Tissue specific expression and genomic organization of bitter sesquiterpene lactone biosynthesis in Cichorium intybus L. (Asteraceae)

Bogdanovic, M., Cankar, K., Todorovic, S., Dragicevic, M., Simonovic, A., van Houwelingen, A. M. M. L., ... Beekwilder, M. J.

Tissue specific expression and genomic organization of bitter sesquiterpene lactone biosynthesis in *Cichorium intybus L.* (Asteraceae)

Authors:

Milica Bogdanović2*, Katarina Cankar1*, Slađana Todorović2, Milan Dragicević2, Ana Simonović2, Adele van Houwelingen1, Elio Schijlen1, Bert Schipper1, David Gagneul3$, Theo Hendriks3$$, Marie-Christine Quillet3$$, Harro Bouwmeester4, Dirk Bosch1, Jules Beekwilder1, #

* : authors contributed equally to this work

#: author for correspondence: Jules.beekwilder@wur.nl

Affiliations

1 Wageningen Plant Research, Wageningen UR, P.O. Box 16, 6700 AA Wageningen, The Netherlands

2 Institute for Biological Research Siniša Stanković, Department for Plant Physiology, University of Belgrade, Bulevar Despota Stefana 142, 11060 Belgrade, Serbia

3 Univ Lille Nord de France, Stress Abiotiques et Différenciation des Végétaux Cultivés (SADV). UMR INRA-USTL 1281, Bât. SN2, F-59655 Villeneuve d'Ascq, France

4 Laboratory of Plant Physiology, Wageningen University, 6708PD Wageningen, The Netherlands

$ Present address : EA 7394, USC INRA 1411, Institut Charles Viollette (ICV), Agro-food and biotechnology research institute, Université de Lille, INRA, ISA, Univ. Artois, Univ. Littoral Côte d'Opale, Cité Scientifique, 59655 Villeneuve d'Ascq, France

$$ present address: University of Lille, Unité Eco-Evo-Paléo, UMR UDL-CNRS. Bât. SN2, F-59655 Villeneuve d'Ascq, France

Abstract

Chicory (*Cichorium intybus* L.) produces bitter sesquiterpene lactones (STLs). Some enzymes in the biosynthetic pathway towards these compounds have been characterized. However, the genomic organization and tissue specificity of their biosynthesis is largely unknown. Concentrations of two sesquiterpene lactones and expression of genes involved in the first dedicated biosynthetic step were measured in different chicory tissues. BAC clones containing different genes encoding germacrene A synthase were sequenced, and revealed several tightly linked paralogs. Promoters of genes encoding two germacrene A synthases were fused to GFP and expressed in plants regenerated from transformed chicory hairy root cultures. Highest expression was observed in the epidermis of leaves and external root tissue. This work opens the possibility to select for chicory germplasm diversified in STL content, and to study their role in chicory in defence and physiology.

Keywords

Cichorium intybus; sesquiterpene lactone; genomic organisation; gene expression; biosynthesis

Introduction

Chicory (*Cichorium intybus* L.) is a perennial plant from the *Asteraceae* family. Wild chicory, with its characteristic blue flowers growing in roadsides in most European countries, has its origin in the Mediterranean basin, but has become common in temperate regions world-wide. Characteristically, it forms a strong taproot, which allows the plant to persist during periods of drought and temperature stress (Cranston et al., 2016).

C. intybus is cultivated for numerous different applications, and can be divided into varieties or cultigroups according to their use (Barcaccia et al., 2016; Cadalen et al., 2010). It is cultivated as a vegetable in the region around Belgium, the north of France and the Netherlands, to produce chicons. A chicon is an etiolated compact leaf structure that is consumed as white "witlof" or red "Radicchio" chicons. The related species *C. endivia* is consumed as a green leafy vegetable. The common feature of these vegetable forms is that their leaves are slightly bitter. Another variety, *C. intybus* v. *sativa*, is produced for industrial applications. Its taproot, which can amount to over a kilo of biomass per plant, is used for extracting inulin. Inulin is a fructose polymer, used as food fiber, but also as a low-calorie sweetener. Industrial chicory needs only very low input of crop protection agents. One of the reasons for its robust growth is the presence of bitter

compounds in leaves and roots, which belong to the class of sesquiterpene lactones (STLs).

STLs form a class of compounds that predominantly occur in Asteraceae species. STLs have a variety of bioactivities, ranging from allelopathic activity (Molinaro et al., 2016) to protective activity against herbivorous insects in roots (Huber et al., 2016) and flowers (Prasifka et al., 2015). In chicory, STLs provide bitterness, which gives flavour to the vegetables, and their presence in roots has been deployed to convert chicory root into a coffee substitute (Street et al., 2013). Also, STLs have a number of health associated properties, such as antimicrobial activity, and are used as chemotherapeutic agents (Ghantous et al., 2010; Popovic et al., 2015). The presence of STLs in chicory is associated with latex, which is exuded from both leaves and roots upon tissue damage (Sessa et al., 2000). The major STLs of chicory belong to the class of guaianolide sesquiterpene lactones and commonly derive from a single sesquiterpene, germacrene A (Fig. 1) (de Kraker et al., 1998; de Kraker et al., 2001). They are diversified by a large set of modifications, including oxidations, lactone ring closures and conjugations to oxalate, hydroxyphenylacetate and/or glycosyl moieties. These modifications diversify their biological properties. In chicory, the most predominant STLs are lactucin, lactucopicrin and 8-deoxylactucin, including their oxalates and glycosides (Fig. 1) (Sessa et al., 2000). The genetic control of STL biosynthesis in chicory is yet poorly understood at the genetic level. For the first dedicated biochemical step, two cDNAs encoding germacrene A synthases (GAS) converting farnesyl diphosphate (FPP) to germacrene A, hereafter named *CiGAS-short* and *CiGAS-long*, have been isolated (Bouwmeester et al., 2002). More recently, also cytochrome P450 enzymes that are able to oxidize germacrene A were described. Two germacrene A oxidases (GAO) were functionally characterised and shown to be capable to convert germacrene A to its acid, hereafter named CYP71AV4 and CYP71AV8 (Cankar et al., 2011; Nguyen et al., 2010). The costunolide synthase (COS, CYP71BL3), which also belongs to the CYP71 family of cytochrome P450 enzymes, was isolated and shown to convert germacrene A acid further to costunolide (Ikezawa et al., 2011; Liu et al., 2011).

Implementation of this knowledge into development of breeding tools has been hampered by the complex genetic structure of chicory: *C. intybus* is a self-incompatible species, and its genome is highly heterozygous (Zavada et al., 2017). Moreover, from other species it is known that relevant biosynthetic genes (both terpene synthase- and cytochrome P450 enzymes) occur as members of large gene families (Kulheim et al., 2015).

For the development of genotypes that differ in bitter compound content and composition, it would be of interest to better understand the genomic organisation of

terpene synthase genes involved in STL biosynthesis in chicory. In this work we analysed sequences and genomic organisation of genes in *C. intybus* that are relevant for STL biosynthesis. We used promoter sequences of these genes to study their expression in different tissues.

Materials and methods

qPCR

Chicory plants of industrial root chicory variety Orchies (Florimond-Deprez), which was chosen because it is frequently used in the field for inulin production, were grown in the greenhouse at 20 °C and 16 h daylight conditions for a period of three months. Leaves were collected and immediately frozen in liquid nitrogen. The root tissues were first separated into root epidermis, cortex and pith, and then snap frozen in liquid nitrogen. For seedlings, Orchies seeds were germinated on a filter paper at room temperature. Five days after germination, seedlings were pooled in three groups and frozen in liquid nitrogen. RNA was isolated from three biological replicates for each tissue using the protocol described by Chang et al. (Chang et al., 1993). One microgram of total RNA was used for cDNA synthesis using the iScript cDNA Synthesis Kit (Bio-Rad) after the treatment with DNAseI enzyme (Invitrogen). Primers used for the amplification of the GAS-short (NCBI accession number: AF498000) and GAS-long (AF497999) are given in the supplemental table 1. The transcript level of ribosomal protein L19 was used as endogenous reference gene (van Arkel et al., 2012). Quantitative PCR reactions were carried out in a total volume of 20 µL containing 10 µL of iQ SYBR Green Supermix (Bio-Rad), 0.3 µM of forward and reverse primer and cDNA corresponding to 20 ng RNA in a MyiQ real-Time PCR instrument (Bio-Rad). The following PCR program was used: 95 °C for 3 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. No unspecific amplification was observed by the melting curve analysis. Relative gene expression was calculated as: $2^{-\Delta Ct}$, where ΔCt = Ct (target gene) - Ct (reference gene)(Livak and Schmittgen, 2001). Data were subjected to statistical analysis using SPSS software (version 23 for Windows, IBM) for the analysis of variance (ANOVA). Duncan's multiple range test was used to analyse differences between treatments.

Metabolite analysis

Aliquot of 250 mg of chicory tissue material in three biological replicates (same materials as were used for the qPCR analysis) were extracted with 750 µl of 99.9% MeOH and 0.13% formic acid (v/v). Samples were vigorously vortexed and sonicated for 15 minutes. The samples were centrifuged for 15 min at 13.000 rpm in a table top centrifuge and the clear supernatant was transferred to a fresh tube and used for LC-MS analysis. A LC-LTQ-Orbitrap FTMS system (Thermo Scientific) consisting of an Acquity HPLC (Waters) connected to an LTQ/Orbitrap hybrid mass spectrometer (Thermo Fisher Scientific) equipped with an ESI source was used. Chromatographic separation took place on an analytical column (Luna 3µ C18/2 100A; 2.0 × 150 mm; Phenomenex, USA). 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. A

linear gradient from 5 to 75% acetonitrile (v/v) in 45 min was applied, which was followed by 15 min of washing and equilibration. FTMS full scans (m/z 95.00–1300.00) were recorded with a resolution of 60000. Injection volume was 5 µl. The quantification of lactucin and lactucopicrin was performed using an external standard curve prepared from authentic standards of these compounds purchased from Extrasynthese. Standards for other STLs were not commercially available. Data were subjected to statistical analysis using SPSS software (version 23 for Windows, IBM) for the analysis of variance (ANOVA). Duncan's multiple range test was used to analyze differences between treatments.

PCR-SSCP analysis of GAS-short and GAS long genes

Establishment of the mapping populations for the construction of a consensus genetic map for chicory and the method of SSCP analysis are described by (Cadalen et al., 2010). For CiGAS-short SSCP detection was performed after vertical gel electrophoresis rather than capillary electrophoresis. PCR was performed in a total volume of 20 µl containing 0.2 µM of each forward and reverse primer (primer sequences are given in supplemental table 1; primers synthesized by Integrated DNA Technologies BVBA, Leuven (B)), 1 mM $MgCl_2$, 0.1 mM dNTPs, 1 ng genomic DNA, 0.6 unit Taq DNA polymerase (Appligene, 15 units/µl), and 1x PCR buffer. The reaction mixture was incubated in a thermocycler (PE 9600, Perkin Elmer) with 1 min denaturation at 94°C, 1 min annealing at 55°C and 1 min elongation at 72°C during 40 cycles. After thermal cycling, the PCR products were denatured at 94°C for 5 min in presence of formamide. For CiGAS-short, one-tenth (v/v) of the PCR reactions were applied onto a non-denaturing polyacrylamide gel (0.5x MDE gel; TEBU), 20 cm in length and 0.75 mm thick, and electrophoresis was performed at 10°C in a Protean IIxi (Biorad) apparatus with 0.6x TBE (Tris-borate-EDTA) running buffer for 3000 Vh. After the run, the bands were visualized by silver staining (Bassam et al., 1991).

BAC selection

To obtain genomic gene sequences for *CiGAS-long* (AF797999), *CiGAS-short1* (AF498000) and *CiGAS-short2* (EH705708), one of two BAC- libraries (Gonthier et al., 2010) was screened. Two filters of the chicory CinS1S4 BAC library with in total 55,296 unique clones in duplo (representing 68% of the complete library, corresponding to 4.4x haploid genome equivalents) were screened with two sets of probes. One set contained the probes for *CiGAS-long* and *CiGAS-short2*; the second set contained the probes for *CiGAS-short1*. Each of the gene probes were obtained by PCR using specific primer pairs on genomic DNA of the K59 genotype (Supplemental table 1). Screening and validation were performed as described previously (Gonthier et al., 2010) at CNRGV in Toulouse, France. The designed primer pairs corresponded to cDNA sequences that were supposed

to flank introns as defined by sequence comparisons with genomic Arabidopsis gene sequences (for details see Cadalen et al., 2010).

Chicory BAC sequencing

BAC DNA was prepared using the QIAGEN large construct kit. The isolated DNA was subsequently analysed on a Roche GS FLX Sequencing device, using a long run. Overlapping sequences were assembled into 'contigs', using Newbler software (Roche). This resulted in 10-25 contigs per BAC.

The sequence assembly for BACs 59A14 and 83A09 resulted in several contigs after the 454 sequencing, therefore, additional Sanger sequencing of these BAC clones was performed using the DYEnamic ET Terminator Cycle Sequencing Kit (GE Healthcare, UK). The additional sequences were used to bridge the gaps between the contigs. The primer list is given in supplemental table 1. Sequences of the BACs were submitted to NCBI Genbank under accession numbers MH350853-MH350858.

Promoter characterization

Promoter regions of *CiGAS-L1b* (1738 bp), *CiGAS-S1* (1393 bp) and *CiGAS-S2* (1948 bp) were amplified from corresponding BACs using Advantage PCR kit (Clontech, USA) according to manufacturer's protocol. Primers used for amplification are given in supplementary table 1. Amplified fragments were first cloned into pDONR207 vector (Invitrogen), and recombinant colonies were selected on LB plates supplemented with gentamicin. The presence of the insert was confirmed by colony PCR. The cloned fragment was analysed by sequencing using DYEnamic ET Terminator Cycle Sequencing Kit (GE Healthcare, UK) and primers given in supplementary table 1. Selected clones were further transferred into pKGW-GGRR-C destination vector for promoter analysis in plants (Op den Camp et al., 2011). The vector contains a streptomycin/spectinomycin bacterial resistance gene, a kanamycin plant resistance gene, DsRED as a fluorescent marker for transformation under the control of AtUBQ10 promoter and eGFP-GUS fusion driven by the cloned promoter. The insertion of the promoter fragments in the pKGW-GGRR-C vector was confirmed by colony PCR. The plasmids obtained by this procedure were named: pKGW-PCiGAS-L1, pKGW-PCiGAS-S1 and pKGW-PCiGAS-S2. A control pKGW-GGRR-C plasmid was included.

Chicory transformation and regeneration of transformed plants

For all *in vitro* experiments, chicory (*C. intybus* L. Blue) was used (Samen Mauser, Winterthur, Switzerland). Seeds were surface sterilized with 20% (v/v) commercial bleach (0.8% active chlorine w/v) for 10 min, followed by five rinses with sterile distilled water. Seeds were germinated on Murashige and Skoog (MS) medium without sucrose

solidified with 0.6% agar (w/v). Seedlings and mature plants were grown in a climate room on solid MS medium containing 2% sucrose (w/v) under long-day conditions (16/8 light/dark), at 25 ± 2 °C, 60-70% relative humidity and photon flux rate of 42 µmol m^{-2} s^{-1} at the cultures level. *Rhizobium rhizogenes* A4M70GUS, obtained from Plant Physiology Department of IBISS, contained pRiA4 plasmid with integrated GUS cassette in TL region (Tepfer and Casse-Delbart, 1987). *R. rhizogenes* A4M70GUS strains carrying promoter-eGFP constructs were incubated overnight in liquid YEB medium with neomycin 100 mg·l^{-1} and spectinomycin 100 mg·L^{-1}. Leaves of 5-week old chicory plants were cut and placed in a petri dish containing solid MS medium, and inoculated along the leaf veins with a sterile needle dipped in bacterial culture. Roots emerging from leaves were excised and grown separately as clones on solid MS medium containing 500 mg·l^{-1} cefotaxime. Roots were subcultured every month and cefotaxime concentration was gradually reduced over the period of 6 months. Regenerated shoots forming spontaneously on root cultures were excised and grown separately. One month old plants were used for genomic DNA isolation to check their transgenic nature. Genomic DNA of transformants was extracted from leaves using a mini-prep CTAB method (Haymes, 1996) and treated with RNase A (Fermentas, USA) using manufacturer's protocol. PCR was performed using Taq recombinant polymerase (Fermentas, USA). The PCR mixtures consisted of 100 ng of genomic DNA, 1 µM specific primers (supplemental table 1) and standard components according to Fermentas protocol, in a 25-µl volume. The Primers used were DsRED_F and R, rolA1_F and R to establish transformation of the pKGW-GGRR-C plasmid and A4GUS plasmid, and virD1_F and R to exclude bacterial contamination. Selected transgenic plants were grown for 10 weeks before being used for experiments.

Confocal microscopy

Confocal microscopy was used for fluorescence analysis of chicory tissues. Thin hairy roots and leaves were placed on slides and kept moist with half strength liquid MS medium. Leica TCS SPE microscope (Leica Microsystems, Wetzlar, Germany) was used with sequential scanning for DsRED and GFP channel. DsRED was excited with 532 nm laser line and the signal was recorded in a 570-650 nm range, while eGFP was excited with a 488 nm laser line and recorded in a 500-530 nm range.

Results

Sesquiterpene lactone metabolites and gene expression

Several tissues from chicory plants were analyzed for content of STLs lactucin and lactucopicrin (Fig. 2A). Root epidermis and leaf tissue were found to have the highest levels of the STLs, while seedlings and inner root tissues (root cortex and root pith) showed lower concentrations. Subsequently, the expression of early biosynthetic genes for STLs was tested by quantitative RT PCR. Relative gene expression of the two germacrene A synthases (*CiGAS-short*, *CiGAS-long*) in different tissues is shown in fig 2B.

In the root tissues, the amount of STLs largely paralleled the gene expression profiles of *CiGAS-long*, and *CiGAS-short*. Differences in gene expression are less significant compared to STL concentrations, due to high variation between biological replicates. As a trend, gene expression and STL accumulation is highest in the outer tissues. Interestingly, the root cortex, which represents a collection of root cell types including laticifers, but also cells that mediate production and storage of high amounts of fructose polymers (inulin), displays an intermediate expression level of STL biosynthetic genes, not significantly different from other root tissues. In leaves, concentrations of STLs were relatively high, in spite of very low *CiGAS-short* expression. In seedlings, on the other hand, STL levels were relatively low, while expression of STL biosynthetic genes was high, suggesting that the STLs have not yet accumulated in the seedlings. In lettuce, the increase of expression of STL biosynthetic genes before STL accumulation in seedlings has also been reported (Bestwick et al., 1995).

The *CiGAS-short* gene is poorly expressed in leaf, while the *CiGAS-long* gene, which encodes an enzyme with the same activity, is expressed in both leaf and root tissues, which is in agreement with previous results (Bouwmeester et al., 2002).

Genomic organization of STL genes

Genetic loci corresponding to cDNAs of *CiGAS-long* and *CiGAS-short* were previously mapped in two mapping populations to construct the reference chicory genetic map (Cadalen et al., 2010). For *CiGAS-short*, a single locus was mapped to linkage group 3 (LG3) in the 'Rubis 118' F2 mapping population. However, a single-strand conformation polymorphism (SSCP) analysis of the 96 individuals of this population revealed multi-banded profiles, with 5 bands for each of the parental genotypes and 9 for the hybrid genotype (Fig. 3). The segregation of the three profiles, two corresponding to parental genotypes and one to their heterozygous hybrid, and the absence of recombination

between these profiles, suggested the presence of at least 3 tightly linked *CiGAS-short*-like genes at this locus, with each allele responsible for 1 or 2 bands in the SSCP profiles.

In a previous analysis, markers based on *CiGAS-long* suggested two loci for this gene on LG9 (Cadalen et al., 2010). However, a more extensive marker analysis resolved this to a single locus for *CiGAS-long* on LG9. The revised LG9 map is shown in Supplemental Figure 1. In contrast to *CiGAS-short*, only simple SSCP profiles for *CiGAS-long* were obtained (not shown), suggesting that this locus contains a single copy of the gene.

To zoom in on genomic sequences representing sesquiterpene biosynthetic genes, the chicory CinS2S4 BAC libary (Gonthier et al., 2010) was screened using PCR probes generated on genomic DNA for *CiGAS-long* and *CiGAS-short* genes. In total 40 BACs hybridizing to one of the probes were identified, and 11 of those were selected for sequencing, using 454 sequencing technology (Table 1). In the assembled sequences, each BAC was covered by a set of contigs, separated by gaps of unidentified sequences. PCR amplification and Sanger sequencing was performed to fill these gaps. Subsequently, the positions of open reading frames (ORFs) encoding the STL genes and their intron-exon structures were established using the cDNA sequences of *CiGAS-long* and *CiGAS-short* as queries.

Three sequences containing ORFs corresponding to the *CiGAS-short* gene (termed *CiGAS-S1, S2* and *S3*) were found to map on BAC 83A09, and overlapping BACs 36D10, 94D20 and 29O10 (Table 1). In addition, two physically unlinked pseudogenes resembling *CiGAS-short* were found (*CiGAS-S4a* and *CiGAS-S4b*) on BACs 73J10 and 05O04 respectively. For the *CiGAS-long* gene, two gene copies were found, *CiGAS-L1a* on BAC 05I22 and *CiGAS-L1b* on BAC 105O22.

Coding capacity of terpene biosynthetic genes

CiGAS genes appear to have conserved a 7-exon/6-intron gene structure (Figure 4). *CiGAS-S1* and *CiGAS-S2* encode full-length proteins. For *CiGAS-S3*, the first two exons map outside the analyzed BACs, but otherwise it encodes an uninterrupted protein. *CiGAS-S4a* carries a premature stop codon in exon 1. Both *CiGAS-S4a* and *CiGAS-S4b* contain a 4 bp insertion in exon 3, leading to a frameshift and premature termination of the protein. In addition, *CiGAS-S4b* is interrupted by a 4862 bp insertion in exon 3, related to the TNT-194 transposon from *Nicotiana tabacum*. Both *CiGAS-L1a* and *CiGAS-L1b* appear to represent uninterrupted ORFs, though exon7 from *CiGAS-L1b* mapped outside the analyzed BACs.

Sequence comparison of terpene biosynthetic genes

The homology of the identified genes was further investigated by sequence comparison. First, alignments were made for open reading frames and untranslated regions (introns, promoter segments and terminator). Data on the percentage of identity found in the sequence comparisons can be found in Supplemental figure 2.

GAS-short genes *CiGAS-S1* and *CiGAS-S2*, which encode full-length proteins, share between 80% and 100% sequence identity in the ORF (98%), the intron sequences and the region of 350 bp upstream of the start-codon. *CiGAS-S3*, *CiGAS-S4a* and *CiGAS-S4b* have a lower identity to *CiGAS-S1* (86-91% identity in the ORF). For *CiGAS-S4a* and *CiGAS-S4b*, sequences share 98% identity. In the *CiGAS-S4b* promoter sequence, a large deletion of 7.5 kb is observed, relative to *CiGAS-S4a*. Apart from this deletion, the regions in which *CiGAS-S4a* and *CiGAS-S4b* are embedded can be readily aligned, suggesting that these genes are alleles of a pseudogene.

The identity between *CiGAS-L1a* and *CiGAS-L1b* is very high in ORF and non-coding regions (82-100%), except for an insertion of a 4260 bp retrotransposon of the Ty1/Copia family in *CiGAS-L1a*, located 700 bp upstream of its start-codon. Given the high co-linearity and sequence identity of *CiGAS-L1a* and *CiGAS-L1b*, it can be concluded that they encode alleles of the same gene.

Indeed the genetic data obtained previously also suggest the presence of two alleles for CiGAS-L1. The plant used for the creation of the BAC library is a descendant of the K28K59 cross, in which the *CiGAS-long* gene was mapped, indicating heterozygosity. In contrast, for *CiGAS-short* no polymorphism was detected in the K28K59 population, indicating that the parents, and thus their progeny was homozygous for the *CiGAS-S1-3*. An exception is formed by *CiGAS-S4*, the pseudogene; the primers used for GAS-short mapping would not amplify the *CiGAS-S4* sequence.

Thus, in the chicory genome, a single *CiGAS-long* gene has been observed, with at least two alleles (*CiGAS-L1a* and *CiGAS-L1b*). At least four *CiGAS-short* paralogs were observed to exist (*CiGAS-S1,2,3* and *4*), of which at least two encode full-length proteins (*CiGAS-S1* and *CiGAS-S2*) and one is a pseudogene (*CiGAS-S4*). From the pseudogene, at least two alleles seem to be observed (*CiGAS-S4a* and *CiGAS-S4b*).

The genomic regions flanking the *CiGAS-long* and *CiGAS-short* genes are rich in fragmented retrotransposon elements. Therefore synteny of these genes with other coding genes is difficult to address.

Promoter analysis of Germacrene A synthase and CYP71AV8 genes

Promoter fragments for *CiGAS-S1* (1393 bp), *CiGAS-S2* (1948 bp) and *CiGAS-L1b* (1738 bp) were amplified from the BACs. These promoter fragments were cloned in vector pKGW-GGRR-C (Gavrin et al., 2014), in such a way that they could drive expression of green fluorescent protein (GFP). Constructs were named $P_{CiGAS\text{-}L1b}$, $P_{CiGAS\text{-}S1}$ and $P_{CiGAS\text{-}S2}$ and were introduced in *R. rhizogenes*.

In order to localize expression of the STL biosynthetic genes, the *R. rhizogenes* strains hosting $P_{CiGAS\text{-}L1b}$, $P_{CiGAS\text{-}S1}$ and $P_{CiGAS\text{-}S2}$ were introduced into chicory, thereby producing hairy root cultures. Strains containing either a vector expressing DsRED or GFP were used as controls in transformation. Roots emerging on inoculated leaves were grown separately as independent clones without growth regulators and displayed fast growth and extensive branching (supplemental Figure 4). Subcultured roots spontaneously formed shoots, which, when excised and grown separately, regenerated whole plants. The presence of promoter constructs was confirmed by PCR (Supplemental table 2).

Since each construct contains a DsRED gene driven by the constitutive *UBI10* promoter, DsRED fluorescence was used to evaluate the presence of ectopic constructs in roots (Supplemental Fig. 3). DsRED fluorescence was stable in most clones over the course of several months and a good indication of stable transgene expression in transformed lines. Only for $P_{CiGAS\text{-}S2}$, no lines that displayed consistent DsRED fluorescence was obtained.

To observe promoter activity in different cell types in roots and leaves, GFP expression was monitored in *in vitro* plants carrying the $P_{CiGAS\text{-}L1b}$ and $P_{CiGAS\text{-}S1}$ promoter constructs by fluorescence confocal microscopy. In roots, all three promoters were able to drive GFP fluorescence. GFP expression was not localized specifically to vascular tissue, but was observed in all cells, and was most pronounced in epidermal layer. In leaf tissue, only $P_{CiGAS\text{-}L1b}$ plants displayed a strong GFP-fluorescence, which corresponds to the RT-PCR results. Again, fluorescence was not localized to specific cell types, but could be observed in mesophyll cells. The most prominent fluorescence was observed in the epidermal layer (Figure 5).

Discussion

In this work, we address the genetic organization of genes that perform the first dedicated steps in the synthesis of STLs in chicory. Analysis of genetic maps, SSCP data and BAC sequencing lead to the identification of a number of paralogous genes and their alleles. The paralogues are expressed in different tissues, as appears from a promoter-

GFP analysis in hairy roots. These data form an important step in the understanding of the regulation of STL biosynthesis in Asteraceae, and provide leads for understanding their evolution.

Genomic organisation of STL biosynthesis in chicory

The biosynthesis of STLs is complex, and involves many steps. Previously two different cDNAs were identified, which both encode germacrene A synthases mediating the first dedicated step in STL biosynthesis. Our work is focussed on the genomic organization of these two genes. The *CiGAS-long* transcript seems to be encoded by a single gene *CiGAS-L1*, which is now mapped on one locus on LG9 (supplemental fig 1). Two alleles (*CiGAS-L1a* and *b*) were observed for this gene. *CiGAS-L1* seems to control STL biosynthesis in leaf, while the *CiGAS-short* gene is hardly expressed in this tissue. *CiGAS-L1* also contributes strongly to STL biosynthesis in root, along with *CiGAS-short* genes. GFP-promoter studies with the *CiGAS-L* promoter are in agreement with the qRT-PCR data.

Also *CiGAS-short* was mapped on a single locus, on LG3. However, this locus appears to be more complex, with at least 3 gene copies, according to the SSCP analysis. BAC sequencing showed that indeed at least three copies of *CiGAS-short* are physically linked. Of these, *CiGAS-S1* and *CiGAS-S2* seem to be the result of a recent duplication. This duplication event also includes the core promoter. The more distal parts of the promoter and the terminator vary more, suggesting that subtle differences in regulation may exist. *CiGAS-S3* and *S4* seem to have resulted from earlier duplications from *CiGAS-S1*. *CiGAS-S4* does not encode a functional protein, and there is no visible selection pressure for intact coding: the region overlapping this gene is prone to rearrangements by shifts in the reading frame, stop codons, small indels, large deletions and transposon insertions. Based on the BAC sequencing we cannot yet make a physical link between *CiGAS-S4* and the other *CiGAS-S*s.

Recently, the genome of *Lactuca sativa* (lettuce) was published (Reyes-Chin-Wo et al., 2017). *Lactuca* is highly related to chicory, both belonging to the subfamily Cichorioideae of the Asteraceae and produces STLs such as lactucin. When compared to the *Lactuca* genome, *CiGAS-S* and *CiGAS-L* are mapping on the corresponding linkage groups, as shown in Supplemental Table 3. *Lactuca* has a single copy of a *CiGAS-L*-like gene on chromosome 2. Two orthologues of *CiGAS-S1-4* can be found on the *Lactuca* chromosome 8.

We have been investigating the conservation of the local synteny for the GAS genes between *Lactuca* and chicory. On the chicory BAC sequences, we observe that the genomic regions containing the *CiGAS-S* and *CiGAS-L* sequences are both rich in

fragments of transposable elements. Therefore, conservation of the gene synteny between *Lactuca* and Chicory could not be observed. Two physically linked copies of a homologue of CiGAS-S1 are present in lettuce, while chicory has at least three physically linked *CiGAS-S* genes, and one additional pseudogene *CiGAS-S4*. This indicates that the region where *CiGAS-S* genes are located has undergone recent rearrangements in both species.

Localization of STL biosynthesis in chicory

Chicory contains branched anastomosed laticifers (Vertrees and Mahlberg, 1978), in which latex with high concentrations of STLs is stored (Sessa et al., 2000). Commonly, latex is proposed to function in defence against herbivory. In chicory roots, latex is mainly produced in the cortex (Vertrees and Mahlberg, 1978), although the concentrations of STL and the expression of biosynthetic genes are highest in the root external tissues, compared to the root cortex (Figure 2). Promoter-GFP analysis for both *CiGAS-L1* and *CiGAS-S1* seem to confirm this. These promoter-GFP analyses were performed in roots and leaves of plants regenerated from hairy root culture, which are known to produce STLs (Bogdanovic et al., 2014; Malarz et al., 2002). Still, in a taproot with a probably higher cell differentiation, the localization of GAS expression may be different from hairy roots. More cell-type specialization is expected in the taproot. Nonetheless, the promoter-GFP results from transformed plants seem to parallel the qRT-PCR results showing a high expression in the epidermal part of the taproot. Also in leaves, latex extrusion can be observed, and also in leaves most intensive fluorescence was observed in epidermal tissues. It remains unresolved what drives the accumulation of STLs in the latex. It is clear that more detailed studies are needed to resolve the control of accumulation of STLs in the latex of chicory.

Conclusion

In this study we identified genomic regions of chicory involved in STL biosynthesis. This work allows to address the role of STLs in chicory in defence and physiology of the plant. Moreover, the availability of promoter sequences should provide tools to drive localized expression of genes in tissues that have high expression of the terpene pathway. This could enable the production of different terpenes in chicory tap root.

References

Barcaccia, G., Ghedina, A., Lucchin, M., 2016. Current Advances in Genomics and Breeding of Leaf Chicory (*Cichorium intybus* L.). Agriculture-Basel 6, 50. doi:10.3390/agriculture6040050

Bassam, B.J., Caetano-Anolles, G., Gresshoff, P.M., 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal Biochem 196, 80-83. doi:10.1016/0003-2697(91)90120-I

Bestwick, L., Bennett, M.H., Mansfield, J.W., Rossiter, J.T., 1995. Accumulation of the phytoalexin lettucenin a and changes in 3-hydroxy-3-methylglutaryl coenzyme a reductase activity in lettuce seedlings with the red spot disorder. Phytochemistry 39, 775-777. doi:10.1016/0031-9422(95)00089-P

Bogdanovic, M.D., Todorovic, S.I., Banjanac, T., Dragicevic, M.B., Verstappen, F.W.A., Bouwmeester, H.J., Simonovic, A.D., 2014. Production of guaianolides in *Agrobacterium rhizogenes* - transformed chicory regenerants flowering in vitro. Ind Crop Prod 60, 52-59. doi:10.1016/j.indcrop.2014.05.054

Bouwmeester, H.J., Kodde, J., Verstappen, F.W., Altug, I.G., de Kraker, J.W., Wallaart, T.E., 2002. Isolation and characterization of two germacrene A synthase cDNA clones from chicory. Plant Physiol 129, 134-144. doi:10.1104/pp.001024

Cadalen, T., Morchen, M., Blassiau, C., Clabaut, A., Scheer, I., Hilbert, J.L., Hendriks, T., Quillet, M.C., 2010. Development of SSR markers and construction of a consensus genetic map for chicory (*Cichorium intybus* L.). Molecular Breeding 25, 699-722. doi:10.1007/s11032-009-9369-5

Cankar, K., van Houwelingen, A., Bosch, D., Sonke, T., Bouwmeester, H., Beekwilder, J., 2011. A chicory cytochrome P450 mono-oxygenase CYP71AV8 for the oxidation of (+)-valencene. FEBS Lett 585, 178-182. doi:10.1016/j.febslet.2010.11.040

Chang, S., Puryear, J., Cairney, J., 1993. A Simple and Efficient Method for Isolating RNA from Pine Trees. Plant Mol Biol Report 11, 113-116. doi:10.1007/BF02670468

Cranston, L.M, Kenyon P.R., Morris, S.T., Lopez-Villalobos, N., Kemp, P.D. (2016) Morphological and physiological responses of plantain (*Plantago lanceolata*) and chicory (*Cichorium intybus*) to water stress and defoliation frequency. J Agro Crop Sci 202, 13–24. doi:10.1111/jac.12129

de Kraker, J.W., Franssen, M.C., Dalm, M.C., de Groot, A., Bouwmeester, H.J., 2001. Biosynthesis of germacrene A carboxylic acid in chicory roots. Demonstration of a cytochrome P450 (+)-germacrene a hydroxylase and NADP+-dependent sesquiterpenoid dehydrogenase(s) involved in sesquiterpene lactone biosynthesis. Plant Physiol 125, 1930-1940. doi:10.1104/pp.125.4.1930

de Kraker, J.W., Franssen, M.C., de Groot, A., Konig, W.A., Bouwmeester, H.J., 1998. (+)-Germacrene A biosynthesis. The committed step in the biosynthesis of bitter sesquiterpene lactones in chicory. Plant Physiol 117, 1381-1392. doi:10.1104/pp.117.4.1381

Gavrin, A., Kaiser, B.N., Geiger, D., Tyerman, S.D., Wen, Z., Bisseling, T., Fedorova, E.E., 2014. Adjustment of host cells for accommodation of symbiotic bacteria: vacuole defunctionalization, HOPS suppression, and TIP1g retargeting in Medicago. Plant Cell 26, 3809-3822. doi:10.1105/tpc.114.128736

Ghantous, A., Gali-Muhtasib, H., Vuorela, H., Saliba, N.A., Darwiche, N., 2010. What made sesquiterpene lactones reach cancer clinical trials? Drug Discov Today 15, 668-678. doi:10.1016/j.drudis.2010.06.002

Gonthier, L., Bellec, A., Blassiau, C., Prat, E., Helmstetter, N., Rambaud, C., Huss, B., Hendriks, T., Berges, H., Quillet, M.C., 2010. Construction and characterization of two BAC libraries representing a deep-coverage of the genome of chicory (*Cichorium intybus* L., Asteraceae). BMC Res Notes 3, 225. doi:10.1186/1756-0500-3-225

Haymes, K.M., 1996. Mini-prep method suitable for a plant breeding program. Plant Molecular Biology Reporter 14, 280-284. doi:10.1007/Bf02671664

Huber, M., Epping, J., Schulze Gronover, C., Fricke, J., Aziz, Z., Brillatz, T., Swyers, M., Kollner, T.G., Vogel, H., Hammerbacher, A., Triebwasser-Freese, D., Robert, C.A., Verhoeven, K., Preite, V., Gershenzon, J., Erb, M., 2016. A Latex Metabolite Benefits Plant Fitness under Root Herbivore Attack. PLoS Biol 14, e1002332. doi:10.1371/journal.pbio.1002332

Ikezawa, N., Gopfert, J.C., Nguyen, D.T., Kim, S.U., O'Maille, P.E., Spring, O., Ro, D.K., 2011. Lettuce costunolide synthase (CYP71BL2) and its homolog (CYP71BL1) from sunflower catalyze distinct regio- and stereoselective hydroxylations in sesquiterpene lactone metabolism. J Biol Chem 286, 21601-21611. doi:10.1074/jbc.M110.216804

Kulheim, C., Padovan, A., Hefer, C., Krause, S.T., Kollner, T.G., Myburg, A.A., Degenhardt, J., Foley, W.J., 2015. The Eucalyptus terpene synthase gene family. BMC Genomics 16, 450. doi:10.1186/s12864-015-1598-x

513 Liu, Q., Majdi, M., Cankar, K., Goedbloed, M., Charnikhova, T., Verstappen, F.W., de Vos,
514 R.C., Beekwilder, J., van der Krol, S., Bouwmeester, H.J., 2011. Reconstitution of the
515 costunolide biosynthetic pathway in yeast and *Nicotiana benthamiana*. PLoS One 6,
516 e23255. doi:10.1371/journal.pone.0023255

517 Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-
518 time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402-408.
519 doi:10.1006/meth.2001.1262

520 Malarz, J., Stojakowska, A., Kisiel, W., 2002. Sesquiterpene lactones in a hairy root
521 culture of *Cichorium intybus*. Z Naturforsch C 57, 994-997. doi:10.1515/znc-2002-11-
522 1207

523 Molinaro, F., Monterumici, C.M., Ferrero, A., Tabasso, S., Negre, M., 2016. Bioherbicidal
524 activity of a germacranolide sesquiterpene dilactone from *Ambrosia artemisiifolia* L. J
525 Environ Sci Health B 51, 847-852. doi:10.1080/03601234.2016.1208466

526 Nguyen, D.T., Gopfert, J.C., Ikezawa, N., Macnevin, G., Kathiresan, M., Conrad, J.,
527 Spring, O., Ro, D.K., 2010. Biochemical conservation and evolution of germacrene A
528 oxidase in asteraceae. J Biol Chem 285, 16588-16598. doi:10.1074/jbc.M110.111757

529 Op den Camp, R.H., De Mita, S., Lillo, A., Cao, Q., Limpens, E., Bisseling, T., Geurts, R.,
530 2011. A phylogenetic strategy based on a legume-specific whole genome duplication
531 yields symbiotic cytokinin type-A response regulators. Plant Physiol 157, 2013-2022.
532 doi:10.1104/pp.111.187526

533 Popovic, V., Stojkovic, D., Nikolic, M., Heyerick, A., Petrovic, S., Sokovic, M., Niketic, M.,
534 2015. Extracts of three *Laserpitium* L. species and their principal components laserpitine
535 and sesquiterpene lactones inhibit microbial growth and biofilm formation by oral
536 Candida isolates. Food Funct 6, 1205-1211. doi:10.1039/c5fo00066a

537 Prasifka, J.R., Spring, O., Conrad, J., Cook, L.W., Palmquist, D.E., Foley, M.E., 2015.
538 Sesquiterpene Lactone Composition of Wild and Cultivated Sunflowers and Biological
539 Activity against an Insect Pest. J Agric Food Chem 63, 4042-4049.
540 doi:10.1021/acs.jafc.5b00362

541 Reyes-Chin-Wo, S., Wang, Z.W., Yang, X.H., Kozik, A., Arikit, S., Song, C., Xia, L.F.,
542 Froenicke, L., Lavelle, D.O., Truco, M.J., Xia, R., Zhu, S.L., Xu, C.Y., Xu, H.Q., Xu, X.,
543 Cox, K., Korf, I., Meyers, B.C., Michelmore, R.W., 2017. Genome assembly with in vitro
544 proximity ligation data and whole-genome triplication in lettuce. Nature Communications
545 8. doi:10.1038/ncomms14953

546 Sessa, R.A., Bennett, M.H., Lewis, M.J., Mansfield, J.W., Beale, M.H., 2000. Metabolite
547 profiling of sesquiterpene lactones from *Lactuca* species. Major latex components are
548 novel oxalate and sulfate conjugates of lactucin and its derivatives. J Biol Chem 275,
549 26877-26884. doi:10.1074/jbc.M000244200

550 Street, R.A., Sidana, J., Prinsloo, G., 2013. *Cichorium intybus*: Traditional Uses,
551 Phytochemistry, Pharmacology, and Toxicology. Evid Based Complement Alternat Med
552 2013, 579319. doi:10.1155/2013/579319

553 Tepfer, M., Casse-Delbart, F., 1987. *Agrobacterium rhizogenes* as a vector for
554 transforming higher plants. Microbiol Sci 4, 24-28. doi:10.1385/0-89603-321-X:49

555 van Arkel, J., Vergauwen, R., Sevenier, R., Hakkert, J.C., van Laere, A., Bouwmeester,
556 H.J., Koops, A.J., van der Meer, I.M., 2012. Sink filling, inulin metabolizing enzymes and
557 carbohydrate status in field grown chicory (*Cichorium intybus* L.). J Plant Physiol 169,
558 1520-1529. doi:10.1016/j.jplph.2012.06.005

559 Vertrees, G.L., Mahlberg, P.G., 1978. Structure and Ontogeny of Laticifers in *Cichorium*
560 *intybus* (Compositae). Am J Bot 65, 764-771. doi:10.2307/2442152

561 Zavada, T., Malik, R.J., Kesseli, R.V., 2017. Population structure in chicory (*Cichorium*
562 *intybus*): A successful U.S. weed since the American revolutionary war. Ecol Evol 7,
563 4209-4219. doi:10.1002/ece3.2994

564

565

566

567 Author contributions

568 The study was conceived by MB, KC, HB, DB, TH and JB. MB, ST, MD and AS performed
569 *C. intybus* transformation, regeneration and testing of transgenes. KC, AvH, BS and JB
570 performed qPCR and metabolite analyses. DG, MCQ and TH performed BAC selection and
571 SSCP analysis. KC, MB, ES and JB performed genomics analysis. MB, KC, TH and JB
572 prepared the manuscript. All authors contributed to revising the manuscript. All authors
573 have read and approved the final manuscript.

574 Conflict of interest

575 The authors declare that the research was conducted in the absence of any commercial
576 or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

The authors thank the former SADV-CartoChic team members for their contributions to the genetic re-mapping of GAS-long locus on LG9: Thiery Cadalen, Isabelle Scheer, Monika Mörchen, Christelle Blassiau, Aline Clabaut and Meriem Bahri. The team of Hélène Bérges, CNRGV Toulouse, is acknowledged for screening the chicory BAC libraries, and providing the selected BAC clones used for further analyses. Marian Bemer is acknowledged for support with confocal microscopic analysis. Rene Geurts is acknowledged for help with setting up hairy root transformation.

This project was partly funded by the European Commission (EU FP7 project TerpMed, Plant Terpenoids for Human Health: a chemical and genomic approach to identify and produce bioactive compounds, Grant agreement no.: 227448).

Tables

Table 1: characteristics of BAC clones

BAC number	Probe hybridization	Gene found	BAC insert size
05I22	GAS-long	CiGAS-L1a	95 kb
105O22	GAS-long	CiGAS-L1b	140 kb
83A09	GAS-short	CiGAS-S1, GAS-S2, GAS-S3	125 kb
73J10	GAS-short	CiGAS-S4a	107 kb
94D20	GAS-short	CiGAS-S1	115 kb
36D10	GAS-short	CiGAS-S2	95 kb
29O10	GAS-short	CiGAS-S1	95 kb
05O04	GAS-short	CiGAS-S4b	85 kb

Figure legends

Figure 1: Biosynthetic pathway of guaianolide STLs in chicory. FPP: farnesyl diphosphate; GAS: germacrene A synthase; GAO: germacrene A oxidase, COS: costunolide synthase.

Figure 2: Gene expression and quantification of STLs in chicory tissues. A) Quantification of lactucin and lactucopicrin amount in chicory tissues. Different letters above bars represent statistically significant differences ($p < 0.05$) according to Duncan's test; B) gene expression of germacrene A synthase genes in chicory tissues normalised to the expression of the ribosomal protein L19. For both plots the mean ± SD for three biological replicates is shown. C) Chicory root producing latex. D) Chicory root cross section with indicated regions taken for analysis.

Figure 3: SSCP screening of CiGAS-Long. Silver-stained PAGE gel showing SSCP profiles for *CiGAS-CiGAS-S1* in 18 descendants of the Rubis 118 mapping population. On the right is a schematic representation of the profiles observed. Lanes indicated with a or b represent plants showing one of the two parental banding patterns (a = MS8, b = Cassel), while the a/b lanes represent heterozygous plants (= F1 'Rubis').

Figure 4: Schematic representation of the genomic gene structure for the GAS genes. Open box: exon; filled black box: intron; filled grey box: retrotransposon; arrow: promoter; //: BAC end; asterisk: premature stopcodon.

Figure 5. Confocal analysis of whole chicory root (A) and leaf (B), as seen on GFP channel (green), DsRED channel (red) and composite image. All images were acquired using the same parameters. Promoter clone constructs analyzed include pKGW-PCiGAS-L1 and pKGW-PCiGAS-S1, as well as an untransformed plant (WT) and a $DsRED^{+}$ clone.

Supplementary material

Supplemental table 1: Primers used in this study

Supplemental table 2. Summary of chicory transformation efficiency using *R. rhizogenes* A4M70GUS strains carrying promoter constructs or control vectors.

Supplemental table 3 Comparison *Lactuca sativa* gene models, compared to chicory proteins

Supplemental figure 1: Genetic map of LG9 of the K28K59 population

Supplemental figure 2: Percentage of identity matrices found in the sequence comparisons between germacrene A synthases for exons, introns, promoter regions (bp 0-350 and 351-1000) and terminator regions.

Supplemental figure 3. Chicory hairy root transformation and selection.

Figure 1

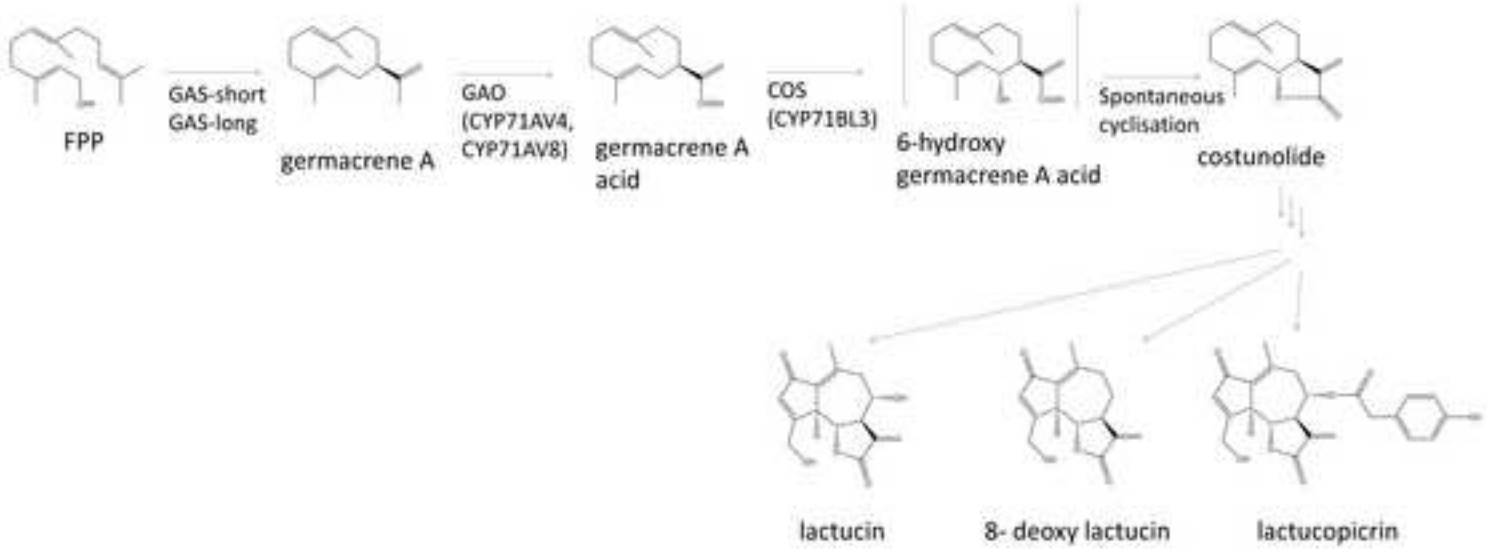
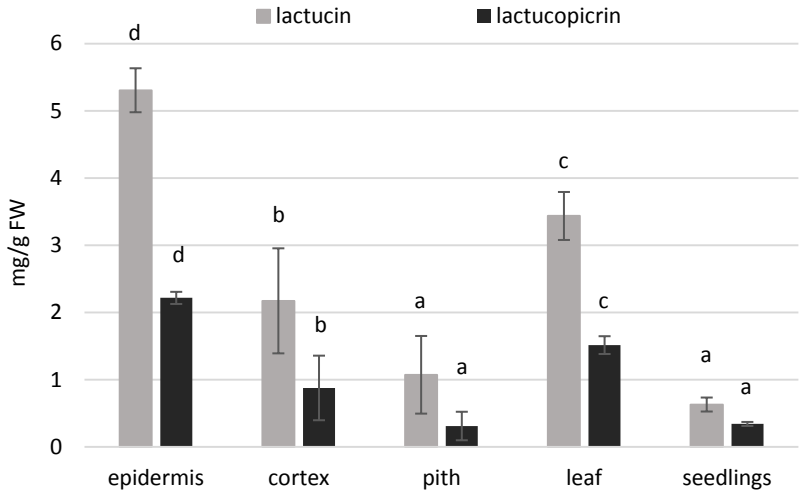
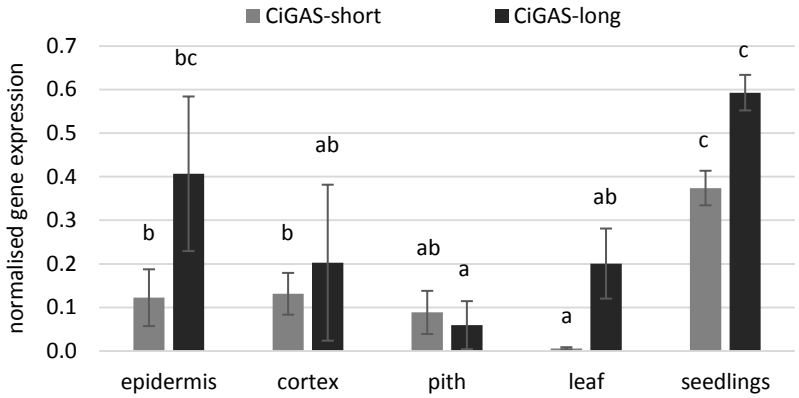

FPP
GAS-short
GAS-long
germacrene A
GAO
(CYP71AV4,
CYP71AV8)
germacrene A
acid
COS
(CYP71BL3)
6-hydroxy
germacrene A acid
Spontaneous
cyclisation
costunolide
lactucin
8- deoxy lactucin
lactucopicrin

Figure 2

A



B



C

D

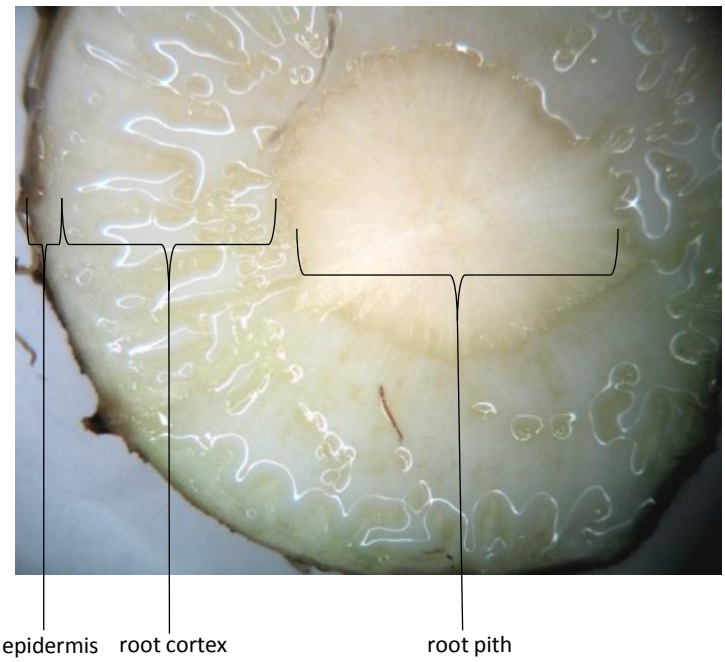


Figure 3

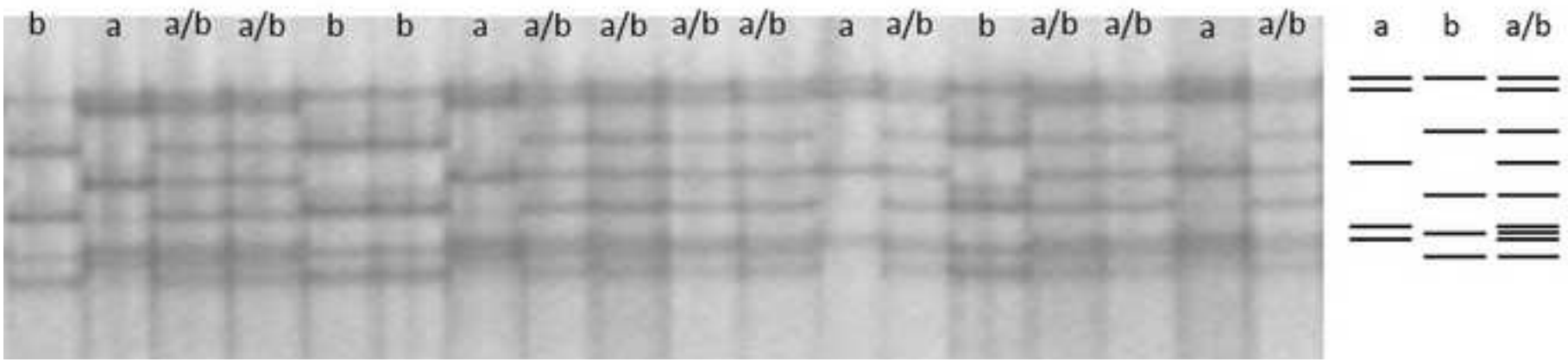


Figure 4

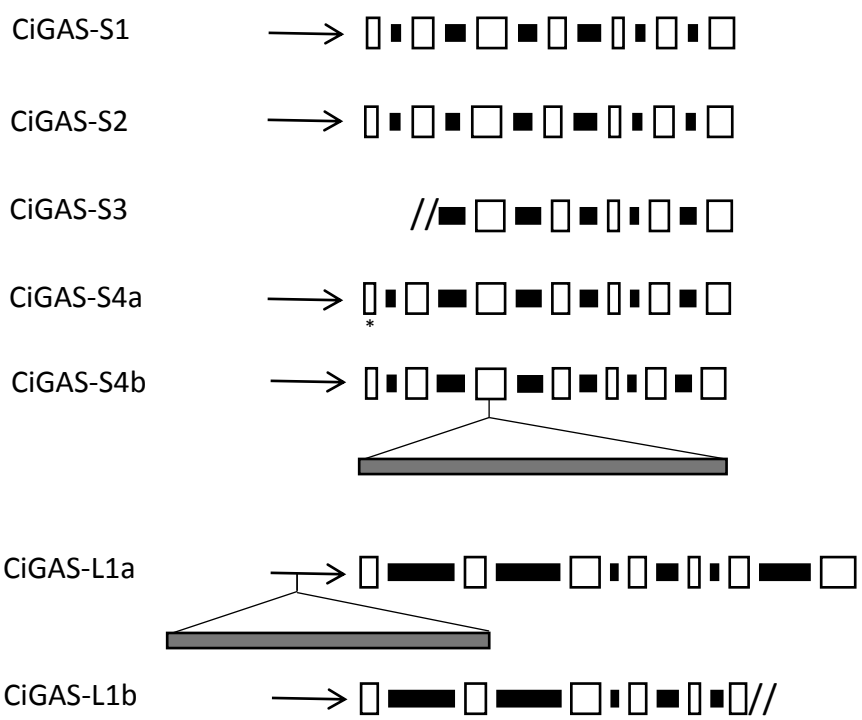
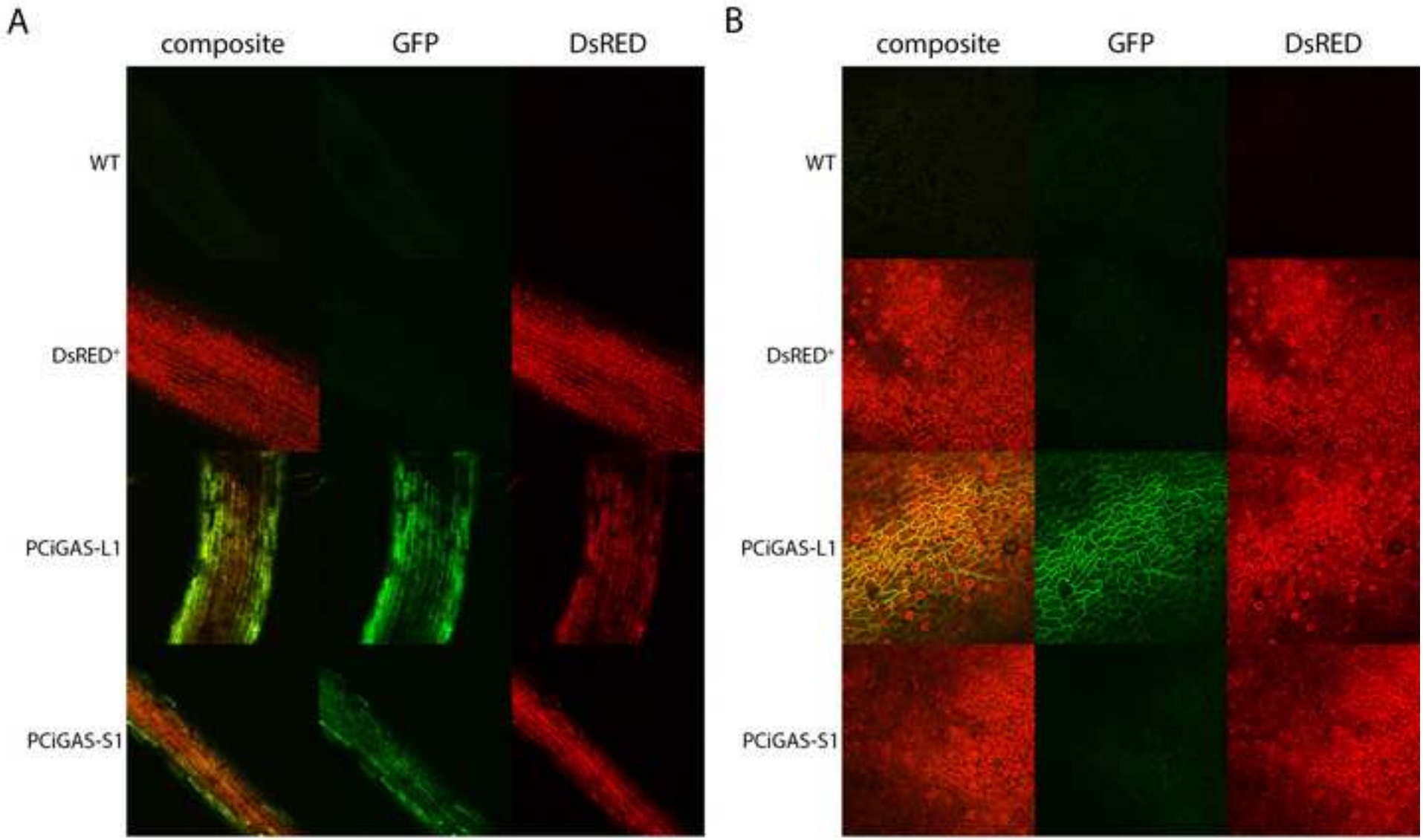

CiGAS-S1
CiGAS-S2
CiGAS-S3
CiGAS-S4a
*
CiGAS-S4b
CiGAS-L1a
CiGAS-L1b

Figure 5


A
composite
GFP
DsRED
WT
DsRED+
PCiGAS-L1
PCiGAS-S1
B
composite
GFP
DsRED
WT
DsRED+
PCiGAS-L1
PCiGAS-S1

Supplemental materials

Supplemental table 1: Primers used in this study

Primer name	Primer sequence	Purpose
GAS-long-as	AGACTCGGTGGAGGACTAACG	qPCR
GAS-short-s	ACCATTGAAGAGATTCGACTTCTG	qPCR
GAS-short-as	GCTTCTCAAGTTCAGCATACTCATC	qPCR
CH50F32-L19	CTGCCAGCGTCCTCAAGTG	qPCR
CH50R82-L19	CATTGGGATCAAGCCAAACCT	qPCR
ProGasL_R	ggggaccactttgtacaagaaagctgggtc TTTTCGGATTGGATGATCTTTAATG	Amplification of promoter from BAC 105O22, attB sites added for Gateway cloning
ProGasL_F	ggggacaagtttgtacaaaaaagcaggcttg TTACCAAACTGTGGTAGTTTGT	Amplification of promoter from BAC 105O22, attB sites added for Gateway cloning
ProGasS83_R	ggggaccactttgtacaagaaagctgggtc TTCCTGAAGATGAAAGGATATAGC	Amplification of promoter from BAC 83A09, attB sites added for Gateway cloning
ProGasS83_F	ggggacaagtttgtacaaaaaagcaggcttg GACTAATCTCGTCCACATGA	Amplification of promoter from BAC 83A09, attB sites added for Gateway cloning
ProGasS29_R	ggggaccactttgtacaagaaagctgggtcTTCTTGAAGGTGA AAGGATATAG	Amplification of promoter from BAC 29O10, attB sites added for Gateway cloning
ProGasS29_F	ggggacaagtttgtacaaaaaagcaggcttgGTACAACCGTTG CTTCTG	Amplification of promoter from BAC 29O10, attB sites added for Gateway cloning

243_F_6355	AGCTTGAGCTCTCCCATA	Colony PCR, promoter constructs in 243 pKGW-GGRR-C vector
+243_GasL_6642	GACGGTAGCCATTGGATT	Colony PCR, promoter constructs in 243 pKGW-GGRR-C vector
+243_GasS29_6647	GCAGTGGAAGGAGAATATGC	Colony PCR, promoter constructs in 243 pKGW-GGRR-C vector
+243_GasS83_6474	ACGTCCACCGTAAGGTTT	Colony PCR, promoter constructs in 243 pKGW-GGRR-C vector
seqProGasL_497_F	GCTCCAAAGGTTGTTAGTGA	Sequencing promoter fragment
seqProGasL_905_F	GTCAAACTCGCTTTACTCTC	Sequencing promoter fragment
seqProGasL_622_R	GCAAGAACATTCGTTCCATA	Sequencing promoter fragment
seqProGasL_1369_R	CATCGATTGGCTTTTGTATG	Sequencing promoter fragment
seqProGasS29_537_F	CAGAACTACAACCAGATCTA	Sequencing promoter fragment
seqProGasS29_992_F	CGCGTGTTTACAATTCATAC	Sequencing promoter fragment
seqProGasS29_1501_F	GTTTGTTACGGTTCACGTTT	Sequencing promoter fragment

seqProGasS29_642_R	TTCCTAACAAGTGGTATCAG	Sequencing promoter fragment
seqProGasS29_1124_R	AAAACTCAATTACACAGCAAAA	Sequencing promoter fragment
seqProGasS29_1650_R	TTTCGCTAAGTGCAGGTTAT	Sequencing promoter fragment
seqProGasS83_542_F	CACACGTAGAGGATTATACA	Sequencing promoter fragment
seqProGasS83_1117_F	GTCATTGTGAATGCCAACAG	Sequencing promoter fragment
seqProGasS83_683_R	CAAGAAAAGCTGCACTTATG	Sequencing promoter fragment
seqProGasS83_1186_R	TCCGCTAGTTGAGATTTACA	Sequencing promoter fragment
SEQ_0-1587_1_FORWARD	CTCGTGCTCATACCTGCTTG	BAC 59A14 sequencing
SEQ_0-1587_1_REVERSE	TCCCTTCAGCTCAGGTACAAA	BAC 59A14 sequencing
SEQ_102506-124907_1_FORWARD	TCATTGGAGTTTGCAGTTTGA	BAC 59A14 sequencing
SEQ_102506-124907_1_REVERSE	GACATGGACCTCTGGGCTAA	BAC 59A14 sequencing
SEQ_11273-15280_1_FORWARD	GCTCCACCTTTGGTCCAGTA	BAC 59A14 sequencing
SEQ_11273-15280_1_REVERSE	ACCCTAATCGACCCAAATCC	BAC 59A14 sequencing
SEQ_124907-132809_1_FORWARD	AGCCACCAAATGAACCAAAC	BAC 59A14 sequencing
SEQ_124907-132809_1_REVERSE	TTGCAAACATGCATGAAAGA	BAC 59A14 sequencing
SEQ_15280-18597_5_FORWARD	TTGTAGGACCGAGGCTCATC	BAC 59A14 sequencing

SEQ_15280-18597_5_REVERSE	TGGAAAGCCTTTTGGGTTTT	BAC 59A14 sequencing
SEQ_1587-6366_1_FORWARD	TTGCCTCTTACTGGGTCTGG	BAC 59A14 sequencing
SEQ_1587-6366_1_REVERSE	ACACGTACACACATGCTTGC	BAC 59A14 sequencing
SEQ_18597-25734_1_FORWARD	TAGCCCCATACTTCGGTTGA	BAC 59A14 sequencing
SEQ_18597-25734_1_REVERSE	CTTCAAGAACATCGGCTGCT	BAC 59A14 sequencing
SEQ_25734-35456_8_FORWARD	TGCTTACCGACTTCCACAAA	BAC 59A14 sequencing
SEQ_25734-35456_8_REVERSE	GGCCTCTCGATGAATTTGAA	BAC 59A14 sequencing
SEQ_35456-42520_4_FORWARD	AATGGGGTTTAAGGTTATCCA	BAC 59A14 sequencing
SEQ_35456-42520_4_REVERSE	TTCTACGAACACGCAAAATCC	BAC 59A14 sequencing
SEQ_42520-68650_1_FORWARD	CACATGCCTTTAGCGGTGTT	BAC 59A14 sequencing
SEQ_42520-68650_1_REVERSE	TTGTTGGCGTTCCATGTTTA	BAC 59A14 sequencing
SEQ_6366-11273_1_FORWARD	AAGCAAGCCTTTTCCAGACA	BAC 59A14 sequencing
SEQ_6366-11273_1_REVERSE	TGATGGGTACTCCCAATGGT	BAC 59A14 sequencing
SEQ_68650-97690_1_FORWARD	GCACGACTTAGATGCCATGA	BAC 59A14 sequencing
SEQ_68650-97690_1_REVERSE	CAGTGAAAGAAGCCCTCTGG	BAC 59A14 sequencing
SEQ_97690-102506_1_FORWARD	CGGCATTTTCTCCATTCCTA	BAC 59A14 sequencing
SEQ_97690-102506_1_REVERSE	GCAAGGCTATTTGGGGTTTT	BAC 59A14 sequencing
SEQ_0-14970_1_FORWARD	TGCGGACAGTGCTATTATGTG	BAC 83A09 sequencing
SEQ_0-14970_1_REVERSE	TGTCATTGTGAATGCCAACA	BAC 83A09 sequencing

SEQ_108466-109906_1_FORWARD	TCAAATTTGCACACCAAGGA	BAC 83A09 sequencing
SEQ_108466-109906_1_REVERSE	GCTTTGACAGGTTGCACTCA	BAC 83A09 sequencing
SEQ_109906-112306_1_FORWARD	ACCCTTAACTGGCCGATTTC	BAC 83A09 sequencing
SEQ_109906-112306_1_REVERSE	CACGTATGCATCCCTGAAAA	BAC 83A09 sequencing
SEQ_112306-115827_1_FORWARD	TGGTGACTATGGAGCCTGTG	BAC 83A09 sequencing
SEQ_112306-115827_1_REVERSE	TCTCGCGATCTCAACCTTTT	BAC 83A09 sequencing
SEQ_115827-125976_1_FORWARD	GAACTCACGTCGCAAATCAA	BAC 83A09 sequencing
SEQ_115827-125976_1_REVERSE	AGCAACTTCCTCCACTCGAA	BAC 83A09 sequencing
SEQ_125976-129895_1_FORWARD	AATGATTGGTCCACGTTTGG	BAC 83A09 sequencing
SEQ_125976-129895_1_REVERSE	AATGCCAATGTCGAATGCTT	BAC 83A09 sequencing
SEQ_129895-133651_4_FORWARD	AGAGGGGAGACCCCAACATA	BAC 83A09 sequencing
SEQ_129895-133651_4_REVERSE	CTTCGTTAGAGCCTCGTGCT	BAC 83A09 sequencing
SEQ_133651-137630_1_FORWARD	CATGATAGGCTGGGTCATCC	BAC 83A09 sequencing
SEQ_133651-137630_1_REVERSE	CCTCTTGCTGAAATTGAGGTG	BAC 83A09 sequencing
SEQ_137630-143861_1_FORWARD	ATTCGCACCAGTAGGATTGC	BAC 83A09 sequencing
SEQ_137630-143861_1_REVERSE	GTGCAACTGGATGCTCAACA	BAC 83A09 sequencing
SEQ_143861-147283_1_FORWARD	GCAACAATGCAGCAGATCTTT	BAC 83A09 sequencing
SEQ_143861-147283_1_REVERSE	GCCGGTGAATTTGTTTCTTT	BAC 83A09 sequencing
SEQ_14970-21361_1_FORWARD	AGAAGTGTTCGGCAAAAGCA	BAC 83A09 sequencing

SEQ_14970-21361_1_REVERSE	TGCAAGAAATAGCAACCCATT	BAC 83A09 sequencing
SEQ_21361-26091_1_FORWARD	GGAGGGCGTGTCCAATAATA	BAC 83A09 sequencing
SEQ_21361-26091_1_REVERSE	TGGGTGTTTCGGATGATTTT	BAC 83A09 sequencing
SEQ_26091-27459_1_FORWARD	AACCCATTTCTAGGGTTCTTGA	BAC 83A09 sequencing
SEQ_26091-27459_1_REVERSE	TTGGCAAGAGAGGGAAGTTG	BAC 83A09 sequencing
SEQ_27459-31613_1_FORWARD	GCGAAATTCGACTTGCCTAA	BAC 83A09 sequencing
SEQ_27459-31613_1_REVERSE	CCCAGAACAACCAAAACACC	BAC 83A09 sequencing
SEQ_31613-36257_1_FORWARD	CATGTGGGACTCAAATGGTG	BAC 83A09 sequencing
SEQ_31613-36257_1_REVERSE	TGTCCCATGAATTTCGTAAGC	BAC 83A09 sequencing
SEQ_36257-39500_2_FORWARD	TCACCCCTAGAATTGCCATC	BAC 83A09 sequencing
SEQ_36257-39500_2_REVERSE	TCTCCTGCCAAGTGATTTTG	BAC 83A09 sequencing
SEQ_39500-40674_1_FORWARD	AACCCATTTCTAGGGTTCTTGA	BAC 83A09 sequencing
SEQ_39500-40674_1_REVERSE	GAGAGAGCAGGGAGTGCAAC	BAC 83A09 sequencing
SEQ_40674-53846_1_FORWARD	GTTGGGGTGTGTGTGTGTGT	BAC 83A09 sequencing
SEQ_40674-53846_1_REVERSE	ACGGCTAACGGAAATGAAGA	BAC 83A09 sequencing
SEQ_53846-56201_1_FORWARD	CCTGGCCATCAGATTTCATT	BAC 83A09 sequencing
SEQ_53846-56201_1_REVERSE	TGACCACGATGGTATTGTTTT	BAC 83A09 sequencing
SEQ_56201-62778_1_FORWARD	TTTTGGTCCCTGAAGTTTGC	BAC 83A09 sequencing
SEQ_56201-62778_1_REVERSE	TGGGAAACATGTGTGTATGGA	BAC 83A09 sequencing

SEQ_62778-64149_1_FORWARD	CTCCCTCGTCCTCATTTCTG	BAC 83A09 sequencing
SEQ_62778-64149_1_REVERSE	GATGTGACGGTGAAGTGACG	BAC 83A09 sequencing
SEQ_64149-71972_1_FORWARD	TTGGGAAGTTCAAGGAGTGG	BAC 83A09 sequencing
SEQ_64149-71972_1_REVERSE	AAATCCCTTCGCCACTAACC	BAC 83A09 sequencing
SEQ_71972-78416_1_FORWARD	TAGGTTGCTTTTGCGTTGTG	BAC 83A09 sequencing
SEQ_71972-78416_1_REVERSE	AAGTCGAGGTGTGGGACAGT	BAC 83A09 sequencing
SEQ_78416-84007_1_FORWARD	GGTTGTTTCGCGAACGTAGT	BAC 83A09 sequencing
SEQ_78416-84007_1_REVERSE	ATGCCAAACCACACCAAAAC	BAC 83A09 sequencing
SEQ_84007-90090_1_FORWARD	TGTGCCCTTACACATCATCC	BAC 83A09 sequencing
SEQ_84007-90090_1_REVERSE	TATCTTTTGGTGGGCGTCTC	BAC 83A09 sequencing
SEQ_90090-96283_1_FORWARD	ATTGAGTGGACCGCCAGTTA	BAC 83A09 sequencing
SEQ_90090-96283_1_REVERSE	GATTTACGAACGCACACACG	BAC 83A09 sequencing
SEQ_96283-99247_1_FORWARD	TGTCCAACCACCAATCAACT	BAC 83A09 sequencing
SEQ_96283-99247_1_REVERSE	TCGATAGCGAATGAAGATGG	BAC 83A09 sequencing
SEQ_99247-108466_1_FORWARD	GCACCCAGAACAAACAAACA	BAC 83A09 sequencing
SEQ_99247-108466_1_REVERSE	TCGTTGTCTTGGTTCAATTTGT	BAC 83A09 sequencing
AF498000 L2	CGTTACTCCTTGGCACGAAT	Genetic mapping & SSCP analysis *CiGAS-CiGAS-S1* (Cadalen et al 2010)
AF498000 R2	TAGCTCGTCAATCGCTTCCT	Genetic mapping & SSCP analysis

		CiGAS-CiGAS-S1 (Cadalen et al 2010)
AF497999 L2	CAAGCCATGAGTCGGTTGTA	Genetic mapping & SSCP analysis *CiGAS-long* (Cadalen et al 2010)
AF497999 R2	ACGTGCCCGAGAGTAATACG	Genetic mapping & SSCP analysis *CiGAS-long* (Cadalen et al 2010)
AF497999 L3	GGGCACGTATCATAGCCACT	Genetic mapping & SSCP analysis *CiGAS-long* (Cadalen et al 2010)
AF497999 R3	TTGAAGCGTGGACACTGAAC	Genetic mapping & SSCP analysis *CiGAS-long* (Cadalen et al 2010)
AF497999 For	GATCCGATCGATCACAGATG	BAC screening and validation
AF497999 Rev	CGGTTTGATATATTCAGGAAGTTG	BAC screening and validation
AF498000 For	CCTCCTTCTGTATGGGGTGA	BAC screening and validation
AF498000 Rev	AAGCTGCGCTTCAATCTCTT	BAC screening and validation
EH705708 For	AAAGACTCTACAAGCTTCCGAGTTA	BAC screening and validation
EH705708 Rev	CCCGTCATCATATCTATACACAACA	BAC screening

		and validation
ADF43080 For	ATAGCGCTGAGTTCCCGTTA	BAC screening and validation
ADF43080 Rev	GACTGTGGCAGATGACGTGT	BAC screening and validation
HQ166835 For	AAGAAAGCGCAGAGTTTCCA	BAC screening and validation
HQ166835 Rev	GGATCCCTGTTTATGGCAAA	BAC screening and validation
DsRED_F	TCGTTTGTGGGAGGTGATGTCCA	Screening of the hairy root transformants
DsRED_R	CTTCAAGTGGGAGCGCGTGA	Screening of the hairy root transformants
rolA1_F	GCGAACGCGACCATCTTGCT	Screening of the hairy root transformants
rolA1_R	GGTCCCTTCGCAGCAACTCG	Screening of the hairy root transformants
virD1_F	ATGTCGCAAGGCAGTAAG	Screening of the hairy root transformants
virD1_R	CAAGGAGTCTTTCAGCATG	Screening of the hairy root transformants

Supplemental table 2. Summary of chicory transformation efficiency using *R. rhizogenes* A4M70GUS strains carrying promoter constructs or control vectors.

Transformed construct	Initial number of hairy root clones	Number of hairy root clones that regenerated plants	Number of hairy root clones tested positive by PCR	Number of regenerated plants tested by PCR	Number of regenerated plants tested positive by PCR
DsRED+	53	19	12	62	37
GFP+	54	14	9	66	14
PCiGAS-L1b	60	27	19	93	67
PCiGAS-S2	48	3	3	11	11
PCiGAS-S1	85	29	26	129	100
total	349	103	74	391	245

Supplemental table 3 Comparison *Lactuca sativa* gene models, compared to chicory proteins

Lactuca gene model	Best chicory protein hit (BLAST-X)	% identity (protein)	Localization on Lactuca genome
Lsat_1_v5_gn_2_29221.1	CiGAS-L1	89%	Chromosome 2
Lsat_1_v5_gn_8_116340.1	CiGAS-S1	73%	Chromosome 8
Lsat_1_v5_gn_8_116421.1	CiGAS-S1	72%	Chromosome 8

Supplemental figure 1: Genetic map of LG9 of the K28K59 population

The revised genetic map of the LG9 is shown, showing a single locus for *CiGAS-long* gene. Linkage analysis and map calculations were performed using the program JoinMap v 3.0 (van Ooijen and Voorrips 2001). Linkage grouping was determined using a LOD threshold = 3. JoinMap setting parameters were: maximum recombination frequency = 0.45, LOD = 1.0, jump = 5. Pair-wise recombination frequencies were converted in map distances (cM) using the Kosambi mapping function (Kosambi 1944). The map was drawn using the program MapChart v 2.32 (Voorrips 2002). SSR and STS markers were described previously (Cadalen et al, 2010). Amplified fragment length polymorphism (AFLP, Vos et al. 1995) analysis was performed as described in Gonthier et al, 2013, and AFLP markers are indicated by the four selective nucleotides, i.e. two on the left and right primer, respectively, followed by the length of the amplified fragment.

References:

Gonthier L, Blassiau C, Mörchen M, Cadalen T, Poiret M, Hendriks T, Quillet M-C (2013) High-density genetic maps for loci involved in nuclear male sterility (NMS1) and sporophytic self-incompatibility (S-locus) in chicory (Cichorium intybus L., Asteraceae). Theor Appl Genet 126:2103–2121.

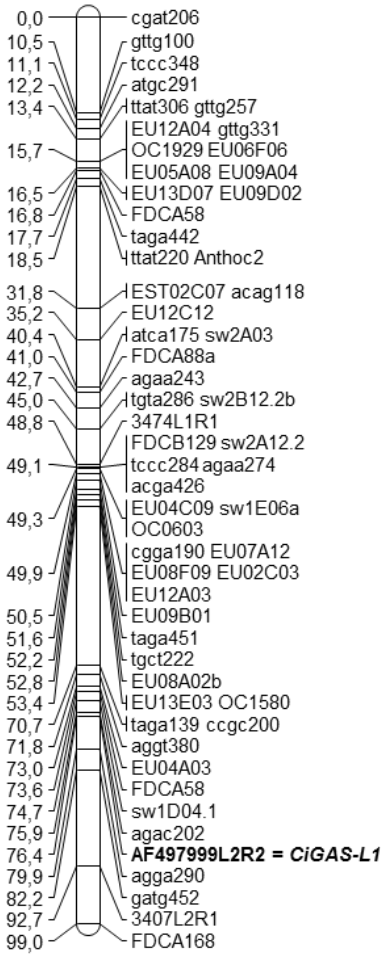
Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eug 12:172–175.

van Ooijen JW, Voorrips RE (2001) JoinMap 3.0, software for the calculation of genetic linkage maps. Plant Research International, Wageningen.

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77– 78.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23(21): 4407–4414.

LG9 K28K59


0,0
10,5
11,1
12,2
13,4
15,7
16,5
16,8
17,7
18,5
31,8
35,2
40,4
41,0
42,7
45,0
48,8
49,1
49,3
49,9
50,5
51,6
52,2
52,8
53,4
70,7
71,8
73,0
73,6
74,7
75,9
76,4
79,9
82,2
92,7
99,0
cgat206
gttg100
tccc348
atgc291
ttat306 gttg257
EU12A04 gttg331
OC1929 EU06F06
EU05A08 EU09A04
EU13D07 EU09D02
FDCA58
taga442
ttat220 Anthoc2
EST02C07 acag118
EU12C12
atca175 sw2A03
FDCA88a
agaa243
tgta286 sw2B12.2b
3474L1R1
FDCB129 sw2A12.2
tccc284 agaa274
acga426
EU04C09 sw1E06a
OC0603
cgga190 EU07A12
EU08F09 EU02C03
EU12A03
EU09B01
taga451
tgct222
EU08A02b
EU13E03 OC1580
taga139 ccgc200
aggt380
EU04A03
FDCA58
sw1D04.1
agac202
AF497999L2R2 = CiGAS-L1
agga290
gatg452
3407L2R1
FDCA168

Supplemental figure 2

Percentage of identity matrices found in the sequence comparisons between germacrene A synthases for exons, introns, promoter regions (bp 0-350 and 351-1000) and terminator regions. Shading reflects the percentage of homology.

Exon1

	CiGAS-L1a	CiGAS-L1b	CiGAS-S1	CiGAS-S2	CiGAS-S4a	CiGAS-S4b
CiGAS-L1a		100,0	59,7	58,9	61,0	60,0
CiGAS-L1b	100,0		59,7	58,9	61,0	60,0
CiGAS-S1	58,9	58,9		97,7	68,6	68,6
CiGAS-S2	59,7	59,7	97,7		70,5	70,5
CiGAS-S4a	60,0	60,0	70,5	68,6		94,3
CiGAS-S4b	61,0	61,0	70,5	68,6	94,3	

Exon2

	CiGAS-L1a	CiGAS-L1b	CiGAS-S4a	CiGAS-S4b	CiGAS-S2	CiGAS-S1
CiGAS-L1a		100,0	69,4	69,4	68,6	68,6
CiGAS-L1b	100,0		69,4	69,4	68,6	68,6
CiGAS-S4a	69,4	69,4		98,2	84,9	84,5
CiGAS-S4b	69,4	69,4	98,2		86,0	85,2
CiGAS-S2	68,6	68,6	84,9	86,0		98,2
CiGAS-S1	68,6	68,6	84,5	85,2	98,2	

Exon3

	CiGAS-L1a	CiGAS-L1b	CiGAS-S1	CiGAS-S2	CiGAS-S3	CiGAS-S4a	CiGAS-S4b
CiGAS-L1a		98,9	71,5	71,8	71,3	69,2	68,1
CiGAS-L1b	98,9		72,1	72,3	71,8	69,7	68,6
CiGAS-S1	71,8	72,3		98,7	91,0	86,4	87,0
CiGAS-S2	71,5	72,1	98,7		91,5	87,0	87,5
CiGAS-S3	71,3	71,8	91,5	91,0		93,6	93,4
CiGAS-S4a	68,1	68,6	87,5	87,0	93,4		96,6
CiGAS-S4b	69,2	69,7	87,0	86,4	93,6	96,6	

Exon4

	CiGAS-L1a	CiGAS-L1b	CiGAS-S1	CiGAS-S2	CiGAS-S3	CiGAS-S4b	CiGAS-S4a
CiGAS-L1a		98,2	77,2	76,7	76,7	74,9	74,0
CiGAS-L1b	98,2		77,6	77,2	76,3	75,3	74,4
CiGAS-S1	76,7	77,2		97,7	92,7	92,7	91,8
CiGAS-S2	77,2	77,6	97,7		92,2	92,2	91,3
CiGAS-S3	76,7	76,3	92,2	92,7		93,2	92,2
CiGAS-S4b	74,9	75,3	92,2	92,7	93,2		99,1
CiGAS-S4a	74,0	74,4	91,3	91,8	92,2	99,1	

Exon5

	CiGAS-L1a	CiGAS-L1b	CiGAS-S1	CiGAS-S2	CiGAS-S3	CiGAS-S4a	CiGAS-S4b
CiGAS-L1a		98,6	73,4	71,9	70,5	69,8	71,2
CiGAS-L1b	98,6		72,7	71,2	70,5	69,1	71,2
CiGAS-S1	71,9	71,2		96,4	89,2	89,9	89,9
CiGAS-S2	73,4	72,7	96,4		89,2	89,9	90,7
CiGAS-S3	70,5	70,5	89,2	89,2		93,5	95,0
CiGAS-S4a	71,2	71,2	90,7	89,9	95,0		97,1
CiGAS-S4b	69,8	69,1	89,9	89,9	93,5	97,1	

Exon6

	CiGAS-L1a	CiGAS-L1b	CiGAS-S4b	CiGAS-S4a	CiGAS-S3	CiGAS-S2	CiGAS-S1
CiGAS-L1a		100,0	73,1	73,1	78,5	81,4	81,0
CiGAS-L1b	100,0		71,7	71,7	77,3	79,8	78,8
CiGAS-S4b	73,1	71,7		98,8	92,0	85,9	86,4
CiGAS-S4a	73,1	71,7	98,8		91,6	85,1	85,5
CiGAS-S3	78,5	77,3	92,0	91,6		93,2	93,6
CiGAS-S2	81,4	79,8	85,9	85,1	93,2		98,4
CiGAS-S1	81,0	78,8	86,4	85,5	93,6	98,4	

Exon7

	CiGAS-L1a	CiGAS-S2	CiGAS-S1	CiGAS-S3	CiGAS-S4b	CiGAS-S4a
CiGAS-L1a		74,4	74,4	73,4	73,1	73,1
CiGAS-S2	74,4		98,0	86,9	85,9	85,9
CiGAS-S1	74,4	98,0		86,2	85,5	85,5
CiGAS-S3	73,4	86,9	86,2		93,9	93,6
CiGAS-S4b	73,1	85,9	85,5	93,9		99,7
CiGAS-S4a	73,1	85,9	85,5	93,6	99,7	

Intron1

	CiGAS-S2	CiGAS-S1	CiGAS-L1a	CiGAS-L1b	CiGAS-S4b	CiGAS-S4a
CiGAS-S2		87,6	52,3	52,3	57,1	57,1
CiGAS-S1	87,6		51,4	51,4	52,4	52,4
CiGAS-L1a	52,3	51,4		99,8	60,2	60,2
CiGAS-L1b	52,3	51,4	99,8		60,2	60,2
CiGAS-S4b	57,1	52,4	60,2	60,2		100,0
CiGAS-S4a	57,1	52,4	60,2	60,2	100,0	

Intron2

	CiGAS-L1a	CiGAS-L1b	CiGAS-S2	CiGAS-S1	CiGAS-S3	CiGAS-S4b	CiGAS-S4a
CiGAS-L1a		100,0	42,3	41,4	47,8	49,9	49,4
CiGAS-L1b	100,0		42,3	41,4	47,8	49,9	49,4
CiGAS-S2	42,3	42,3		92,0	50,0	57,7	54,9
CiGAS-S1	41,4	41,4	92,0		55,1	58,1	55,8
CiGAS-S3	47,8	47,8	50,0	55,1		82,4	80,3
CiGAS-S4b	49,9	49,9	57,7	58,1	82,4		97,6
CiGAS-S4a	49,4	49,4	54,9	55,8	80,3	97,6	

Intron3

	CiGAS-L1a	CiGAS-L1b	CiGAS-S2	CiGAS-S1	CiGAS-S3	CiGAS-S4b	CiGAS-S4a
CiGAS-L1a		84,4	57,5	57,5	58,4	60,5	60,5
CiGAS-L1b	84,4		56,0	56,0	59,5	59,0	59,0
CiGAS-S2	57,5	56,0		98,6	66,2	66,4	66,4
CiGAS-S1	57,5	56,0	98,6		67,6	67,3	66,8
CiGAS-S3	58,4	59,5	66,2	67,6		88,5	89,8
CiGAS-S4b	60,5	59,0	66,4	67,3	88,5		96,7
CiGAS-S4a	60,5	59,0	66,4	66,8	89,8	96,7	

Intron4

	CiGAS-L1a	CiGAS-L1b	CiGAS-S3	CiGAS-S4b	CiGAS-S4a	CiGAS-S2	CiGAS-S1
CiGAS-L1a		82,5	39,6	41,6	43,1	52,6	47,0
CiGAS-L1b	82,5		37,3	38,6	41,6	51,3	47,4
CiGAS-S3	39,6	37,3		84,3	83,6	55,8	52,3
CiGAS-S4b	41,6	38,6	84,3		95,5	53,1	51,0
CiGAS-S4a	43,1	41,6	83,6	95,5		55,9	52,9
CiGAS-S2	52,6	51,3	55,8	53,1	55,9		82,6
CiGAS-S1	47,0	47,4	52,3	51,0	52,9	82,6	

Intron5

	CiGAS-L1a	CiGAS-L1b	CiGAS-S2	CiGAS-S1	CiGAS-S3	CiGAS-S4b	CiGAS-S4a
CiGAS-L1a		100,0	48,3	50,0	44,7	48,2	47,0
CiGAS-L1b	100,0		48,3	50,0	44,7	48,2	47,0
CiGAS-S2	48,3	48,3		83,8	60,4	60,7	59,6
CiGAS-S1	50,0	50,0	83,8		58,2	58,4	57,3
CiGAS-S3	44,7	44,7	60,4	58,2		82,6	81,5
CiGAS-S4b	48,2	48,2	60,7	58,4	82,6		98,9
CiGAS-S4a	47,0	47,0	59,6	57,3	81,5	98,9	

Intron6

	CiGAS-S3	CiGAS-S4b	CiGAS-S4a	CiGAS-L1a	CiGAS-S2	CiGAS-S1
CiGAS-S3		80,7	80,1	48,1	33,3	33,3
CiGAS-S4b	80,7		99,5	47,3	33,3	33,3
CiGAS-S4a	80,1	99,5		47,3	33,3	33,3
CiGAS-L1a	48,1	47,3	47,3		57,1	51,5
CiGAS-S2	33,3	33,3	33,3	57,1		91,3
CiGAS-S1	33,3	33,3	33,3	51,5	91,3	

Prom350

	CiGAS-L1a	CiGAS-L1b	CiGAS-S4b	CiGAS-S4a	CiGAS-S2	CiGAS-S1
CiGAS-L1a		100	44,2	43,8	45,6	43,3
CiGAS-L1b	100		44,2	43,8	45,6	43,3
CiGAS-S4b	44,2	44,2		98,3	52,9	51,0
CiGAS-S4a	43,8	43,8	98,3		53,0	51,2
CiGAS-S2	45,6	45,6	52,9	53,0		85,8
CiGAS-S1	43,3	43,3	51,0	51,2	85,8	

Promo1000

	CiGAS-S4a	CiGAS-S4b	CiGAS-L1a	CiGAS-L1b	CiGAS-S2	CiGAS-S1
CiGAS-S4a		37,3	40,2	40,0	42,4	40,2
CiGAS-S4b	37,3		42,5	42,5	38,9	43,3
CiGAS-L1a	40,2	42,5		99,3	44,8	46,6
CiGAS-L1b	40,0	42,5	99,3		44,2	46,1
CiGAS-S2	42,4	38,9	44,8	44,2		52,1
CiGAS-S1	40,2	43,3	46,6	46,1	52,1	

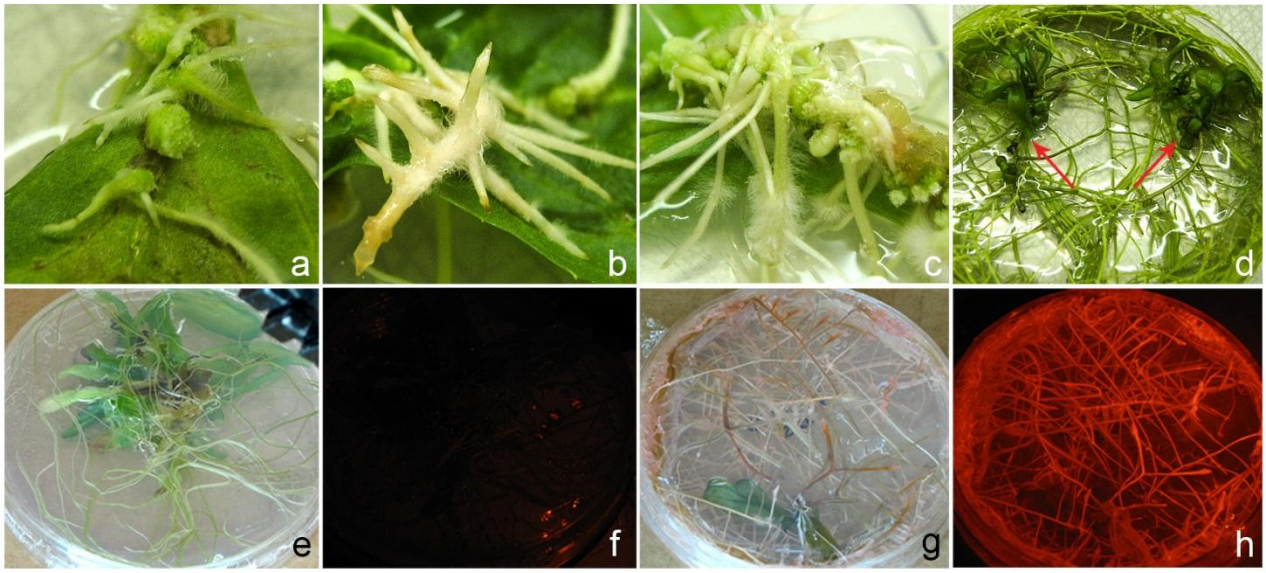
Termi200

	CiGAS-S4b	CiGAS-S4a	CiGAS-S3	CiGAS-L1a	CiGAS-S2	CiGAS-S1
CiGAS-S4b		98,6	39,7	39,7	39,6	39,7
CiGAS-S4a	98,6		39,7	39,7	38,8	39,7
CiGAS-S3	39,7	39,7		43,5	47,2	41,9
CiGAS-L1a	39,7	39,7	43,5		45,9	42,1
CiGAS-S2	39,6	38,8	47,2	45,9		51,4
CiGAS-S1	39,7	39,7	41,9	42,1	51,4	

ORF

	CiGAS-L1b	CiGAS-L1a	CiGAS-S2	CiGAS-S1	CiGAS-S3	CiGAS-S4b	CiGAS-S4a
CiGAS-L1b		99,3	72,0	71,6	73,4	70,1	69,8
CiGAS-L1a	99,3		72,8	72,5	73,7	70,8	70,5
CiGAS-S2	72,0	72,8		98,0	90,6	86,2	86,3
CiGAS-S1	71,6	72,5	98,0		90,5	85,9	85,9
CiGAS-S3	73,4	73,7	90,6	90,5		93,3	93,1
CiGAS-S4b	70,1	70,8	86,2	85,9	93,3		98,0
CiGAS-S4a	69,8	70,5	86,3	85,9	93,1	98,0	

Supplemental figure 3



Supplemental figure 3. Chicory hairy root transformation and selection. Emerging roots 10 days (a) and four weeks (b and c) after inoculation. Spontaneously formed shoots (arrows) appeared on a month old hairy root culture (d). DsRED visualization in the chicory hairy root culture transformed with the *R. rhizogenes* A4M70GUS strain (e, f) and *R. rhizogenes* A4M70GUS carrying the pKGW-GGRR-C plasmid (g, h) under white light (e, g) and green LED (515-530 nm) with red LP (600 nm) filter (f, h).