

Isolation and mapping of a C3'H gene (CYP98A49) from globe artichoke, and its expression upon UV-C stress

Andrea Moglia · Cinzia Comino · Ezio Portis ·
Alberto Acquadro · Ric C. H. De Vos ·
Jules Beekwilder · Sergio Lanteri

Received: 16 January 2009 / Revised: 10 February 2009 / Accepted: 1 March 2009 / Published online: 20 March 2009
© Springer-Verlag 2009

Abstract Globe artichoke represents a natural source of phenolic compounds with dicaffeoylquinic acids along with their biosynthetic precursor chlorogenic acid (5-caffeoylquinic acid) as the predominant molecules. We report the isolation and characterization of a full-length cDNA and promoter of a globe artichoke *p*-coumaroyl ester 3'-hydroxylase (CYP98A49), which is involved in both chlorogenic acid and lignin biosynthesis. Phylogenetic analyses demonstrated that this gene belongs to the CYP98 family. CYP98A49 was also heterologously expressed in yeast, in order to perform an enzymatic assay with *p*-coumaroylshikimate and *p*-coumaroylquininate as substrates. Real Time quantitative PCR analysis revealed that CYP98A49 expression is induced upon exposure to UV-C radiation. A single nucleotide polymorphism in the CYP98A49 gene sequence of two globe artichoke varieties used for genetic mapping allowed the localization of this gene to linkage group 10 within the previously developed maps.

Keywords Globe artichoke · C3'H · Genetic mapping · UV-C radiation

Introduction

Globe artichoke ($2n = 2x = 34$, *Cynara cardunculus* L. var. *scolymus*) is a perennial, allogamous, primarily vegetatively propagated vegetable, native to the Mediterranean basin. Its cultivation makes a considerable contribution to the agricultural economy of southern Europe, with Italy being the major world producer (about 470 Mt per year). Globe artichoke is used mostly for human food, although the polyphenolic content of its leaves is known to have therapeutic properties. The predominant phenolics present in the leaf are the dicaffeoylquinic acids and chlorogenic acid (5-caffeoylquinic acid) (Lattanzio et al. 1994; Wang et al. 2003). These metabolites possess antioxidative, hepatoprotective, diuretic and choleric activity (Adzet et al. 1987; Gebhardt 1997; Brown and Rice-Evans 1998). Leaf extracts have also been reported to inhibit cholesterol biosynthesis, to contribute to the prevention of arteriosclerosis and other cardiovascular diseases (Gebhardt 1998), and may inhibit HIV integrase, a key enzyme in viral replication and insertion into host DNA (Slanina et al. 2001).

The reconstruction of the phenylpropanoid pathway and the elucidation of the biological functions of secondary metabolites is an area of active research (Douglas 1996). Phenylpropanoid compounds, which include flavonoids, lignin, coumarins and many small phenolic compounds, contribute to a multiplicity of plant functions, such as the strengthening of the cell wall, pigmentation of the flower, defense against pathogens and cell signaling (Boudet 2007).

Nucleotide sequence data reported are available in the GenBank database under accession number: FJ225121

Communicated by H. Judelson.

A. Moglia · C. Comino · E. Portis · A. Acquadro ·
S. Lanteri (✉)
DiVaPRA, Plant Genetics and Breeding, University of Torino,
via L. da Vinci 44, 10095 Grugliasco (TO), Italy
e-mail: sergio.lanteri@unito.it

R. C. H. De Vos · J. Beekwilder
Plant Research International, P.O. Box 16,
6700 AA Wageningen, The Netherlands

To date few studies on phenylpropanoid biosynthetic pathway have been performed in globe artichoke. Three sequences with high similarity to PAL (phenylalanine ammonia-lyase), the first enzyme involved in the phenylpropanoid pathway, have been isolated (De Paolis et al. 2008). Moreover, gene sequences encoding hydroxycinnamoyltransferase (HCT and HQT), involved in the synthesis of chlorogenic acid, have been recently identified (Comino et al. 2007, 2009). New insights are thus required for identifying other genes involved in caffeoylquinic acid synthesis and elucidating the causal relationships between the various enzymes.

The 3'-hydroxylation of the phenylpropanoid ring is essential not only for the synthesis of lignin (Franke et al. 2002; Abdulrazzak et al. 2006; Chen and Dixon 2007), but is also an important step in the synthesis of chlorogenic acid (Schoch et al. 2006; Mahesh et al. 2007). The CYP98 family of plant cytochromes P450 has been designated as the family of enzyme that performs the *meta*-hydroxylation step in the phenylpropanoid pathway (Fig. 1). This *meta*-hydroxylation is not catalyzed on free *p*-coumaric acid, but on its conjugates with shikimic, quinic or phenyllactic acids (Schoch et al. 2006). The protein and encoding genes in this family are also referred to as C3'H.

The phenylpropanoid pathway is activated upon exposure to abiotic and biotic stresses, such as wounding, UV irradiation and pathogen attack (Dixon and Paiva 1995; Treutter 2005). The involvement in the stress response of chlorogenic acid has been reported, with its concentration increasing in lettuce upon wounding, and in

tobacco upon both UV-B irradiation and insect feeding (Cantos et al. 2001; Izaguirre et al. 2007). In globe artichoke, we have recently shown that exposure to UV-C consistently increases the level of the major dicaffeoylquinic acid isomer, while the exogenous application of either methyl jasmonate or salicylic acid has no such effect (Moglia et al. 2008). Furthermore we previously generated the first genetic map of globe artichoke, based on a two-way pseudo-testcross strategy (Lanteri et al. 2006). An F₁ population was created by crossing 'Romanesco clone C3' (a late-maturing, non-spiny type) with 'Spinoso di Palermo' (an early-maturing spiny type), and the progeny was genotyped using a number of marker types, such as AFLP, S-SAP, SSR and M-AFLP (Lanteri et al. 2006).

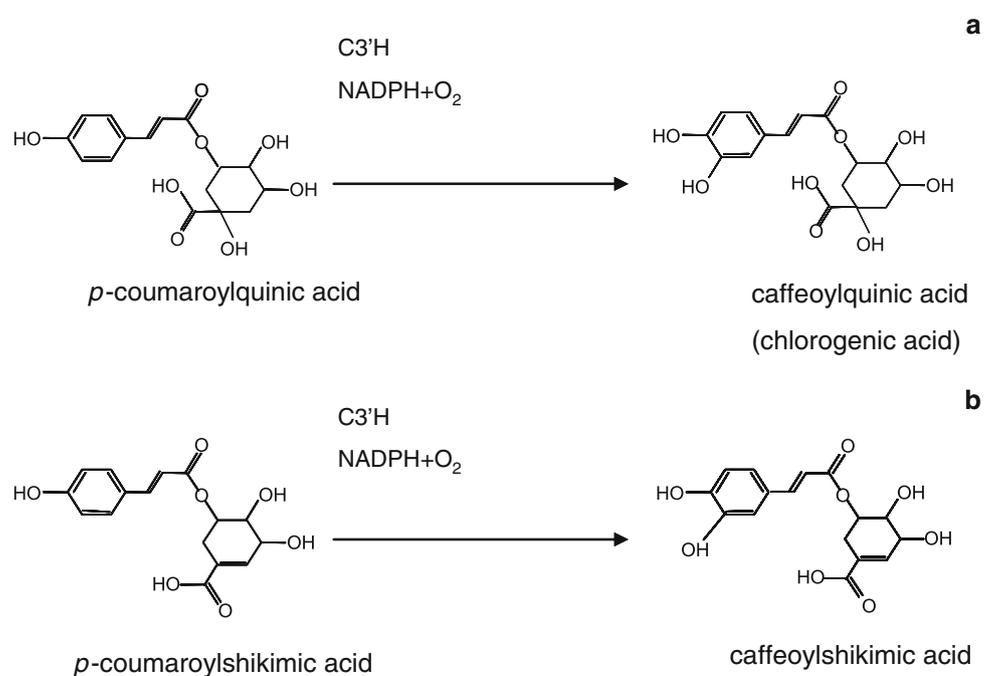
The main objectives of this paper are (a) the cloning of a C3'H gene and its promoter from globe artichoke, (b) the study of its transcriptional regulation in response to UV-C irradiation and (c) the definition of the gene's localization in a previously developed map in relation to other DNA-based markers. The *in vitro* enzymatic assay of C3'H was also performed.

Materials and methods

Plant material, DNA and RNA extraction

Seeds of globe artichoke ('Concerto', Nunhems) were germinated on a 15-cm Petri dish on two layers of wetted filter paper, and after 2 weeks transplanted into soil-filled

Fig. 1 Coumaroyl-ester metabolism affected by recombinant CYP98 genes. Conversion of **a** *p*-coumaroylquinic acid to 5-caffeoylquinic acid (chlorogenic acid), **b** *p*-coumaroylshikimate to caffeoylshikimate



10 cm pots in a glasshouse held at 22–24°C for 10 weeks. DNA was extracted from leaves, following Lanteri et al. (2001). Total RNA was extracted from approximately 100 mg fresh leaf using the Trizol reagent (Invitrogen), following the manufacturer's instructions.

Cloning of the C3'H gene and its promoter

Three sets of degenerate primers (COD C3'H For, COD C3'H Nested and COD C3'H Rev, see Table 1) were designed for the amplification of C3'H from the amino acid sequences of conserved regions of *Arabidopsis thaliana* (NP_850337), *Nicotiana tabacum* (ABC69384), *Coffea canephora* (ABB83677), *Ocimum basilicum* (AAL99200), *Medicago truncatula* (ABC59086) and *Sesamum indicum* (AAL47545) homologs. PCR was performed from 25 ng template of genomic DNA following the CODEHOP strategy (Morant et al. 2002) which employs a 3 min denaturation at 94°C, followed by 20 touch-down cycles [(94°C/1 min, 70°C/2 min (reducing by 1°C each cycle) and 72°C/2 min], and 29 cycles of 94°C/1 min, primer annealing temperature/1 min and 72°C/10 min, and finishing with a 10 min incubation at 72°C. The PCR fragment, A-tailed by the use of *Taq* DNA polymerase (Promega), was extracted from the agarose gel and cloned into the p-GEMTeasy vector (Promega, Madison, WI, USA). Plasmids from two colonies were sequenced using ABI310 capillary sequencer (Applied Biosystem). To assign a putative function, a BLAST search was performed against Viridiplantae GenBank database. To obtain a complete cDNA sequence, the SMART RACE cDNA amplification kit (Clontech) was employed, following the manufacturer's instructions, except for the use of an annealing temperature 5–10°C less than was recommended. Specific primers were designed for 3' and 5' amplification of the C3'H transcript, based on the partial

cDNA sequence. Fragments of the expected size were extracted from an agarose gel, cloned into the p-GEMTeasy vector and sequenced. The CYP number of the isolated gene was provided by Dr Nelson (<http://drnelson.utm.edu/CytochromeP450.html>). Multiple global sequence alignments were performed using ClustalW (<http://www.ebi.ac.uk/clustalw/>), applying default parameters. Phylogenetic analysis was performed using MEGA v3.0 (Kumar et al. 2004). Protein sequences used for alignments were: CYP98A1 (*Sorghum bicolor*, AAC39316), CYP98A2 (*Glycine max*, AAB94587), CYP98A3 (*A. thaliana*, NP_850337), CYP98A4 (*Oryza sativa*, AAU44038), CYP98A6 (*Lithospermum erythrorhizon*, AB017418), CYP98A8 (*A. thaliana*, NP_177594), CYP98A9 (*A. thaliana*, NP_177595), CYP98A10 (*Triticum aestivum*, CAE47489), CYP98A11 (*T. aestivum*, CAE47490), CYP98A12 (*T. aestivum*, CAE47491), CYP98A13v1 (*O. basilicum*, AAL99200), CYP98A13v2 (*O. basilicum*, AAL99201), CYP98A14 (*Solenostemon scutellarioides*, CAD20576), CYP98A19 (*Pinus taeda*, AAL47685), CYP98A20 (*S. indicum*, AAL47545), CYP98A21 (*Ammi majus*, AAT06912), CYP98A27 (*Populus trichocarpa*, ACC63870), CYP98A28 (*Camptotheca acuminata*, AAS57921), CYP98A29 (*Zea mays*, assembled from GSS sequences), CYP98A33v1 (*N. tabacum*, ABC69384), CYP98A35 (*C. canephora*, ABB83676), CYP98A36 (*C. canephora*, ABB83677), CYP98A37 (*M. truncatula*, ABC59086), CYP98A39v1 (*T. aestivum*, AJ585988), CYP98A39v2 (*T. aestivum*, AJ585990), CYP98A39v3 (*T. aestivum*, AJ585991), CYP98A40 (*T. aestivum*, AJ585989), CYP98A41 (*T. aestivum*, AJ585987), CYP98A46 (*Coptis japonica*, BAF98473). The Genome Walker Universal kit (Clontech) was used to isolate the globe artichoke C3'H promoter, following the manufacturer's protocol. Fragments were isolated after agarose gel separation, cloned into the p-GEMTeasy vector and sequenced. The promoter's regulatory motifs were identified according to Narusaka et al. (2004).

Table 1 Primer sequences used in this study for gene isolation (COD), for heterologous expression (Expr), for genetic mapping (C3'H447) and for RT-qPCR (RT)

COD C3'H For	5'-GATGARGTYACHGCYATGGTTGA-3'
COD C3'H For Nested	5'-GAATGGGCNATGGCVGA-3'
COD C3'H Rev	5'-ATATCRTARCCHCCRAT-3'
Expr C3'H For	5'-ATGACCCTCCTACTCCTCCCC-3'
Expr C3'H Rev	5'-TTACACATCCACGGCCACACG-3'
C3'H 447 Outer For	5'-AGTTGTTTTCTCCAAGAGGCTTGAGGC-3'
C3'H 447 Outer Rev	5'-AGTAATGGATGGGTACCTACCCAACCC-3'
C3'H 447 Inner For	5'-GAACGGAATGAGGATCAAACGTAGTTATCG-3'
C3'H 447 Inner Rev	5'-GAACGGAATGAGGATCAAACGTAGTTATCG-3'
RT C3'H For	5'-CTCTATCAGCGCCTCCGATT-3'
RT C3'H Rev	5'-ATATGATCGGGCCGTATTGC-3'
RT Act For	5'-TACTTTCTACAACGAGCTTC-3'
RT Act Rev	5'-ACATGATTTGAGTCATCTTC-3'

Heterologous expression of the C3'H gene in brewers' yeast

The isolated globe artichoke C3'H gene was amplified from cDNA template with primers Expr C3'H For and Expr C3'H Rev (Table 1), modified to introduce *Bgl*III and *Eco*RI cloning sites at the 5' and 3' ends. The amplicons were cloned into p-GEMTeasy vector and sequenced using ABI310 capillary sequencer. After *Bgl*III and *Eco*RI digestion, the fragments were directionally cloned into the expression cassette of pYeDP60 (Pompon et al. 1996) digested with *Bam*HI and *Eco*RI. *Saccharomyces cerevisiae* strain WAT11 is a derivative from strain W303-1B, in which yeast *CPR* gene (coding for a NADPH-cytochrome P450 reductase) was replaced by the *Arabidopsis ATR1* (Urban et al. 1997). WAT11 was transformed using the lithium acetate/single stranded carrier DNA/polyethylene glycol method (Gietz and Woods 2002). Transformants were selected on SGI plates (1 g/l bactocasaminoacids, 20 mg/l tryptophan, 6.7 g/l yeast nitrogen base, 20 g/l glucose, 20 g/l bactoagar) and successful transformation was confirmed by PCR analysis.

Enzymatic assay of the C3'H activities

Yeast microsomes were isolated after 24 h of induction on 20 g/l galactose at 30°C, according to Pompon et al. (1996). The amount of total protein in microsomes was determined according to Bradford protein assay. The substrate specificity of microsomal fractions, extracted from the recombinant yeast culture, was tested with *p*-coumaric acid, *p*-coumaroylshikimate and *p*-coumaroylquininate (kindly provided by Dr. Ullmann, Université Louis Pasteur, Strasbourg). The negative control was represented by the microsomal fraction of yeast transformed with an empty plasmid, and the positive control from yeast expressing CYP98A3 (ortholog of C3'H in *A. thaliana*, obtained from Arabidopsis Biological Resource Center Arabidopsis, Ohio State University). C3'H enzymatic activity was measured in 100 µl reaction tubes containing 600 µM NADPH, a range of *p*-coumaroylshikimate concentrations (0, 2, 5, 8, 10, 15, 20, 40, 60, 80, 100 and 150 µM), 100 mM sodium phosphate buffer (pH 7.4) and 30 µg of total microsomal proteins. The same conditions were applied with *p*-coumaroylquininate or *p*-coumaric acid as substrate. After starting the reaction upon addition of the enzyme, the tubes were shaken in the dark at 30°C for 30 min, and the reaction was stopped by adding 100 µl of trifluoroacetic acid:methanol (1:1,000 v/v). The reaction products were filtered through a 0.45 µM Anotop 10 filter (Whatman) and immediately analyzed by HPLC with photodiode array (PDA) detection, as described below.

HPLC analyses

Enzymatic assays were analyzed by reverse-phase HPLC with PDA detection, as described recently (Moglia et al. 2008). In short, the HPLC system comprised a Waters 600 gradient controller, a Waters 996 PDA detector and a column incubator held at 40°C. For the chromatographic separation an analytical column Luna C18 (2) (2 mm × 150 mm, 100 Å, particle size 3 µM) with a pre-column (2 mm × 4 mm) from Phenomenex was used. The mobile phase consisted of degassed trifluoroacetic acid:ultrapure water (1:1,000 v/v, eluate A), and trifluoroacetic acid:acetonitrile (1:1,000 v/v, eluate B), starting at 5% B, 95% A and increasing linearly to 35% B, 65% A over 45 min. The flow rate was 1 ml/min, the injection volume 10 µl, and the range of detection wavelength was 240–600 nm. Peak areas were calculated using Empower software (Waters).

UV-C treatment, and RT-qPCR assays

Three globe artichoke foliar discs were exposed to UV-C treatment (16 W germicidal lamp during 20 min) as described by Moglia et al. (2008). They were then ground separately to a fine powder in liquid nitrogen, and RNA extraction was performed from 100 mg of each powdered leaf, as described above. Primers (RT C3'H For, RT C3'H Rev, Table 1) were designed on C3'H sequence using the Primer 3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). As a housekeeping gene, actin was chosen for its stability and level of expression, which is comparable to the genes of interest and whose expression remained stable after the UV-C stress. The primers (RT Act For, RT Act Rev, Table 1) were designed on the globe artichoke actin (ACT, AM744951).

For the real time quantitative PCR (RT-qPCR), cDNA was first prepared in a 20 µl RT reaction containing 5× iScript Reaction Mix (Bio-Rad), 1 µl iScript Reverse Transcriptase and 1 µg total RNA. PCR primers were designed to detect globe artichoke C3'H and actin (Table 1). The cDNA was diluted to obtain a threshold cycle (CT) value between 25 and 35. The 20 µl RT-qPCRs, performed in triplicate for each sample, contained 1× iQ Supermix, 1× SYBR-Green I (iQTM, SYBR® Green Supermix), 10 µM primer and 3 µl diluted cDNA. 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 35 cycles of 95°C/15 s and 60°C/60 s. In all experiments, appropriate negative controls containing no template were subjected to the same procedure to detect or exclude any possible contamination. Melting curve analysis was performed at the end of amplification.

Standard curves were analyzed using iCycler iQ software. Amplicons were analyzed by the comparative threshold cycle method, in which $\Delta\Delta C_t$ is calculated as $\Delta C_{tI} - \Delta C_{tM}$, where ΔC_{tI} is the C_t value for the any target gene normalized to the endogenous housekeeping gene and ΔC_{tM} is the C_t value for the calibrator, which is also normalized to housekeeping gene.

SNP detection and linkage analysis

Sequence variation in the C3'H gene was sought by comparing the copy present in the non spiny globe artichoke variety 'Romanesco C3' with the spiny type 'Spinoso di Palermo'. Parental genomic DNAs were amplified with Expr C3'H For and Exp C3'H Rev (Table 1), and the amplicons were directly sequenced to facilitate SNP mining.

SNPs genotyping was carried out with the tetra-primers ARMS-PCR method (Ye et al. 2001) by using two sets of outer and inner primers (Table 1), designed using the software made available on-line (http://cedar.genetics.soton.ac.uk/public_html/primer1.html). PCR products were separated by 2% agarose gel electrophoresis. The segregation data obtained in an established mapping population (Lanteri et al. 2006) were combined with those associated with the markers used to construct the map and used to locate the map position of C3'H gene.

Segregation data of C3'H-SNP marker were monitored and analyzed in the 94 individuals of F1 progeny together with those of 35 AFLP, 41 S-SAP, 38 M-AFLP and 51 SSR markers previously applied for globe artichoke map construction (Lanteri et al. 2006). The goodness of fit between observed and expected segregation data was assessed using the Chi-square (χ^2) test. Independent linkage maps were constructed for each parent using the double pseudo-test-cross mapping strategy (Weeden 1994) by using JoinMap 2.0 software (Stam and Van Ooijen 1995). For both maps, linkage groups were accepted at a LOD threshold of 4.0. To determine marker order within a linkage group, the following JoinMap parameter settings were used: Rec = 0.40, LOD = 1.0, Jump = 5. Map distances were converted to centiMorgans using the Kosambi mapping function (Kosambi 1944). Linkage groups were drawn using MapChart 2.1 software (Voorrips 2002).

Results

Gene and promoter isolation

To isolate globe artichoke C3'H coding sequences, degenerated primers were used to amplify an 800 bp fragment from leaf genomic DNA. Only one amplified

fragment was obtained and showed (Blastx) high amino acid similarity (>90%) with members of the CYP98A family. The original 800 bp genomic sequence included one 171 bp intron.

After RACE-PCR, the sequence was extended to 1,524 bp, representing a 508 residue protein (CYP98A49, accession number FJ225121). The globe artichoke CYP98A49 gene contains four expected P450 conserved domains: the proline rich membrane hinge (PPGP), the I-helix involved in oxygen binding and activation (A/G-G-X-E/D-T-T/S), the clade signature (PERF) and the cysteine-containing region (PFGXGRRXCX) (Fig. 2).

A phylogenetic tree was constructed using 30 sequences from CYP98 family (Fig. 3). Most of the accessions are included in two main clusters with high bootstrap probability: one comprises genes derived from Monocots species and the other one derived from Dicots and a Gymnosperm (*P. taeda*) species. This division into two main groups (as previously observed by Morant et al. 2007) indicates that the encoding genes resulted from early duplication of ancestral gene. As opposite CYP98A8 and CYP98A9, two other CYP98A members from *A. thaliana*, and CYP98A20 are well separated from the main clusters. Globe artichoke CYP98A49 grouped in a sub-cluster of five proteins, and seemed closer to CYP98A46 from *Coptis*, CYP98A27 from *Populus*, CYP98A37 from *Medicago* and CYP98A2 from *Glycine*.

In addition to the coding region, we searched the 2-kb sequence upstream of the ORF (FJ225121). The CYP98A49 putative transcription start site was 24 bp upstream of the ATG start codon. Upstream of this transcription start site are a putative TATA box (−32 bp) and a putative CAAT box (−145 bp). The promoter contains the following regulatory elements: recognition sites of MYB, MYC, DRE-core, W-Box, TGA Box, P-Box, BoxIV and WRE1 (wound responsive element 1). All these elements, except WRE1, were also detected in the promoter region of *Arabidopsis* CYP98A3 (Fig. 4).

In vitro enzymatic activity

In order to test its enzymatic activity, the CYP98A49 protein from globe artichoke was expressed in yeast. The microsomal fraction was subsequently isolated and tested for enzymatic activity in vitro using several substrates at different concentrations. In the presence of *p*-coumaroylshikimate (Fig. 5a), the recombinant protein synthesized a compound that could be identified as caffeoylshikimate. The enzymatic reaction was completely dependent on the presence of NADPH. The same product was generated by microsomes from yeast expressing *A. thaliana* CYP98A3 (Fig. 5b, positive control), but not by yeast expressing the empty vector (Fig. 5c, negative control).

Fig. 3 Neighbor-joining tree phylogenetic analysis of globe artichoke CYP98A49. The length of the lines indicates the relative distance between nodes. The list of cytochrome p450s updated by Dr. Nelson at <http://drnelson.utmem.edu/CytochromeP450.html> was the starting point for the analysis

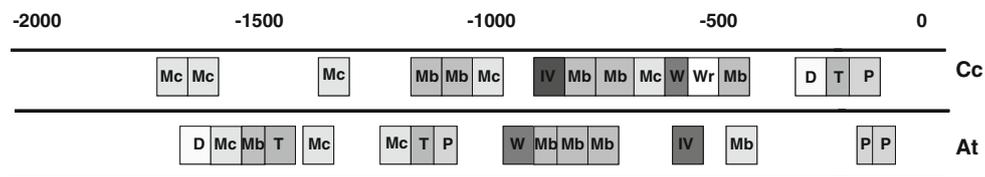
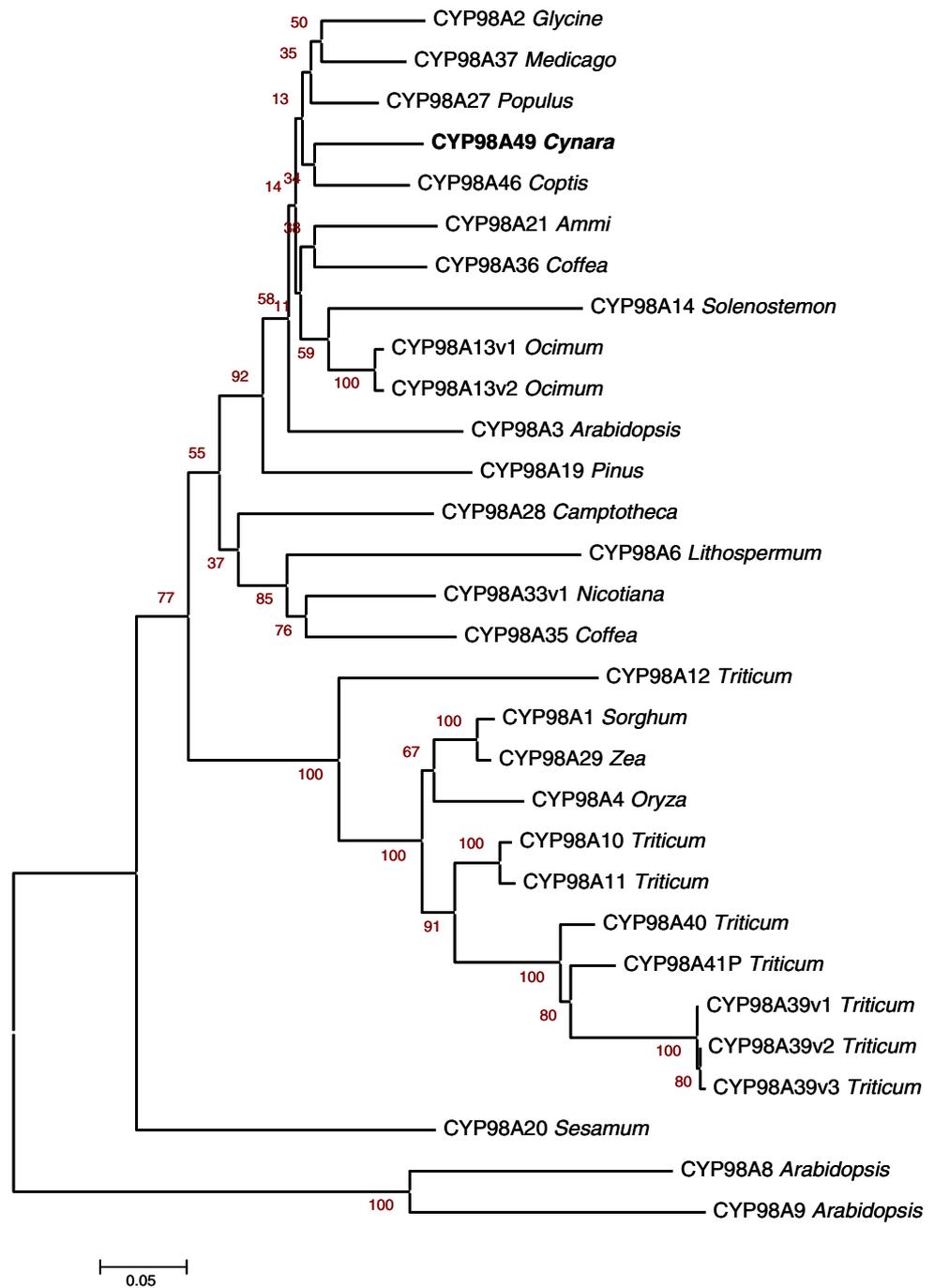


Fig. 4 The localization of *cis*-acting elements in the promoter region of CYP98A3 of *Arabidopsis* (*At*) and of globe artichoke CYP98A49 (*Cc*). In particular are shown recognition sites of MYB (*Mb*), MYC

(*Mc*), DRE-core (*D*), W-box (*W*), TGA box (*T*), P-Box (*P*), BoxIV (*IV*), WRE-1 (*Wr*)

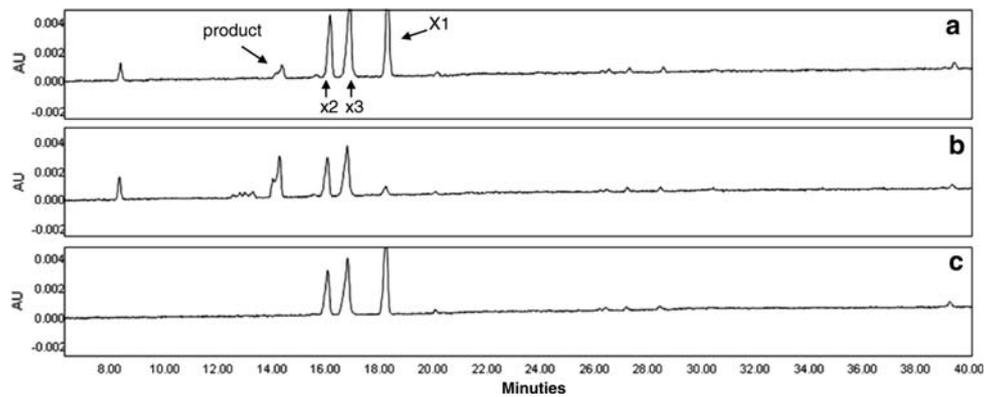


Fig. 5 HPLC-PDA analysis of reaction products of yeast microsomes containing globe artichoke CYP98A49 (a) or *A. thaliana* CYP98A3 (b) using *p*-coumaroylshikimate as a substrate (50 μ M). Control reactions were derived from yeast transformed with the empty

plasmid (c). Chromatograms are recorded at 312 nm. Peak *X1* is the substrate; while the other peaks close to it (*X2*, *X3*) represent other isomers (3 and 4) of *p*-coumaroylshikimate. The injection volume was 10 μ l

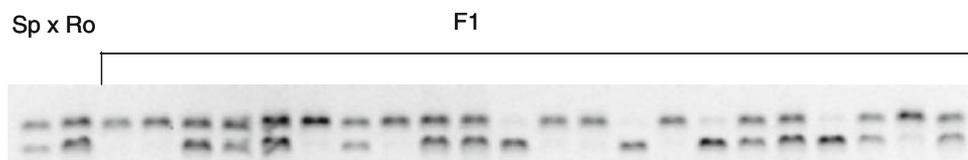


Fig. 6 Single nucleotide polymorphism segregation in a mapping population, as detected by tetra-primers ARMS-PCR on agarose gel. Romanesco C3 (*Ro*) and Spinoso di Palermo (*Sp*)

identification of one nucleotide difference at position 447 (data not shown). The tetra primers ARMS-PCR assay demonstrates that both parents are heterozygous at this base position. The C3'H_{snp447} locus segregated in a 1:2:1 ratio ($\chi^2 = 2.80$, $P > 0.1$) in the mapping population (Fig. 6), and mapped to linkage group 10 in both maps, ~ 2 cM from the microsatellite locus CELMS-39 and 8 cM from the AFLP locus p45/m47-07 (Fig. 7). Twenty-one markers were assigned to the female LG 10: four microsatellites (CELMS-04, -20, -39 and CLIB-04), two S-SAP (cyre5 markers), 2 M-AFLP (polyGA markers) and 12 AFLP together with SNP-C3'H, covering 84.4 cM and a mean inter-marker distance of 4.22 cM. The majority of map interval (70%) was < 5 cM and three gaps of 8 cM were present. The male LG 10 was composed of 18 markers: two microsatellite (CELMS-04 and -39) one S-SAP, 14 AFLP and the SNP-C3'H, spanned 99.4 cM with a marker density of 5.84 cM. Three large caps longer than 12 cM were detected. Nine intercross markers (comprising C3'H gene) were shared between the parents, allowing the alignment of the maternal and paternal LG 10 (Fig. 7). Estimation of markers order and distance of LG10 were improved with the integration of the co-dominant SNP-C3'H markers, increasing the number of bridge markers. The relative orders of some markers previously

determined (Lanteri et al. 2006) are slightly changed, as two inversions and a small shift of a few centimorgans were detected (Fig. 7).

Discussion

The P450 proteins form a large family of enzymes involved in plant metabolism, but the function of about 80% of them remains unknown. The *A. thaliana* genome includes 273 cytochrome P450 genes distributed in 45 families and sub-families (<http://drnelson.utmem.edu/Arablinks.html>).

Globe artichoke CYP98A49 sequence is highly homologous to some of the other CYP98 genes (up to 86% of identity and 93% of similarity to CYP98A46 from *C. japonica*), and contains the conserved domains associated with P450 family members (Fig. 2). The 3'-hydroxylation step is critical in the synthesis of phenolic compounds. Most of the members of the CYP98 family, described to date, metabolize shikimate esters of *p*-coumaric acid more efficiently than quinate esters (Schoch et al. 2001, 2006; Morant et al. 2007). On the other hand CYP98A35 from coffee is capable of metabolizing *p*-coumaroylquininate and *p*-coumaroylshikimate with the same efficiency (Mahesh et al. 2007). CYP98A49 from globe artichoke appears to

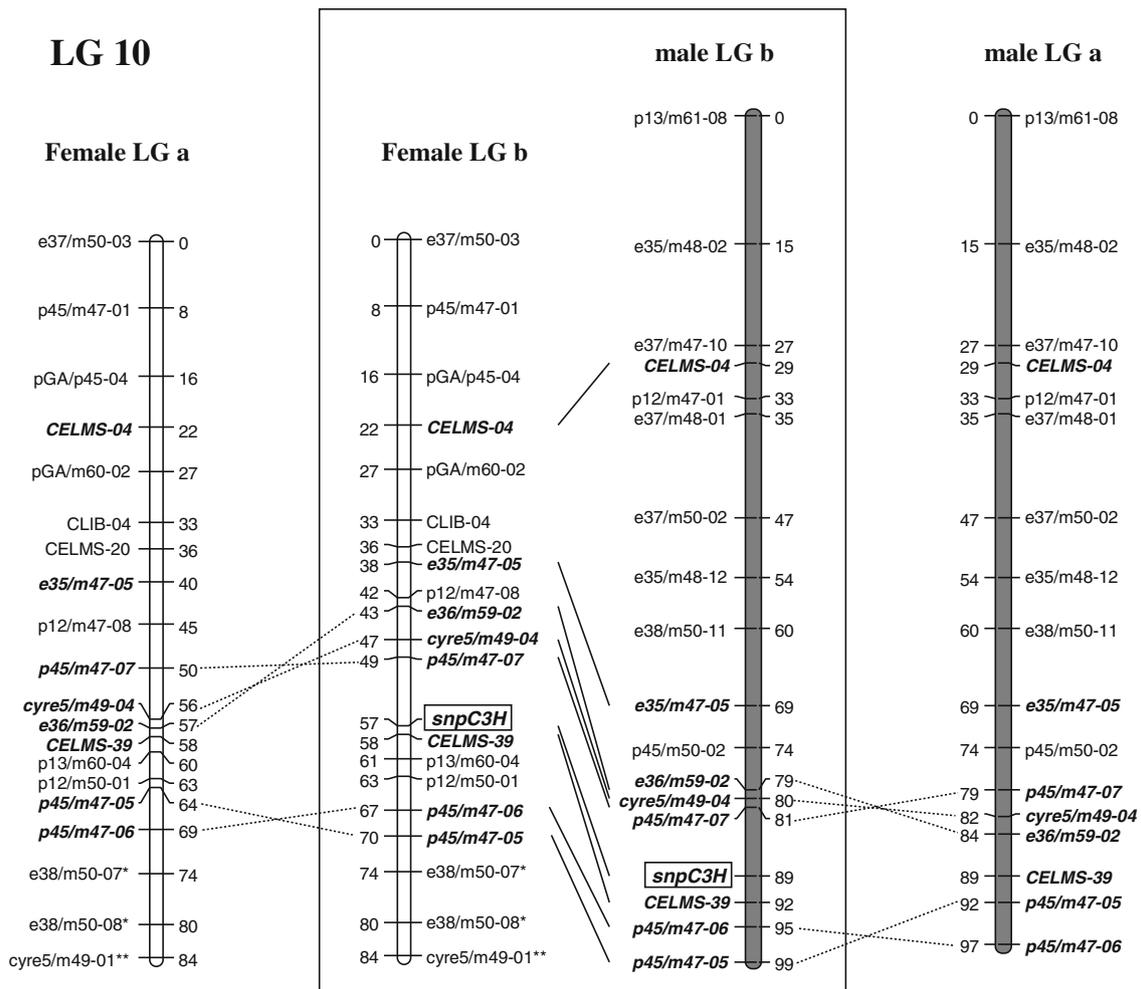


Fig. 7 Linkage group (LG) 10 of the globe artichoke varietal types ‘Romanesco C3’ (female parent, white LGs on the left) and ‘Spinoso di Palermo’ (male parent, gray LGs on the right). Intercross markers are shown in *bold* and are connected by a *solid line*. The LGs

previously reported by Lanteri et al. (2006) are presented to one side, and changed marker orders are indicated by *dotted lines*. *Asterisks* indicate markers showing significant levels of segregation distortion (*single asterisk* $0.1 > P \geq 0.05$, *double asterisk* $0.05 > P \geq 0.01$)

show a lower affinity for quinate esters than shikimate esters (Fig. 5), but the limited activity detected for shikimate esters hampered an accurate evaluation of enzymatic activity with quinate esters.

Shikimate esters are transient intermediates in the formation of more oxygenated compounds such as lignin precursors and chlorogenic acid. Two distinct pathways have been proposed for the synthesis of chlorogenic acid: (1) pathway from *p*-coumaroyl-CoA, involving first a transesterification of this compound and quinic acid via hydroxycinnamoyl-CoA:quinic acid hydroxycinnamoyl transferase (HQT) activity and then hydroxylation of *p*-coumaroylquinic acid to 5-caffeoylquinic acid, catalyzed by C3’H; (2) pathway involving *p*-coumaroyl-CoA transesterification with shikimic acid by means of hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase (HCT), then *p*-coumaroylshikimate hydroxylation to caffeoylshikimate.

This compound is then converted by HCT (Hoffmann et al. 2003; Comino et al. 2007) to caffeoyl-CoA which is conjugated to quinic acid by HQT (Niggeweg et al. 2004) activity, to give 5-caffeoylquinic acid.

It has been long debated whether the HQT enzyme either directly acts on caffeoyl-CoA and quinic acid to produce chlorogenic acid, or whether it synthesizes *p*-coumaroylquinic acid from *p*-coumaroyl-CoA and quinic acid, which is converted to chlorogenic acid by C3’H. Strong support for the first alternative has been provided in tomato, in which silencing of the HQT gene resulted in a 98% reduction in the level of chlorogenic acid (Niggeweg et al. 2004).

The gene encoding HCT in globe artichoke has recently been isolated (Comino et al. 2007). It is not clear, in this stage, if there are other C3’H genes and for this reason is not yet possible to conclude what is the route involved in the synthesis of chlorogenic acid in globe artichoke.

Some CYP98 genes are expressed constitutively, while others (particularly those involved in the stress response) are inducible. *A. thaliana* CYP98A3 is constitutively expressed and is upregulated by wounding (Schoch et al. 2001). *Phaseolus vulgaris* CYP98A5 is inducible by treatment with either 3,5-dichlorosalicylic acid or 2,6-dichloroisonicotinic acid (Basson and Dubery 2007), while the activity of *Daucus carota* 5-O-(4-coumaroyl)-D-quinic/shikimate 3'-hydroxylase could be greatly increased by irradiation with blue/UV light (Kühnl et al. 1987). In order to evaluate the changes of CYP98A49 expression levels in response to UV-C induction, we performed RT-PCR experiments. We have shown in a previous work (Moglia et al. 2008) that in globe artichoke the production of dicaffeoylquinic acids, which are powerful antioxidants, can be induced upon exposure to UV-C, which suggests a role for these compounds in the protection of young leaf tissue from reactive oxygen species generated by excess light. The expression level of CYP98A49 gene was strongly increased (greater than fourfold higher) in UV-C treated leaves, as compared to non-treated control leaves, thus indicating not only activation in response to UV light, but likely also a putative role of this enzyme in dicaffeoylquinic acid accumulation.

Transcription factors are important in the regulation of plant responses to environmental stresses. Most of cytochrome P450 genes induced by abiotic and biotic stresses contain the recognition sites of MYB, MYC, TGA-box and W-box for WRKY factors in their promoters (Narusaka et al. 2004). The sequence analysis of the upstream region of globe artichoke CYP98A49 gene revealed the presence of most of these regulatory regions (Fig. 4). Moreover, the same kind of motifs was found in the promoter of *Arabidopsis* CYP98A3 (not tested for UV response), which is homologous to the globe artichoke CYP98A49 gene. In a previous work (Narusaka et al. 2004) the distribution of *cis*-acting elements in the regulatory region of P450 *Arabidopsis* genes activated in response to UV-C radiation, was analyzed. Interestingly, all the UV-induced P450s in *Arabidopsis* share with the globe artichoke CYP98A49 promoter the recognition sites of MYB, MYC and the binding site of WRKY factors (W-box). Therefore, it is possible that the *cis*-acting elements recognized by MYB, MYC and WRKY transcription factors may regulate the expression of genes induced upon UV-C.

The identification of the genetic basis of metabolite variation in *A. thaliana* has been pioneered by Keurentjes et al. (2006), by applying quantitative trait loci (QTL) analyses on a large metabolomics data set. This approach, if applied to crop species, may lead to the development of informative genetic markers that could be exploited in breeding programs aimed at increasing the level of specific phytochemicals. The *C. cardunculus* genome is still poorly

mapped. In order to move to a crossing strategy for breeding, a greater knowledge of globe artichoke genome will be essential. In particular it will be advantageous to establish a framework of linkage relationships for reaching a better knowledge of the genetic bases of the phenylpropanoid pathway. The linkage relationships we established for the globe artichoke CYP98A49 gene may thus represent an initial step in this direction (Fig. 7). The precision of both marker order and inter-marker distances of LG10 has been improved with the integration of CYP98A49 gene.

Future efforts of our research will go in the direction of studying the role of the CYP98A49 gene in globe artichoke development, by means of forward genetic approaches, and in the identification of QTLs associated with the production of phenolic compounds such as chlorogenic acid and dicaffeoylquinic acids. Indeed we are proceeding to the construction of genetic maps based on F1 populations involving combinations between 'Romanesco clone C3' with either cultivated as wild cardoon accessions; these populations will allow comparative QTL mapping studies.

Acknowledgments The authors kindly thank Harry Jonker (PRI) for his excellent technical assistance. The authors kindly thank Dr. Ullmann (Université Louis Pasteur, Strasbourg) for providing the substrates of the reaction, *p*-coumaroylquinic and *p*-coumaroylshikimate and for critical reading of the paper. The authors kindly thank Dr Nelson for providing CYP number to the new gene. The authors kindly thank Prof G. Mauromicale and Dr R. Mauro for providing plant material. Andrea Moglia acknowledges MIUR, for its financial support. Jules Beekwilder was financially supported by the EU 6th Frame FLORA project (2005-FOOD-CT-01730). Ric C. H. De Vos acknowledges initial support from the Centre for BioSystems Genomics, an initiative under the auspices of the Netherlands Genomics Initiative (NGI/NWO).

References

- Abdulrazzak N, Pollet B, Ehling J, Larsen K, Asnaghi C, Ronseau S, Proux C, Erhardt M, Seltzer V, Renou J, Ullman P, Pauly M, Lapierre C, Werck-Reichhart D (2006) A coumaroyl-ester-3-hydroxylase insertion mutant reveals the existence of non redundant meta-hydroxylation pathways and essential roles for phenolic precursors in cell expansion and plant growth. *Plant Physiol* 140:30–48. doi:10.1104/pp.105.069690
- Adzet T, Camarassa J, Laguna CJ (1987) Hepatoprotective activity of polyphenolic compounds from *Cynara scolymus* against CCl₄ toxicity in isolated rat hepatocytes. *J Nat Prod* 50:612–617
- Basson A, Dubery I (2007) Identification of a cytochrome P450 cDNA (CYP98A5) from *Phaseolus vulgaris*, inducible by 3, 5-dichlorosalicylic acid and 2, 6-dichloro isonicotinic acid. *J Plant Physiol* 164:421–428. doi:10.1016/j.jplph.2006.02.006
- Boudet A (2007) Evolution and current status of research in phenolic compounds. *Phytochemistry* 68:2722–2735. doi:10.1016/j.phytochem.2007.06.012
- Brown J, Rice-Evans C (1998) Luteolin-rich artichoke extract protects low density lipoprotein from oxidation in vitro. *Free Radic Res* 29:247–255. doi:10.1080/10715769800300281
- Cantos E, Espin J, Tomas-Barberan F (2001) Effect of wounding on phenolic enzymes in six minimally processed lettuce cultivars

- upon storage. *J Agric Food Chem* 49:322–330. doi:[10.1021/jf000644](https://doi.org/10.1021/jf000644)
- Chen F, Dixon R (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nat Biotechnol* 25:759–761. doi:[10.1038/nbt1316](https://doi.org/10.1038/nbt1316)
- Comino C, Lanteri S, Portis E, Acquadro A, Romani A, Hehn A, Larbat R, Bourgaud F (2007) Isolation and functional characterization of a cDNA coding a hydroxycinnamoyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus* L. *BMC Plant Biol* 7:14. doi:[10.1186/1471-2229-7-14](https://doi.org/10.1186/1471-2229-7-14)
- Comino C, Hehn A, Moglia A, Menin B, Bourgaud F, Lanteri S, Portis E (2009) The isolation and mapping of a novel hydroxycinnamoyltransferase in the artichoke chlorogenic acid pathway. *BMC Plant Biol* 9
- Paolis De, Pignone D, Morgese A, Sonnante G (2008) Characterization and differential expression analysis of artichoke phenylalanine ammonia-lyase-coding sequences. *Physiol Plant* 132:33–43. doi:[10.1111/j.1399-3054.2007.00996.x](https://doi.org/10.1111/j.1399-3054.2007.00996.x)
- Dixon R, Paiva N (1995) Stress-induced phenylpropanoid metabolism. *Plant Cell* 7:1085–1097
- Douglas C (1996) Phenylpropanoid metabolism and lignin biosynthesis: from weeds to trees. *Trends Plant Sci* 1:171–178. doi:[10.1016/1360-1385\(96\)10019-4](https://doi.org/10.1016/1360-1385(96)10019-4)
- Franke R, Humphreys J, Hemm M, Denault J, Ruegger M, Cusumano J, Chapple C (2002) The *Arabidopsis* REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. *Plant J* 30:33–45. doi:[10.1046/j.1365-313X.2002.01266.x](https://doi.org/10.1046/j.1365-313X.2002.01266.x)
- Gebhardt R (1997) Antioxidative and protective properties of extracts from leaves of the artichoke (*Cynara scolymus* L.) against hydroperoxide-induced oxidative stress in cultured rat hepatocytes. *Toxicol Appl Pharmacol* 144:279–286. doi:[10.1006/taap.1997.8130](https://doi.org/10.1006/taap.1997.8130)
- Gebhardt R (1998) Inhibition of cholesterol biosynthesis in primary cultured rat hepatocytes by artichoke (*Cynara scolymus* L.) extracts. *J Pharmacol Exp Ther* 286:1122–1128
- Gietz R, Woods R (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* 350:87–96
- Hoffmann L, Maury S, Martz F, Geoffroy P, Legrand M (2003) Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *J Biol Chem* 278:95–103. doi:[10.1074/jbc.M209362200](https://doi.org/10.1074/jbc.M209362200)
- Izaguirre M, Mazza C, Svatos A, Baldwin I, Ballare C (2007) Solar ultraviolet-B radiation and insect herbivory trigger partially overlapping phenolic responses in *Nicotiana attenuata* and *Nicotiana longiflora*. *Ann Bot* 99:103–109. doi:[10.1093/aob/mcl226](https://doi.org/10.1093/aob/mcl226)
- Keurentjes J, Fu J, de Vos C, Lommen A, Hall R, Bino R, van der Plas L, Jansen R, Vreugdenhil D, Koornneef M (2006) The genetics of plant metabolism. *Nat genet* 38:842–849. doi:[10.1038/ng0706-737](https://doi.org/10.1038/ng0706-737)
- Kosambi D (1944) The estimation of map distances from recombination values. *Ann Eugen* 12:172–175
- Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* 5:150–163
- Kühnl T, Koch U, Heller W, Wellmann E (1987) Chlorogenic acid biosynthesis: characterization of a light-induced microsomal 5-O-(4-coumaroyl)-D-quinic/shikimate 3'-hydroxylase from carrot (*Daucus carota* L.) cell suspension cultures. *Arch Biochem Biophys* 258:226–232. doi:[10.1016/0003-9861\(87\)90339-0](https://doi.org/10.1016/0003-9861(87)90339-0)
- Lanteri S, di Leo I, Ledda L, Mamei M, Portis E (2001) RAPD variation within and among populations of globe artichoke cultivar “Spinoso sardo”. *Plant Breed* 120:243–246. doi:[10.1111/j.1439-0523.2001.tb01994.x](https://doi.org/10.1111/j.1439-0523.2001.tb01994.x)
- Lanteri S, Acquadro A, Comino C, Mauro R, Mauromicale G, Portis E (2006) A first linkage map of globe artichoke (*Cynara cardunculus* var. *scolymus* L.) based on AFLP, S-SAP, M-AFLP and microsatellite markers. *Theor Appl Genet* 112:1532–1542. doi:[10.1007/s00122-006-0256-8](https://doi.org/10.1007/s00122-006-0256-8)
- Lattanzio V, Cardinali A, Di Venere D, Linsalata V, Palmieri S (1994) Browning phenomena in stored artichoke (*Cynara scolymus* L.) heads: enzymic or chemical reactions? *Food Chem* 50:1–7
- Mahesh V, Million-Rousseau R, Ullmann P, Chabrilange N, Bustamante J, Mondolot L, Morant M, Noirot M, Hamon S, de Kochko A, Werck-Reichhart D, Campa C (2007) Functional characterization of two p-coumaroyl ester 3'-hydroxylase genes from coffee tree: evidence of a candidate for chlorogenic acid biosynthesis. *Plant Mol Biol* 64:145–159. doi:[10.1007/s11103-007-9141-3](https://doi.org/10.1007/s11103-007-9141-3)
- Moglia A, Lanteri S, Comino C, Acquadro A, de Vos R, Beekwilder J (2008) Stress-induced biosynthesis of dicaffeoylquinic acids in globe artichoke. *J Agric Food Chem* 56:8641–8649. doi:[10.1021/jf801653w](https://doi.org/10.1021/jf801653w)
- Morant M, Hehn A, Werck-Reichhart D (2002) Conservation and diversity of gene families explored using the CODEHOP strategy in higher plants. *BMC Plant Biol* 2:7. doi:[10.1186/1471-2229-2-7](https://doi.org/10.1186/1471-2229-2-7)
- Morant M, Schoch GA, Ullmann P, Ertunç T, Little D, Olsen CE, Petersen M, Negrel J, Werck-Reichhart D (2007) Catalytic activity, duplication and evolution of the CYP98 cytochrome P450 family in wheat. *Plant Mol Biol* 63:1–19. doi:[10.1007/s11103-006-9028-8](https://doi.org/10.1007/s11103-006-9028-8)
- Narusaka M, Narusaka Y, Seki M, Umezawa T, Ishida J, Nakajima M, Enju A, Shinozaki K (2004) Crosstalk in the responses to biotic and abiotic stresses in *Arabidopsis*: analysis of gene expression in *cytochrome* P450 gene superfamily by cDNA microarray. *Plant Mol Biol* 55:327–342. doi:[10.1007/s11103-004-0685-1](https://doi.org/10.1007/s11103-004-0685-1)
- Niggeweg R, Michael A, Martin C (2004) Engineering plants with increased levels of the antioxidant chlorogenic acid. *Nat Biotechnol* 22:746–754. doi:[10.1038/nbt966](https://doi.org/10.1038/nbt966)
- Pompon D, Louerat B, Bronine A, Urban P (1996) Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol* 272:51–64
- Schoch G, Goepfert S, Morant M, Hehn A, Meyer D, Ullmann P, Werck-Reichhart D (2001) CYP98A3 from *Arabidopsis thaliana* is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. *J Biol Chem* 276:36566–36574. doi:[10.1074/jbc.M104047200](https://doi.org/10.1074/jbc.M104047200)
- Schoch G, Morant M, Abdulrazzak N, Asnaghi C, Goepfert S, Petersen M, Ullmann P, Werck-Reichhart D (2006) The meta-hydroxylation step in the phenylpropanoid pathway: a new level of complexity in the pathway and its regulation. *Environ Chem Lett* 4:127–136. doi:[10.1007/s10311-006-0062-1](https://doi.org/10.1007/s10311-006-0062-1)
- Slanina J, Taborska E, Bochorakova H, Slaninova I, Humpa O, Robinson W, Schram K (2001) New and facile method of preparation of the anti-HIV-1 agent, 1, 3-dicaffeoylquinic acid. *Tetrahedron Lett* 42:3383–3385. doi:[10.1016/S0040-4039\(01\)00448-8](https://doi.org/10.1016/S0040-4039(01)00448-8)
- Stam P, Van Ooijen J (1995) JoinMap version 2.0: software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen
- Treutter D (2005) Significance of flavonoids in plant resistance and enhancement of their biosynthesis. *Plant Biol* 7:581–591. doi:[10.1055/s-2005-873009](https://doi.org/10.1055/s-2005-873009)
- Urban P, Mignotte C, Kazmaier M, Delorme F, Pompon D (1997) Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. *J Biol Chem* 272:19176–19186

- Voorrips R (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered* 93:77–78
- Wang M, Simon J, Aviles I, He K, Zheng Q, Tadmor Y (2003) Analysis of antioxidative phenolic compounds in artichoke (*Cynara scolymus* L.). *J Agric Food Chem* 51:601–608. doi: [10.1021/jf020792bS0021-8561\(02\)00792-6](https://doi.org/10.1021/jf020792bS0021-8561(02)00792-6)
- Weeden N (1994) Approaches to mapping in horticultural crops. In: Gresshoff PM (ed) *Plant genome analysis*. CRC, Boca Raton, pp 51–60
- Ye S, Dhillon S, Ke X, Collins A, Day I (2001) An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acid Res* 29:E88–8