visible damages to leaf epidermal cells and the trichomes looked shrunken.

3.3 In vivo expression of genes for subcellular localization studies

On the basis of Target P [31] and PREDOTAR software [30] a cytosolic targeting was predicted for CcGAS protein. The two cytochrome P450s proteins CcGAO and CcCOS were predicted to be targeted to the endoplasmic reticulum (ER) (Table 2). For the *in vivo* assessment of subcellular localization of the enzymes, *N. benthamiana* plants were infiltrated with agrobacteria suspension harboring the expression constructs pK7-35S:GFP:GAS, pK7-35S:GFP:GAO and pK7-35S:GFP:COS as well as the pK7WGF2 and the endoplasmic reticulum-targeted GFP-KDEL vector alone as controls. Expression of the fusion genes was confirmed using semi quantitative RT-PCR with gene specific primers (Fig. 4).

The subcellular localization of the different proteins was analyzed by confocal laser scanning microscopy (Fig. 5). pK7-35S: GFP: CcGAS transformed cells showed a GFP fluorescence pattern compatible with cytosolic diffusible proteins: fluorescence accumulated at the periphery of the cells, in cytoplasmic strands, and in the nucleoplasm (Fig. 5A, B). This pattern was very similar to that observed in cells transformed with free cytosolic GFP (Fig. 5G, H). By contrast, pK7-35S: GFP:CcGAO (Fig. 5 C, D) and pK7-35S: GFP: CcCOS (Fig. 5E, F) accumulated in the endoplasmic reticulum and nuclear envelope, with the same pattern observed in cells expressing the endoplasmic reticulum-targeted GFP-KDEL construct (Fig. 5I, J). Our confocal microscopy imaging of GFP fusion constructs thus confirmed the predicted localization of CcGAS, CcGAO and CcCOS proteins

3.4 Promoter motives analysis

We analysed the sequence upstream of the ORF of the *CcGAS*, *CcGAO* and *CcCOS* genes: respectively 2171 bp, 1941pb, 2602bp in length. The promoter regions of the biosynthetic genes were analyzed for motives that could indicate regulation of the gene expression (Table 3). A

number of hormone-, abiotic stress- and light-responsive elements were observed, in particular in the CcGAS upstream region. Notably, an L1-box motif, which is associated with trichome specific expression in Arabidopsis [34] was observed in the promoter regions of CcGAO and CcCOS, suggesting a role of trichomes in the biosynthesis of STLs in *C. cardunculus*.

4 Discussion

The major STL present in *Cynara cardunculus* is cynaropicrin, as has been demonstrated by several studies [24, 35]. In this work we show that (i) the highest concentration of cynaropicrin is found in leaves, regardless of their stage of development and age; (ii) during the inflorescence development the cynaropicrin content progressively decreases in the receptacle and it is always absent in external bracts; (iii) almost undetectable amounts of cynaropicrin are present in the roots and stems at the commercial stage of the inflorescence development. These results are consistent with previous results reported by Schneider and Thiele [35] in globe artichoke. These authors also detected a high concentration of cynaropicrin in leaves but could not detect it in roots and in completely developed inflorescences. More recently, Ramos et al. [36] confirmed that in cultivated cardoon cynaropicrin is mainly accumulated in leaf tissues and present in much lower concentrations in the receptacle as well as the outer inflorescence bracts. Since the cynaropicrin is responsible for the bitter taste of globe artichoke, its high accumulation in leaves might be related to the role of STLs as antifeedants: the bitter taste repel chewing insects and birds which break open cells when feeding [37].

For other *Asteraceae* species, different tissue-specific accumulation has been reported. Majdi *et al*. [38] could not detect sequiterpene lactone parthenolide in *T. parthenium* L. roots but found high STL concentrations in flowers, and less in leaves and stems. Similarly, in *Arnica montana* L, Douglas *et al*. [39] reported that the STL levels were higher in flowers and increased progressively during their development. In these plants, the concentration of STLs was by far lower in receptacles and was very low in stems. De Kraker *et al*. [40] detected a high concentration of guianolides lactupicrin, lactucin, and 8-deoxylactucin in leaves and roots of *Cichorium intybus* L.

We also assessed the cynaropicrin content in leaf tissues of two globe artichoke genotypes (Concerto and C3), a cultivated cardoon genotype and a wild cardoon genotype. Interestingly, the highest amount of this compound was found in the ancestral wild form of the species, from which the two cultivated forms have been derived. Apparently, during domestication, humans have progressively selected forms with a lower content of sesquiterpenes, possibly due to their marked bitter taste.

Trichomes are glandular specialized structures covering the surface of plant leaves and have been studied in many *Asteraceae* species to elucidate the biosynthesis of terpenes [38]. We assessed the content in cynaropicrin of both leaf apoplastic cavity and trichomes, and we found out that the highest amount of this compound was present in the trichomes, suggesting they are likely the place where cynaropicrin is synthetized and stored. This hypothesis is strengthened by the presence, in CcGAO and CcCOS promoters, of the L1-box motif, that has been found to confer trichome-specific expression in Arabidopsis [34]. The presence of STLs in glandular trichomes has been

previously reported in many other plant species such as *A. annua* [41], *T. parthenium* L. [38] and *T. cinerariifolium* [23]. High throughput transcript sequencing on mRNA extracted from globe artichoke glandular trichomes might provide a tool for identifying new genes involved in the cynaropicrin biosynthetic pathway. Such an approach has been successfully adopted in other plant species for the identification of genes involved in the biosynthesis of several secondary metabolites [17, 42, 43]. In addition, trichomes, as a specialized structure with restricted communication to the rest of plant, represent a particularly interesting target for metabolic engineering [44]. Studying trichome-specific promoters can allow to bioengineer biochemical pathways present only in trichome cells without affecting plant development or productivity [45].

Costunolide is considered to be the common precursor of STLs and three enzymes involved in its biosynthetic pathway were recently described: i.e. the germacrene A synthase (CcGAS), the germacrene A oxidase (CcGAO) and the costunolide synthase (CcCOS)[25]. As expected, quantitative expression analyses showed that all three enzyme-coding genes are markedly more expressed in the leaves than in other tissues (Fig. 2), supporting a role in controlling the STL biosynthesis. Consistently with our work, many reports in different plant species showed a strong correlation between terpene amount (or emission) and the level of the corresponding mRNA, indicating that the regulation of terpenoid biosynthesis mainly occurs at the transcript level [46].

Most plant P450s were described to be targeted to the ER. For example, subcellular fractionation studies and expression of a green fluorescent protein (GFP) fusion showed that a well-characterized plant P450 enzyme, cinnamate-4-hydroxylase (C4H), is exclusively localized to the ER [47]. Also CYP86A1, encoding a fatty acid ω-hydroxylase involved in suberin monomer biosynthesis, was exclusively localized in the ER [48]. However, it has also been shown that P450-mediated reactions can also occur in the chloroplast, as suggested by several Arabidopsis P450 genes contain high Ser/Thr content in their N termini, indicative of chloroplast targeting. The synthesis of sesquiterpenes is believed to take place in the cytosol, while monoterpene synthases are supposed to be addressed to plastids as their precursor GPP. An *in silico* analysis using Target P predicts an endoplasmic reticulum (ER) targeting for CcGAO and CcCOS, and a cytoplasm addressing for CcGAS. These predictions were experimentally confirmed by using transient GFP fusion proteins expressed in *N. benthamiana* leaves.

The knowledge acquired on STL metabolism in *C. cardunculus* will contribute to enhance plant breeding programs aimed at higher vegetable quality. Future transcript profiling in combination with metabolite profiling (untargeted large scale metabolomics) of glandular trichomes, isolated from tissues in different developmental stages, will expand our insight into the spatial and temporal

accumulation of cynaropicrin in globe artichoke, and will facilitate studies aimed at identifying genes involved in the STL biosynthetic pathway.

ACKNOWLEDGMENTS

We thank Veronica Volpe (DIBIOS, University of Torino) for providing pK7WGF2 vector and Anna Maria Milani and Danila Valentino (both DISAFA, University of Torino) for excellent technical assistance. We thank Esmer Jongedijk and Desalegn Etalo (both PRI) for their help with trichome dipping and apoplast extraction.

References

[1] F. Perez-Garcia, T. Adzet, S. Canigueral, Activity of artichoke leaf extract on reactive oxygen species in human leukocytes, Free Radical Research, 33 (2000) 661-665.
[2] T. Rodriguez, D. Gimenez, R. Vazquez, Choleretic activity and biliary elimination of lipids and bile acids induced by an artichoke leaf extract in rats, Phytomedicine, 9 (2002) 687-693.
[3] 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.
[4] V. Lattanzio, P. Kroon, V. Linsalata, A. Cardinali, Globe artichoke: A functional food and source of nutraceutical ingredients, Journal of Functional Foods, 1 (2009) 131-144.
[5] 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.
[6] E. Rodriguez, G.H.N. Towers, J.C. Mitchell, Biological activities of sesquiterpene lactones, Phytochemistry, 15 (1976) 1573-1580.
[7] A. Picman, Biological activities of sesquiterpene lactone, Biochemical Systematics and Ecology, 14 (1986) 255-281.
[8] F.C. Seaman, Sesquiterpene lactones as taxonomic chacarcteristics in the *Asteraceae*, The Botanical Rewiev, 48 (1982) 123-595.
[9] N.H. Fischer, Biochemistry of the mevalonic acid pathway to terpenoids, in, New York: Plenum Press., 1990.
[10] 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.
[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] Y. Tanaka, K. Tanaka, H. Kojima, T. Hamada, T. Masutani, M. Tsuboi, Y. Akao, Cynaropicrin from *Cynara scolymus* L. suppresses photoaging of skin by inhibiting the transcription activity of nuclear factor-kappa B, Bioorganic & Medicinal Chemistry Letters, 23 (2013) 518-523.
[13] E. Yae, S. Yahara, M. El-Aasr, T. Ikeda, H. Yoshimitsu, C. Masuoka, M. Ono, I. Hide, Y. Nakata, T. Nohara, Studies on the Constituents of Whole Plants of *Youngia japonica*, Chemical & Pharmaceutical Bulletin, 57 (2009) 719-723.
[14] O. Spring, Chemotaxonomy based on metabolites from glandular trichomes, Advances in Botanical Research Incorporating Advances in Plant Pathology, 31 (2000) 153-174.
[15] A. Schilmiller, R. Last, E. Pichersky, Harnessing plant trichome biochemistry for the production of useful compounds, Plant Journal, 54 (2008) 702-711.
[16] R. Sessa, M. Bennett, M. Lewis, J. Mansfield, M. Beale, Metabolite profiling of sesquiterpene lactones from *Lactuca* species - Major latex components are novel oxalate and sulfate conjugates of lactucin and its derivatives, Journal of Biological Chemistry, 275 (2000) 26877-26884.

[17] C. Bertea, 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.
[18] N. Apostolo, C. Brutti, B. Llorente, Leaf anatomy of *Cynara scolymus* L. in successive micropropagation stages, In Vitro Cellular & Developmental Biology-Plant, 41 (2005) 307-313.
[19] 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.
[20] 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.
[21] K. Cankar, A. van Houwelingen, D. Bosch, T. Sonke, H. Bouwmeester, J. Beekwilder, A chicory cytochrome P450 mono-oxygenase CYP71AV8 for the oxidation of (+)-valencene, Febs Letters, 585 (2011) 178-182.
[22] 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.J. Bouwmeester, Reconstitution of the Costunolide Biosynthetic Pathway in Yeast and *Nicotiana benthamiana*, PLoS One, 6 (2011) e23255.
[23] A. Ramirez, N. Saillard, T. Yang, M. Franssen, H. Bouwmeester, M. Jongsma, Biosynthesis of Sesquiterpene Lactones in Pyrethrum (*Tanacetum cinerariifolium*), Plos One, 8 (2013).
[24] B. Menin, C. Comino, E. Portis, A. Moglia, K. Cankar, H. Bouwmeester, S. Lanteri, J. Beekwilder, 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, Plant Science, 190 (2012) 1-8.
[25] K. Eljounaidi, K. Cankar, C. Comino, A. Moglia, A. Hehn, F. Bourgaud, H. Bouwmeester, B. Menin, S. Lanteri, J. Beekwilder, Cytochrome P450s from *Cynara cardunculus* L. CYP71AV9 and CYP71BL5, catalyze distinct hydroxylations in the sesquiterpene lactone biosynthetic pathway, Plant Science, 223 (2014) 59-68.
[26] A. Moglia, S. Lanteri, C. Comino, A. Acquadro, R. De Vos, J. Beekwilder, J. (2008) Stress-induced biosynthesis of dicaffeoylquinic acids in globe artichoke. J. Agric. Food Chem. 56 (2008) 8641-8649.
[27] MH Joosten, Isolation of Apoplastic Fluid from Leaf Tissue by the Vacuum Infiltration-Centrifugation Technique, in, Methods Mol Biol, 835 (2012) 603-610.
[28] M. Duke, R. Paul, H. Elsohly, G. Stutzt, S. Duke, Localization of artemisin and artemisitene in foliar tissues of glanded and glandless biotypes of *Artemisia annua* International Journal of Plant Sciences, 155 (1994) 365-372.
[29] M. Karimi, D. Inze, A. Depicker, GATEWAY(TM) vectors for *Agrobacterium*-mediated plant transformation, Trends in Plant Science, 7 (2002) 193-195.
[30] I. Small, N. Peeters, F. Legeai, C. Lurin, Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences, Proteomics, 4 (2004) 1581-1590.
[31] O. Emanuelsson, S. Brunak, G. von Heijne, H. Nielsen, Locating proteins in the cell using TargetP, SignalP and related tools, Nature Protocols, 2 (2007) 953-971.
[32] D. Scaglione, S. Reyes Chin Wo, S. Lanteri, A. Acquadro, E. Portis, M. Tirone, L. Froenicke, C. Beitel, I. Korf, L. Rieseberg, R. Milchemore, De novo genome assembly and ultra-dense mapping of artichoke, chicory and endive, and syntenic comparisons across the Compositae, in, PAG San Diego California USA, 2014.

[33] M. Lescot, P. Dehais, G. Thijs, K. Marchal, Y. Moreau, Y. Van de Peer, P. Rouze, S. Rombauts, PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences, Nucleic Acids Research, 30 (2002) 325-327.
[34] S. Wang, J. Wang, N. Yu, C. Li, B. Luo, J. Gou, L. Wang, X. Chen, Control of plant trichome development by a cotton fiber MYB gene, Plant Cell, 16 (2004) 2323-2334.
[35] G. Schneider, K. Thiele, The distribution of the bitter principle cynaropicrin in the globe artichoke, Planta Medica, 26 (1974) 174-183.
[36] P. Ramos, A. Guerra, O. Guerreiro, C. Freire, A. Silva, M. Duarte, A. Silvestre, Lipophilic Extracts of *Cynara cardunculus* L. var. *altilis* (DC): A Source of Valuable Bioactive Terpenic Compounds, Journal of Agricultural and Food Chemistry, 61 (2013) 8420-8429.
[37] M. Chadwick, H. Trewin, F. Gawthrop, C. Wagstaff, Sesquiterpenoids Lactones: Benefits to Plants and People, International Journal of Molecular Sciences, 14 (2013) 12780-12805.
[38] M. Majdi, Q. Liu, G. Karimzadeh, M. Malboobi, J. Beekwilder, K. Cankar, R. de Vos, S. Todorovic, A. Simonovic, H. Bouwmeester, Biosynthesis and localization of parthenolide in glandular trichomes of feverfew (*Tanacetum parthenium* L. Schulz Bip.), Phytochemistry, 72 (2011) 1739-1750.
[39] J. Douglas, B. Smallfield, E. Burgess, N. Perry, R. Anderson, M. Douglas, V. Glennie, Sesquiterpene Lactones in *Arnica montana*: a rapid analytical method and the effects of flower maturity and simulated mechanical harvesting on quality and yield, Planta Medica, 70 (2004) 166-170.
[40] J. de Kraker, H. Bouwmeester, M. Franssen, A. de Groot, (+)-germacrene A synthesis in chicory (*Cichorium intybus* L.); the first step in sesquiterpene lactone biosynthesis, Acta Botanica Gallica, 146 (1999) 111-115.
[41] C. Bertea, 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.
[42] B. Lange, M. Wildung, E. Stauber, C. Sanchez, D. Pouchnik, R. Croteau, Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes, Proceedings of the National Academy of Sciences of the United States of America, 97 (2000) 2934-2939.
[43] M. McConkey, J. Gershenzon, R. Croteau, Developmental regulation of monoterpene biosynthesis in the glandular trichomes of peppermint, Plant Physiology, 122 (2000) 215-223.
[44] A. Tissier, Glandular trichomes: what comes after expressed sequence tags? Plant Journal, 70 (2012) 51-68.
[45] G. Gutierrez-Alcala, L. Calo, F. Gros, J. Caissard, C. Gotor, L. Romero, A versatile promoter for the expression of proteins in glandular and non-glandular trichomes from a variety of plants, Journal of Experimental Botany, 56 (2005) 2487-2494.
[46] D. Nagegowda, Plant volatile terpenoid metabolism: Biosynthetic genes, transcriptional regulation and subcellular compartmentation, Febs Letters, 35 (2010) 2965-2973.
[47] D. Ro, J. Ehlting, C. Douglas, Cloning, functional expression, and subcellular localization of multiple NADPH-cytochrome P450 reductases from hybrid poplar, Plant Physiology, 130 (2002) 1837-1851.
[48] R. Hofer, I. Briesen, M. Beck, F. Pinot, L. Schreiber, R. Franke, The Arabidopsis cytochrome P450 CYP86A1 encodes a fatty acid omega-hydroxylase involved in suberin monomer biosynthesis, Journal of Experimental Botany, 59 (2008) 2347-2360.

FIGURE CAPTIONS

Fig. 1. Cynaropicrin content (mg g^{-1} DW) in **(A)** different globe artichoke tissues: stem of the primary inflorescence (S), roots (R), external bract (EB stage 1, 2, 3), receptacle (R stage 1, 2, 3), leaves (L, 6, 20 weeks and 1 year); and in **(B)** 14 weeks old leaves from different *Cynara cardunculus* taxa: globe artichoke (varietal types: Concerto and C3), cultivated cardoon (Altilis 41) and wild cardoon (Creta 4). Bars represent means (n=3)±S.D. Different letters associated with the set of means indicate significance based on Duncan's test ($P \leq 0.05$).

Fig. 2. (A) Relative gene expression of CcGAS, CcGAO and CcCOS in leaf and stem tissues of *Cynara cardunculus*. Globe artichoke actin (CcAct) was used as gene of reference. Black bars represent CcGAS, grey bars represent CcGAO and white bars represent CcCOS. Error bars represent SD (n = 3). Asterisk indicates significance based on Duncan's test ($P \leq 0.001$). **(B)** Simplified biosynthetic pathway of cynaropicrin in globe artichoke.

Fig. 3. (A) Cynaropicrin content (mg g^{-1} DW) in globe artichoke leaf with and without apoplast as well as in extracted apoplastic fluids. **(B)** Cynaropicrin content (mg g^{-1} DW) in globe artichoke leaf with and without trichomes as well as in extracted trichomes. Bars represent means ±S.D (n=3). Different letters associated with the set of means indicate significance based on Duncan's test ($P \leq 0.05$). **(C-F)** Confocal microscopy images of trichomes on the surface of non-dipped **(C, D)** and chloroform dipped **(E, F)** globe artichoke leaves. Bar = 50 μm.

Fig. 4 Semi-quantitative RT-PCR analysis in transient transformed tobacco plants of pK7-35S: GFP:CcGAS (upper, left panel), pK7-35S:GFP:CcGAO (upper, middle panel), pK7-35S: GFP:CcCOS (upper, right panel). Controls correspond to transformed tobacco plants with empty vector.

Fig. 5. Subcellular localization of different GFP fusions in N. benthamiana leaf epidermal cells. In the right column panels (B, D, F, H, L) the GFP signal (green) is overlaid to

plastid autofluorescence (red). GFP-CcGAS (A, B) localizes to the cytoplasm and nucleoplasm (n), as expected for free diffusible cytosolic proteins; GFP-CcGAO (C, D) and GFP- CcCOS (E, F), by contrast, label the nuclear envelope (arrowhead), but not the nucleoplasm, strongly suggesting their incorporation in the endoplasmic reticulum/nuclear envelope compartment. The fluorescent patterns obtained with free cytosolic GFP (G, H), and GFP-KDEL, targeted to the endoplasmic reticulum and nuclear envelope (I, J) are shown for comparison. Bars = 20μm.

Supplementary Figure 1

LC–MS chromatogram showing the total ion count of mass signals recorded in a leaf extract of globe artichoke.

Tables

Table 1. Relative concentration of the major metabolites identified in globe artichoke (F1 hybrid Concerto) tissues. Concentrations were compared by measuring mass signals of the molecular ion ([M+H]) in different tissues. The highest MS area for each compound is indicated in bold. Different letters associated with the set of means indicate significance based on Duncan's test ($P \leq 0.05$).

	Cynarin (1,3-Dicaffeoylquinic Acid)	**Caffeoylquinic acid (chlorogenic acid)**	**Luteolin rutinoside**	**Luteolin glucoside**	**Grosheimin**
	m/z= 517.1338 RT=17.94	**m/z 355.1008 RT=10.69**	**m/z 595.1657 RT=15.72**	**m/z=449.1078 RT=16.1**	**m/z=263.516 RT=19.2**
Stem	1459±352 (b)	1438±481 (a)	12±10 (a)	1663±452(ab)	0±0(a)
Roots	661±500 (a)	1436±454 (a)	0±0 (a)	11±1 (a)	0±0(a)
External bracts stage 1	1894±31 (bc)	2678±437 (bc)	536±230 (ab)	5214±1474 (c)	0±0(a)
external bracts stage 2	1611±36 (bc)	2203±82 (ab)	199±37 (ab)	1997±1388(ab)	0±0(a)
External bracts stage 3	1661±514 (bc)	1938±318 (ab)	250±23 (ab)	2316±1321(b)	0±0(a)
Receptacle stage 1	**2247±761 (c)**	3296±307 (cd)	5±0 (a)	1494±143(ab)	0±0(a)
Receptacle stage 2	1496±15 (bc)	2296±504 (ab)	28±26 (a)	496±410(ab)	0±0(a)
Receptacle stage 3	1771±545 (bc)	2626±622 (bc)	84±65 (ab)	1262±954(ab)	0±0(a)
Leaf 6 weeks old	1746±319 (bc)	**4449±262(e)**	4553±496 (c)	**9734±420(d)**	113±16(b)
Leaf 20 weeks old	1809±41 (bc)	3715±115 (de)	1002±192 (b)	5901±1185(c)	144±22(c)
Leaf 1 year old	1572±252 (bc)	3224±108 (cd)	**5433±873 (c)**	9683±165(d)	**206±34(d)**

Table 2. The predicted location of the three proteins CcGAS, CcGAO and CcCOS by Target P (http://www.cbs.dtu.dk/services/TargetP/) and PREDOTAR version 1.03 (http://urgi.versailles.inra.fr/predotar/)

Gene	Full length	Molecular weight	Predicted localization
CcGAS	559 aa	64.55 kDa	Cytosol
CcGAO	486 aa	54.71 kDa	Secretory pathway (endoplasmic reticulum)
CcCOS	494 aa	55.88kDa	Secretory pathway (endoplasmic reticulum)

Table 3. Number of putative *cis*-acting regulatory elements identified in CcGAS, CcGAO, CcCOS promoters using PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) databases.

***Cis*-acting element**	**Signal sequence**	**Position in the Promoter**			**Expected function**
		CcGAS	CcGAO	CcCOS	
ABRE	ACGTG	2 (105, 1348)			*cis*-acting element involved in the abscisic acid responsiveness
Box 4	ATTAAT	1 (456)	4 (768, 858, 868, 1262)	1 (82)	part of a conserved DNA module involved in light responsiveness
Box I	TTTCAAA	1 (1297)		1 (1209)	light responsive element
CAAT-box	C/CAAT/T	23(422,517, 635..)	32(154,251, 275..)	38(248,380, 601..)	common cis-acting element in promoter and enhancer regions
GA-motif	AAAGATGA	1 (580)			part of a light responsive element
GARE-motif	TCTGTTG/AACAGA	2 (1023, 1186)	1 (1387)	1 (861)	gibberellin-responsive element
GATA-motif	GATAGGA/G	1 (444)	1(1025)		part of a light responsive element
GT1-motif	GGTTAA/AATCCACA	1 (1125)	1 (670)		light responsive element
HSE	AAAAAATTTC	2 (858, 941)			*cis*-acting element involved in heat stress responsiveness
I-box	TGATAATGT/CTCTTATGCT/GATAGGG	2 (234,311)	2(199, 1025)		part of a light responsive element
LTR	CCGAAA	1 (949)			*cis*-acting element involved in low temperature responsiveness
MBS	TAACTG	1 (1320)			MYB binding site involved in drought-inducibility
MRE	AACCTAA	1(415)			MYB binding site involved in light responsiveness
O2-site	GTTGACGTGA	1(61)			*cis*-acting regulatory element involved in zein metabolism regulation
Skn-1_motif	GTCAT	2(642,690)	1(1403)	8(271,652, 1292..)	*cis*-acting regulatory element required for endosperm expression
TATA-box	TTTTA/TATAAATT/TATA/TAATA/ATTATA/TATAA/ATATAT/TATATAA/TACATAAA	78 (323,397, 464..)	104(135, 218,471,..)	76(110,147, 473..)	core promoter element around -30 of transcription start
TC-rich repeats	ATTTTCTTCA/ATTTTCTCCA/ATTCTCTAAC	2 (1011, 1232)		3(168,388, 711)	*cis*-acting element involved in defense and stress responsiveness
as-2-box	GATAatGATG	6(232,250, 272,309,340, 391)			involved in shoot-specific expression and light responsiveness
chs-CMA1a	TTACTTAA	1(1238)			part of light responsive element
chs-CMA2a	TCACTTGA	1(74)			part of light responsive element
circadian	CAANNNNATC	4(232,309, 927,1034)	1(1031)	2(526,739)	*cis*-acting regulatory element involved in circadian control
ARE	TGGTTT		1(688)	1(1476)	cis-acting regulatory element essential for the anaerobic induction
AE-box	AGAAACAA/AGAAACTT		4(61,378, 842,1004)		part of a module for light response
CAT-box	GCCACT			2(1087, 1443)	cis-acting regulatory element related to meristem expression
CG-motif	CCATGGGG			1(1463)	part of light responsive element
L-box	TCTCACCTACC			2(638, 1012)	part of light responsive element
TCA-element	CCATCTTTTT/GAGAAGAATA		1 (404)	2 (1034, 1129)	*cis*-acting element involved in the salicylic acid responsiveness
L1-box	TGCATTA		1 (548)	1(2346)	motif confers trichome-specific expression in Arabidopsis
G-box	CACGTG/C	2(64,1348)	1(663)	4 (425,1001, 1089,1092)	*cis*-acting element involved in in light responsiveness

Supplementary Data

Table S1. Primer sequences used for the preparation of the GFP constructs

Primer	Sequence (5'-3')
attB1GAS	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGGCTGCAGTAGAAGCTAAC
attB2GAS	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTACATTGGTAAAGAGGGAAC
attB1GAO	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGGTACTCACTACTTCCATTGTT
attB2GAO	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAAAAACTTGGTACCAGTACCAA
attB1COS	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGGAACCTCTTACCATCGTCT
attB2COS	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCACTTGGCTAACGACTTGG

Table S2. Primer sequences used for the quantitative PCR experiments

Primer	Sequence (5'-3')
qPCR-CcAct_F	TACTTTCTACAACGAGCTTC
qPCR-CcAct_R	ACATGATTTGAGTCATCTTC
qPCR-CcGAS_F	TTGGTTTGACGTATCTATTC
qPCR-CcGAS_R	GTGTACAAATCAGCTTCATC
qPCR-CcGAO_F	AAGGATCAAAAAGAGTTCAC
qPCR-CcGAO_R	TTCGAAGGAAAAATATCAG
qPCR-CcCOS_F	ATATTCTTTTATAAAGCTCAGG
qPCR-CcCOS_R	CTCTAACGTACAAATCTTCTTC
qPCR-NtcEF_F	TGAGATGCACCACGAAGCTC
qPCR-NtEF_R	CCAACATTGTCACCAGGAAGTG

Figure 1

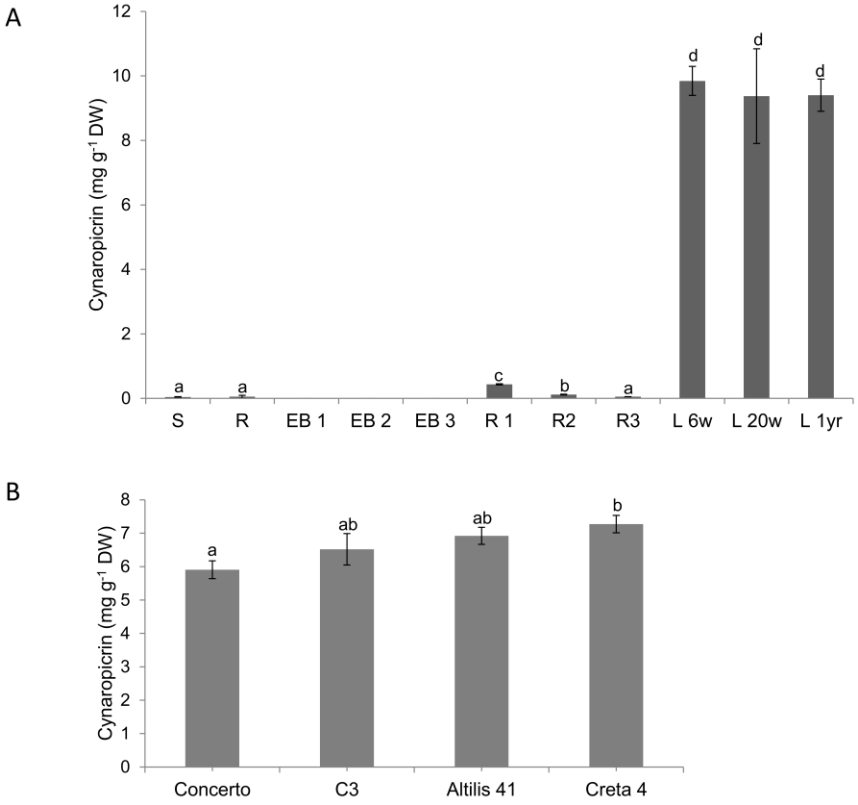

A
Cynaropicrin (mg g-1 DW)
12
10
8
6
4
2
0
a
a
c
b
a
d
d
d
S
R
EB 1
EB 2
EB 3
R 1
R2
R3
L 6w
L 20w
L 1yr
B
Cynaropicrin (mg g-1 DW)
8
7
6
5
4
3
2
1
0
a
ab
ab
b
Concerto
C3
Altilis 41
Creta 4

Figure 2

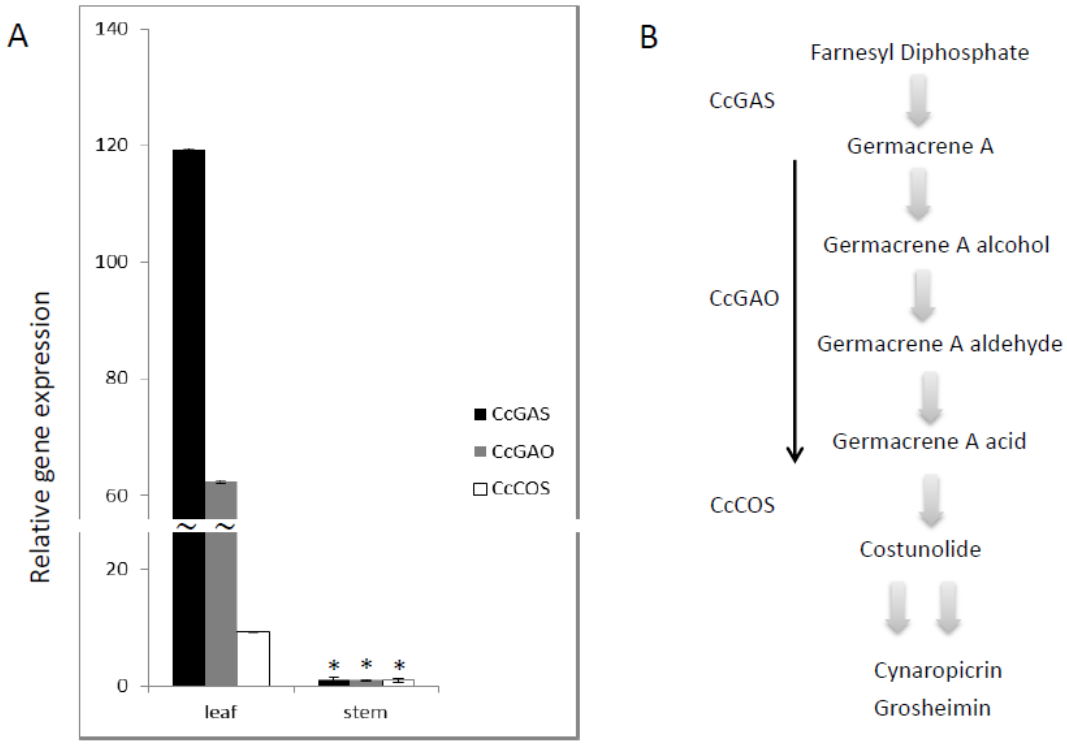

A
140
120
100
80
60
20
0
Relative gene expression
CcGAS
CcGAO
CcCOS
leaf
stem
B
Farnesyl Diphosphate
CcGAS
Germacrene A
Germacrene A alcohol
CcGAO
Germacrene A aldehyde
Germacrene A acid
CcCOS
Costunolide
Cynaropicrin
Grosheimin

Figure 3

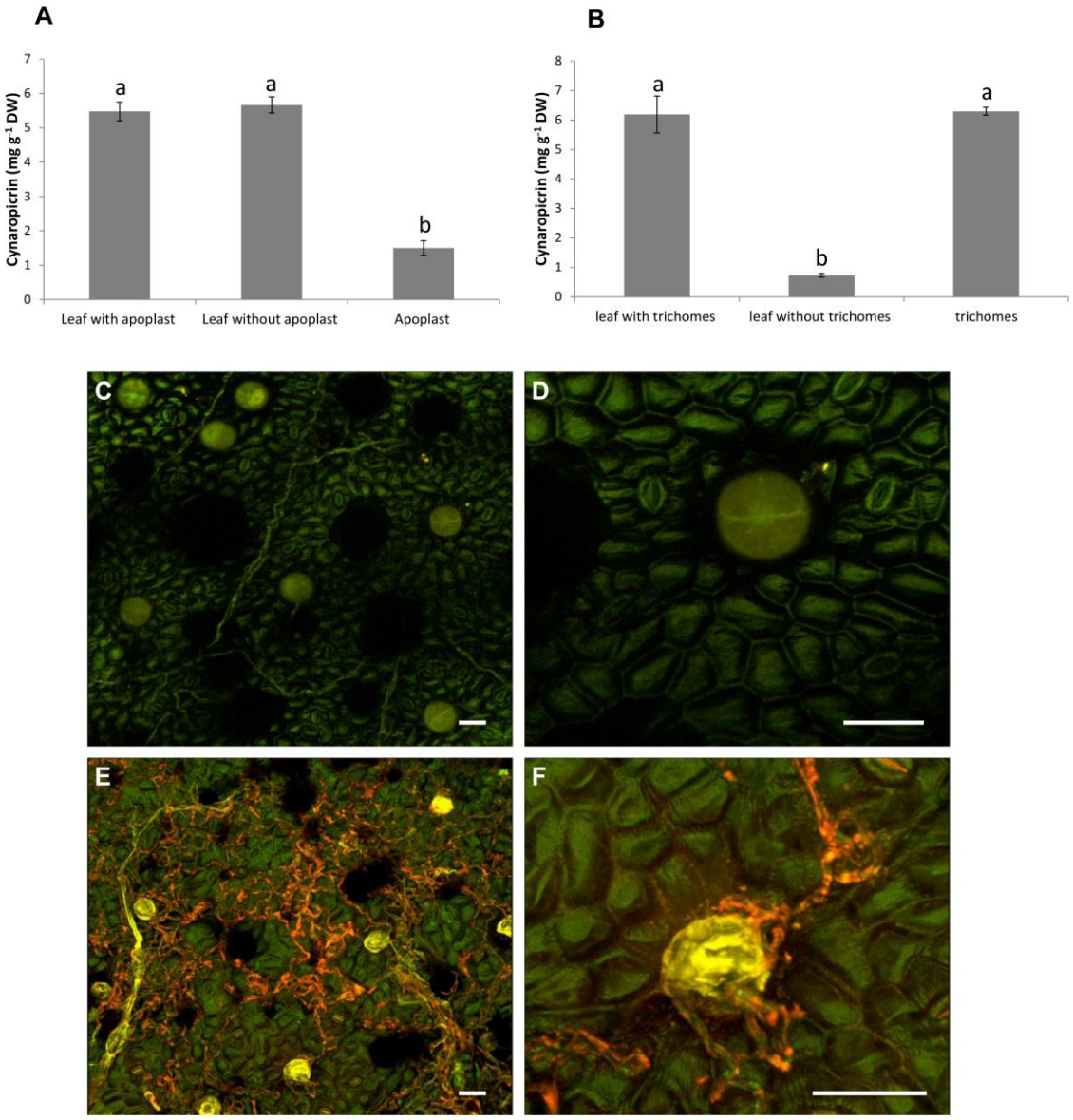

A
Cynaropicrin (mg g-1 DW)
a
a
b
Leaf with apoplast
Leaf without apoplast
Apoplast
B
Cynaropicrin (mg g-1 DW)
a
b
a
leaf with trichomes
leaf without trichomes
trichomes
C
D
E
F

Figure 4

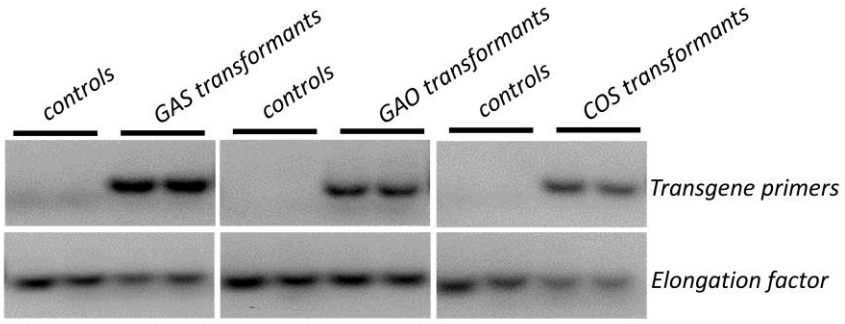

controls
GAS transformants
controls
GAO transformants
controls
COS transformants
Transgene primers
Elongation factor

Figure 5

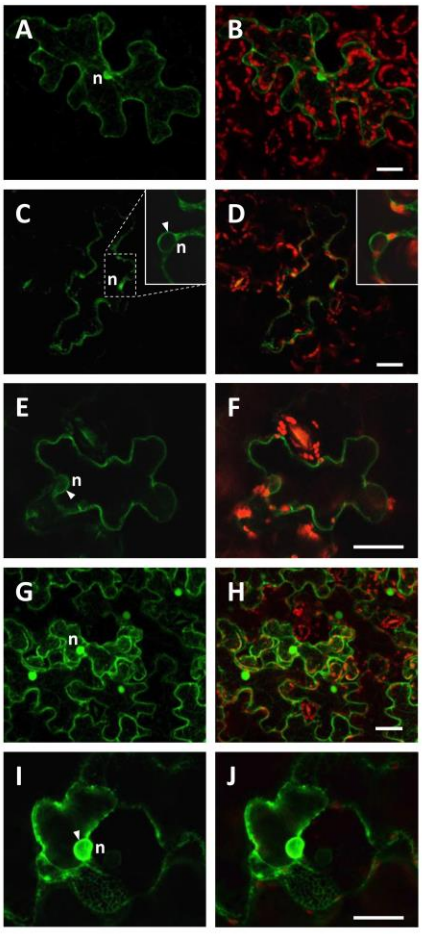
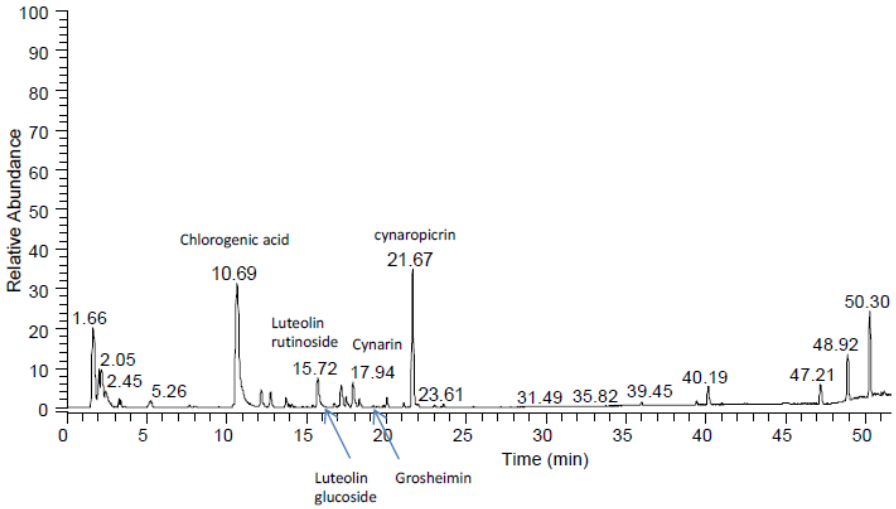


Figure S1


Relative Abundance
Time (min)
Chlorogenic acid
10.69
cynaropicrin
21.67
1.66
2.05
2.45
5.26
Luteolin rutinoside
15.72
Cynarin
17.94
23.61
31.49
35.82
39.45
40.19
47.21
48.92
50.30
Luteolin glucoside
Grosheimin