Genomics 96 (2010) 239–250

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Genomic analysis and gene structure of the plant carotenoid dioxygenase 4 family: A deeper study in *Crocus sativus* and its allies

Oussama Ahrazem [a], Almudena Trapero [a], Mª Dolores Gómez [b], Angela Rubio-Moraga [a], Lourdes Gómez-Gómez [a,*]

[a] Departamento de Ciencia y Tecnología Agroforestal y Genética, Escuela Técnica Superior de Ingenieros Agrónomos, Universidad de Castilla-La Mancha, Campus Universitario s/n, 02071 Albacete, Spain

[b] Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain

ARTICLE INFO

Article history:
Received 4 June 2010
Accepted 7 July 2010
Available online 12 July 2010

Keywords:
Crocus sativus
Carotenoids
Intron
Promoter
Stigma
Volatiles

ABSTRACT

The plastoglobule-targeted enzyme carotenoid cleavage dioxygenase (CCD4) mediates the formation of volatile C_{13} ketones, such as β-ionone, by cleaving the C_9-C_{10} and $C_{9'}$-$C_{10'}$ double bonds of cyclic carotenoids. Here, we report the isolation and analysis of CCD4 genomic DNA regions in *Crocus sativus*. Different *CCD4* alleles have been identified: *CsCCD4a* which is found with and without an intron and *CsCCD4b* that showed the presence of a unique intron. The presence of different *CCD4* alleles was also observed in other *Crocus* species. Furthermore, comparison of the locations of *CCD4* introns within the coding region with *CCD4* genes from other plant species suggests that independent gain/losses have occurred. The comparison of the promoter region of *CsCCD4a* and *CsCCD4b* with available *CCD4* gene promoters from other plant species highlighted the conservation of *cis*-elements involved in light response, heat stress, as well as the absence and unique presence of *cis*-elements involved in circadian regulation and low temperature responses, respectively. Functional characterization of the *Crocus sativus CCD4a* promoter using *Arabidopsis* plants stably transformed with a DNA fragment of 1400 base pairs (P-CsCCD4a) fused to the β-glucuronidase (*GUS*) reporter gene showed that this sequence was sufficient to drive *GUS* expression in the flower, in particular high levels were detected in pollen.

Introduction

Carotenoids are isoprenoid pigments synthesized by all photosynthetic organisms and some nonphotosynthetic bacteria and fungi. In plants, carotenoids are essential in protecting the photosynthetic apparatus from photo-oxidation, and represent essential constituents of the light-harvesting and reaction centre complexes. The oxidative cleavage of carotenoids occurs in plants, animals, and micro-organisms and leads to the release of a range of apocarotenoids that function as signalling molecules with diverse functions [1,2], including the ubiquitous chromophore retinal, plant hormone abscisic acid and strigolactones. Other apocarotenoids with unknown functions in plants but with high economic value are bixin in *Bixa orellana* and saffron in *Crocus sativus* [3], used as a flavouring and colouring agent. *Crocus sativus* is a triploid sterile plant, most probably originated by a human-mediated event, which is propagated by corms. Among all known *Crocus* species, *Crocus sativus* is particularly appreciated due to the high levels of apocarotenoids on the stigmas that determine the commercial value of this plant. Hence, studying apocarotenoid formation in all these species will help to determine the mechanisms underlying apocarotenoid accumulation.

The synthesis of apocarotenoids is initiated by the oxidative cleavage of double bonds in its stigmas carotenoid backbones, catalyzed by carotenoid oxygenases, which are nonheme iron enzymes present in all taxa [1,4]. The first gene identified as encoding a carotenoid cleavage dioxygenase was the *Vp14* maize gene required for the formation of abscisic acid (ABA) [5]. On the basis of their substrate specificity, VP14 and its orthologous have been termed 9-cis epoxy-carotenoid dioxygenases (NCEDs). Plants possess a second group of carotenoid oxygenases, carotenoid cleavage dioxygenases (CCDs), which act on different carotenoid substrates [1]. These include CCD1, CCD4, CCD7 and CCD8.

Plants release volatile apocarotenoids, including C_{13} ketones such as β-ionone and damascone, which constitute an essential aroma note in tea, grapes, roses, tobacco and wine. Two classes of CCD enzymes have been implicated in plant volatile production, CCD1 and CCD4. Both plant enzymes cleave C(40) carotenoids at the C_9-C_{10} and $C_{9'}$-$C_{10'}$ double bonds into C_{14} dialdehydes, which are common to all carotenoid substrates, and two variable end-group-derived C_{13} ketones [1]. The CCD1 enzymes act in the cytosol, where most probably are involved in apocarotenoid cleavage, whereas CCD4 enzymes have been shown to

Abbreviations: CCD, carotenoid cleavage dioxygenase; DSBs, double-stranded breaks; IME, intron-mediated; NCEDs, 9-cis-epoxycarotenoid dioxygenases; NHEJ, nonhomologous end joining.

* Corresponding author. Fax: +34 967599238.
E-mail address: Marialourdes.gomez@uclm.es (L. Gómez-Gómez).

0888-7543/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2010.07.003

reside in plastids, where their substrates are localized, suggesting a direct involvement in volatile formation. The first member of the CCD4 subfamily was identified in *Chrysantemum morifolium* [6] and the enzymatic activity has been recently characterized in *Crocus sativus* [7], *C. morifolium*, *Arabidopsis thaliana*, *Rosa damascena*, and *Malus domestica* [8].

The different *CCD* families are characterized by the heterogeneity in their gene structures, with genes containing no introns or containing as many as 10 introns, as in the human gene β, β-carotene-15,15-dioxygenase (*BCMO*) [9] or 11 and 13 introns for the case of the *CCD1* genes of rice and *Arabidopsis* [1], respectively. The plant group of *CCDs* containing *CCD7* and *CCD8* is characterized by the presence of multiple intron sequences in all the available genes. The *CCD7* genes from petunia, rice and pea are characterized by the presence of 6 introns [10,11], whereas 5 introns are present in the *CCD7* gene of *Arabidopsis* [12], while four of these introns are well conserved among these plant species [11]. The *CCD8* of *Arabidopsis* and petunia also contains 5 intron sequences. (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/av.cgi?db=ara&q=CCD8; [13]. By contrast, compared with the *CCD1* and the *CCD7/8* groups, the *CCD4* group seems to contain genes without intron sequences or with one or two introns [6,8,14].

The genomic structure of a gene determines its regulation. Gene regulation is mainly determined by the promoter region, but several other types of gene regulation, both positive and negative, involve plant introns. Some introns contain enhancer elements or alternative promoters, while many others elevate mRNA accumulation by a different process that has been named intron-mediated enhancement (IME), which is thought to result from synergistic interactions between the factors involved in the various steps of gene expression from transcription to translation [15]. Furthermore, conservation of exon–intron structure in clades of orthologous genes, as well as in families of paralogous genes and protein superfamilies, support the use of gene features as sources for phylogenetic inference [16]. Thus, the knowledge of the genomic structure is important in order to characterize gene families and for the establishment of evolutionary relationships.

In this report, we have isolated and compared the genomic structure of the *CCD4* genes of *Crocus sativus* and compared it with its allies and with other plant species, and determined the intron presence and conservation of intron arrangement within this *CCD* family. In addition, we have isolated and analysed the genomic DNA sequence upstream from the *CCD4* genes of *Crocus sativus*, which were analysed along with other *CCD4* promoters (*Arabidopsis*, rice, tomato, poplar, papaya, *Vitis*, *Medicago* and *Brassica*) in order to identify common cis-regulatory motifs and compare the distribution patterns of these motifs. The spatial and temporal activity of the putative *CsCCD4a* promoter from *Crocus sativus* fused to the β-glucuronidase (*GUS*) gene in stably transformed *Arabidopsis* plants was also assessed, and the localisation of GUS staining was monitored in different plant tissues.

Materials and methods

Plant material

For this study, we included 14 species of *Crocus* with a total of 37 different populations. Specimens were obtained from saffron growers in Tarazona de La-Mancha (Spain), from Dr. U. Jacobsen from the Agricultural University of Denmark and from private collections in the UK (Potterton Nursery) (Table 1). Plant tissues were independently harvested and frozen in liquid nitrogen and stored at −80 °C until required.

Seeds from *Arabidopsis* wild type Columbia ecotype (Col-0), and transgenic lines were sown in pots and watered with nutrient solution under a controlled environment with 16 h light/8 h dark cycles at 22 °C. Seeds from transformed *Arabidopsis* plants were surface sterilized by rinsing them in 70% (v/v) ethanol for 1 min, followed by a 15 min treatment in 10% (v/v) bleach + 0.05% (v/v) Triton X-100 and three rinses in sterile distilled water.

Isolation of CsCCD4a and CsCCD4b genomic clones

Genomic DNA samples were prepared from *Crocus* leaves by using a CTAB (hexadecyltrimethylammonium bromide) method. The isolated DNA was quantified and amplified via PCR in a BioRad MJ

Table 1
Crocus species and populations sampled for *CCD4a* and *CCD4b* genomic sequences: a+: *CCD4a* plus intron; b+: *CCD4b* plus intron; a−: *CCD4a* minus intron; b−: *CCD4b* minus intron.

Series	Nomenclature	Origin of plant material	*CCD4* alleles			
			a+	b+	a−	b−
Crocus	*Crocus sativus* L.	Spain, Tarazona	+	+	+	−
	Crocus sativus L.	China, Yunan (C 395)	+	+	+	−
	Crocus sativus L.	Greece, Kozani, Krokos	+	+	+	−
	Crocus sativus L.	Iran, Mashhad,	+	+	+	−
	Crocus sativus L	(UK)	+	+	+	−
	C. oreocreticus B.L. Burtt	(GKØ00-32) Greece, Creta, Mt Dikti	+	+	−	−
	C. oreocreticus B.L. Burtt	(G94-15) Greece, Creta, Rethimno	+	+	−	−
	C. cartwrightianus Bory and Chaub	(12629) Greece, Nomos Atiki, Plaka	+	+	−	−
	C. cartwrightianus Bory and Chaub	(GJLL 01-56) Greece, Nomos Atiki, Mt. Imitos	+	+	−	−
	C. cartwrightianus Bory and Chaub	(NJG 98-112) Greece, Nomos Atiki, Keratea	+	+	+	−
	C. cartwrightianus Bory and Chaub	(12627) Greece, Nomos Atiki, Ch. Taxiarchiso	+	+	−	−
	C. cartwrightianus Bory and Chaub	(GKØ00-8) Greece, Creta, Hania	+	+	+	−
	C. cartwrightianus Bory and Chaub	(GKØ00-44) Greece, Creta, Hania	+	+	+	−
	C. hadriaticus Herbert	(Ø01-124) Greece, Kefalonia, Mt Etnos	+	+	−	−
	C. hadriaticus Herbert	(J01-37) Greece, Peloponnese, W of Pelei	+	+	−	−
	C. hadriaticus Herbert	(ØJ01-32) Greece, Peloponnese, N of Vasilios	+	+	−	−
	C. hadriaticus Herbert	(GNJ01-141) Greece, Peloponnese, Nomos Arkadia	+	+	−	−
	C. hadriaticus Herbert	(ØJ01-40) Greece, Peloponnese, Mt Didima	++	+	−	−
	C. hadriaticus Herbert	(G98-31) Greece, Peloponnese, Mt Taigetos	++	+	−	−
	C. pallasii Goldb ssp pallasii	(GKØ01-64) Greece, Lesbos, above Agra	−	+	+	−
	C. pallasii Goldb ssp pallasii	(GKØ01-36) Greece, Lesbos, Mt Olymbos	+	+	−	−
	C. pallasii Goldb ssp pallasii	(GKØ01-45) Greece, Lesbos, Petra to Kalloni	−	++	−	−
	C. pallasii Goldb ssp pallasii	(GKØ01-61) Greece, Lesbos, above Agra	+	+	−	−
Longiflori	*C. medius* Balbis	(C 430)	−	−	+	−
	C. medius	(UK)	−	−	+	−
	C. goulimyi	(UK)	−	−	+	−
	C. goulimyi Turrill	(Ø01-12) Greece, Peloponnese, N Pirgos Dirou	−	−	+	−
	C. niveus Bowles	(00-154) Greece, Peloponnese, Nomos Arkadia	−	−	+	−
	C. niveus	(UK)	−	−	+	−
	C. nudiflorus	(UK)	−	−	+	−
Versicolores	*C. imperati*	(UK)	−	−	+	−
Kotschyani	*C. ochloroleucus*	(UK)	−	−	+	+
	C. kostcyanus	(UK)	−	−	+	−
Speciosi	*C. pulchellus* Herbert	(98-81) Greece, Mt Pangaeon	−	−	+	−
	C. pulchellus	(I-19) UK, Kew Gardens	−	−	+	−
	C. speciosus M.	(UK)	−	−	+	+
	C. speciosus M. Bieb	(C12) without origin	−	−	+	+

+: indicates the presence of one copy of that gene; ++: indicates the presence of two copies.